INTRODUCTION TO THE ARITHMETIC THEORY OF QUADRATIC FORMS

SAM RASKIN

Contents

1. Review of linear algebra and tensors

1.1. Linear algebra is assumed as a prerequisite to these notes. However, this section serves to review the language of abstract linear algebra (particularly tensor products) for use in the remainder of these notes.

We assume the reader familiar with these notions, and therefore do not provide a complete and detailed treatment. In particular, we assume the reader knows what a *homomorphism* of algebraic objects is.

1.2. **Rings and fields.** Recall that a *ring* is a set A equipped with associative binary operations $+$ and $\cdot$ called "addition" and "multiplication" respectively, such that $(k, +)$ is an abelian group with identity element 0, such that multiplication distributes over addition, and such that there is a multiplicative identity 1. The ring A is said to be *commutative* if its multiplication is a commutative operation. Homomorphisms of rings are maps are always assumed to preserve 1.

We say that a commutative ring A is a *field* if every non-zero element $a \in A$ admits a multiplicative inverse a^{-1}. We usually denote fields by k (this is derived from the German "Körper").

Example 1.2.1. For example, the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$ are all fields. The integers $\mathbb{Z}$ form a commutative ring but not a field. However, for p a prime number, the integers modulo p form a field that we denote by $\mathbb{F}_p$, and refer to as *the field with p elements.*

For A a ring, we let $A^\times$ denote the group of invertible elements in A.

Exercise 1.1. If k is a field and $G \subseteq k^\times$ is a finite subgroup, show that G is cyclic. Deduce that $\mathbb{F}_p^\times$ is a cyclic group of order $p - 1$.

Date: Spring, 2014.

1.3. **Characteristic.** For a ring A, we have a canonical (and unique) homomorphism $\mathbb{Z} \to A$:

$$n \mapsto \underbrace{1 + \ldots + 1}_{n \text{ times}} \in A$$

and we abuse notation in letting $n \in A$ denote the image of $n \in \mathbb{Z}$ under this map.

For a field k, this homomorphism has a kernel of the form $p \cdot \mathbb{Z}$ for p either a prime number or 0; we say that k has characteristic p accordingly. We see that k has characteristic 0 if $\mathbb{Q} \subseteq k$, and k has characteristic $p > 0$ if $\mathbb{F}_p \subseteq k$. We denote the characteristic of k as $\text{char}(k)$.

1.4. **Vector spaces.** Let k be a field fixed for the remainder of this section.

A *vector space* over k (alias: k-vector space) is an abelian group $(V, +)$ equipped with an action of k by scaling. That is, for $\lambda \in k$ and $v \in V$, we have a new element $\lambda \cdot v \in V$ such that this operation satisfies the obvious axioms. We frequently refer to elements of V as *vectors*.

We sometimes refer to homomorphisms of vector spaces sometimes as *linear transformations*, or as *k-linear maps*.

Remark 1.4.1. Note that the totality of linear transformations $T : V \to W$ can itself be organized into a vector space $\text{Hom}_k(V, W)$, also denoted $\text{Hom}(V, W)$ if the field k is unambiguous. The vector space operations are defined point-wise:

$$(T_1 + T_2)(v) := T_1(v) + T_2(v)$$
$$(\lambda \cdot T)(v) := \lambda \cdot (T(v)).$$

1.5. **Direct sums and bases.** We will primarily be interested in finite-dimensional vector spaces, but for the construction of tensor products it will be convenient to allow infinite direct sums.

Definition 1.5.1. Given two vector spaces V and W, we define their *direct sum* $V \oplus W$ to be the set-theoretic product $V \times W$ equipped with the structure of termwise addition and scalar multiplication.

More generally, given a set I and a collection $\{V_i\}_{i \in I}$ of vector spaces indexed by elements of I, we define the *direct sum* $\oplus_{i \in I} V_i$ as the subset of $\prod_{i \in I} V_i$ consisting of elements $v = (v_i)_{i \in I}$ $(v_i \in V_i)$ such that $v_i = 0$ for all but finitely many $i \in I$. Note that this does indeed generalize the pairwise example above.

Given $\{V_i\}_{i \in I}$ as above, let $\varepsilon_i : V_i \to \oplus_{j \in I} V_j$ denote the linear transformation sending $v_i \in V_i$ to the vector that is v_i in the ith coordinate and 0 in all others.

Proposition 1.5.2. *Given $\{V_i\}_{i \in I}$ as above and W an auxiliary vector space, the map:*

$$\begin{gathered} \{T : \oplus_{i \in I} V_i \to W \text{ a linear transformation}\} \to \{T_i : V_i \to W \text{ linear transformations}\}_{i \in I} \\ T \mapsto \{T \circ \varepsilon_i\}_{i \in I} \end{gathered} \tag{1.1}$$

is a bijection.

Proof. Suppose that we are given a collection $T_i : V_i \to W$ of linear transformations. Note that for every vector $v \in \oplus_{i \in I} V_i$, there exists a unique finite subset $S \subseteq I$ and collection of non-zero vectors $\{v_s \in V_s\}_{s \in S}$ such that:

$$v = \sum_{s \in S} \varepsilon_s(v_s)$$

(if S is empty, we understand the empty sum to mean the zero vector). We then define:

$$T(v) = \sum_{s \in S} T_s(v_s) \in W.$$

It is easy to verify that T is a linear transformation, and that this operation is inverse to the map (1.1).

□

Remark 1.5.3. The above proposition characterizes the direct sum by a *universal property*. This means that we have characterized the vector space $\oplus_{i \in I} V_i$ by describing how it maps to other vector spaces (it would be also be a proper use of the terminology universal property to describe how to map into a given vector space).

Notation 1.5.4. For I a set, we let $V^{\oplus I}$ denote the direct sum $\oplus_{i \in I} V_i$, i.e., the direct sum indexed by I with each "V_i" equal to V.

For $I = \{1, \ldots, n\}$, $n \in \mathbb{Z}^{>0}$, we use the notation $V^{\oplus n}$ instead.

Example 1.5.5. The vector space $k^{\oplus n}$ consists of n-tuples of elements of k, considered as a vector space under termwise addition and scaling given by termwise multiplication. We represent an element $v = (\lambda_i)_{i=1}^n$ of $k^{\oplus n}$ by the vector notation:

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}. \tag{1.2}$$

Definition 1.5.6. For V a vector space, we say that V is *finite-dimensional* (of dimension n) if there exists an isomorphism $V \xrightarrow{\simeq} k^{\oplus n}$. A choice of such isomorphism is called a *basis* for V. The vectors $e_1, \ldots, e_n$ in V corresponding under this isomorphism to the vectors:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

are the *basis vectors* of V corresponding to this basis. We say that n is the *dimension* of V and write $n = \dim V$.

Remark 1.5.7. To say that $e_1, \ldots, e_n$ are basis vectors of a basis of V is to say that every vector $v \in V$ can be written uniquely as:

$$v = \sum_{i=1}^{n} \lambda_i e_i$$

for $\lambda_i \in k$. The expression of v as such corresponds to the expression (1.2) for v.

Notation 1.5.8. Any linear transformation $T : k^{\oplus n} \to k^{\oplus m}$ determines a $(m \times n)$-*matrix* A, i.e., a matrix with m rows and n columns. The ith column of the matrix is the vector $T(e_i)$, written as in (1.2).

1.6. **Linear transformations.** A *subspace* $U \subseteq V$ is a subset closed under addition and scalar multiplication. Subspaces are the same as injective linear transformations $T : U \to V$. Dually, we define *quotient spaces* as surjective linear transformations $T : V \to W$.

Notation 1.6.1. By abuse, we let 0 denote the subspace of any vector space consisting only of the element 0.

Given a subspace U of V, we can form the quotient space $W := V/U$ whose elements are cosets of U in V, i.e., an element $w \in V/U$ is a subset of V of the form $v + U = \{v + u \mid u \in U\}$ for some $v \in V$. Addition and scalar multiplication on V/U are defined in the obvious ways, as is the homomorphism $V \to V/U$. The quotient space satisfies a universal property: giving a homomorphism $W = V/U \to W'$ is equivalent (by means of restriction along $V \to V/U$) to giving a homomorphism $V \to W'$ with the property that U maps to 0 under this transformation.

Construction 1.6.2. Given a linear transformation $T : V \to W$, we can associate various other vector spaces:

- We define the *kernel* $\mathrm{Ker}(T)$ of T as the subspace of V consisting of vectors $v \in V$ such that $T(v) = 0 \in W$.
- We define the *image* $\mathrm{Image}(T)$ of T as the subspace of W consisting of vectors $w \in W$ of the form $T(v)$ for some $v \in V$.
- We define the *cokernel* $\mathrm{Coker}(T)$ as the quotient $W/\mathrm{Image}(T)$ of W.

Notation 1.6.3. A diagram:

$$0 \to U \xrightarrow{T} V \xrightarrow{S} W \to 0 \tag{1.3}$$

is a *short exact sequence of vector spaces* if $\mathrm{Ker}(S) = U$ and $\mathrm{Coker}(T) = W$.

In other words, we say that (1.3) is a short exact sequence if U injects into V, V surjects onto W, the composition $S \circ T$ is equal to 0, and the induced maps:

$$U \to \mathrm{Ker}(S)$$
$$\mathrm{Coker}(S) \to W$$

are equivalences (it suffices to check this for either of the two maps).

Remark 1.6.4. There is clearly no mathematical content here. Still, the notion of short exact sequence is regarded as a useful organizing principle by many mathematicians, and we will use it as such below.

1.7. **Duality.** The *dual* vector space $V^\vee$ to a vector space V is defined as $\mathrm{Hom}_k(V, k)$.

Elements of $V^\vee$ are referred to as *linear functionals.*

Example 1.7.1. For $V = k^{\oplus n}$, the dual is canonically identified with V itself in the obvious way: however, the corresponding isomorphism $V \simeq V^\vee$ depends on the choice of basis.

We have a canonical map:

$$V \to (V^\vee)^\vee$$
$$v \mapsto \Big(\lambda \mapsto \lambda(v)\Big)$$

that is an isomorphism for V finite-dimensional (since we need only verify it for $V = k^{\oplus n}$, where it is obvious).

Given a map $T : V \to W$, we obtain the *dual* (aliases: *transpose* or *adjoint*) map $T^\vee : W^\vee \to V^\vee$ by restriction.

Exercise 1.2. For $V = k^{\oplus n}$ and $W = k^{\oplus m}$, show that the matrix of the dual transformation is the transpose of the matrix of the original linear transformation.

1.8. Let V, W, and U be k-vector spaces.

Definition 1.8.1. A function $B(-,-) : V \times W \to U$ is *(k-)bilinear* if for every $v \in V$ the map $B(v,-) : W \to U$ is linear, and for every $w \in W$ the map $B(-,w) : V \to U$ is linear.

Remark 1.8.2. We emphasize that bilinear maps are not linear (except in degenerate situations).

Proposition 1.8.3. *For every pair of vector spaces V and W, there is a k-vector space $V \otimes_k W$ defined up to unique isomorphism and equipped with a bilinear pairing:*

$$V \times W \to V \underset{k}{\otimes} W \tag{1.4}$$

such that every k-bilinear map $V \times W \to U$ factors uniquely as:

$$\begin{tikzcd}
V \times W \arrow[rr] \arrow[dr] & & V \underset{k}{\otimes} W \arrow[dl, dotted] \\
& U &
\end{tikzcd}$$

with the right map being k-linear.

Proof. Define the set I as $V \times W$. Define the vector space $\widetilde{V \otimes_k W}$ as $k^{\oplus I}$. For each $i = (v,w) \in I$, we abuse notation in letting (v,w) denote the element $\varepsilon_i(1) \in \widetilde{V \otimes_k W}$, i.e., the element of the direct sum that is 1 is the entry i and 0 in all other entries.

Define the subspace K of $\widetilde{V \otimes_k W}$ to be spanned by vectors of the form:

- $(v_1 + v_2, w) - (v_1, w) - (v_2, w)$ for $v_i \in V$, $w \in W$, $\lambda \in k$
- $(v, w_1 + w_2) - (v, w_1) - (v, w_2)$ for $v \in V$, $w_i \in W$, $\lambda \in k$
- $(\lambda \cdot v, w) - \lambda \cdot (v, w)$ for $v \in V$, $w \in W$, $\lambda \in k$
- $(v, \lambda \cdot w) - \lambda \cdot (v, w)$ for $v \in V$, $w \in W$, $\lambda \in k$.

Define $V \otimes_k W$ to be the quotient of $\widetilde{V \otimes_k W}$ by K.

We need to verify the universal property. By the universal property of direct sums, giving a linear transformation $\widetilde{B} : \widetilde{V \otimes_k W} \to U$ is equivalent to giving a function $B : V \times W \to U$: indeed, $B(v,w)$ is the image of (v,w) under $\widetilde{B}$. Moreover, by the universal property of quotients, this map will factor through $V \otimes_k W$ if and only if $\widetilde{B}$ sends each element of K to $0 \in U$. It suffices to verify this for the spanning set of vectors defining K, where it translates to the relations:

$$\begin{gathered}
B(v_1 + v_2, w) - B(v_1, w) - B(v_2, w) = 0 \\
B(v, w_1 + w_2) - B(v, w_1) - B(v, w_2) = 0 \\
B(\lambda \cdot v, w) - \lambda \cdot B(v, w) = 0 \\
B(v, \lambda \cdot w) - \lambda \cdot B(v, w) = 0
\end{gathered}$$

as relations in U. But these are exactly the conditions of bilinearity of the map B.

□

Notation 1.8.4. Where the field k is unambiguous, we use the notation $V \otimes W = V \otimes_k W$.

Notation 1.8.5. For $(v,w) \in V \times W$, let $v \otimes w$ denote the corresponding vector in $V \otimes W$, i.e., the image of (v,w) under the map (1.4).

Notation 1.8.6. For an integer $n \geqslant 0$, we let $V^{\otimes n}$ denote the n-fold tensor product of V with itself, where for the case $n = 0$ we understand the empty tensor product as giving k.

Exercise 1.3.
(1) Show using the proof of Proposition 1.8.3 that vectors of the form $v \otimes w$ for $v \in V$ and $w \in W$ span $V \otimes W$.
(2) Show the same using the universal property of tensor products alone (hint: consider the span of the image of the map $V \times W \to V \otimes_k W$ and show that this subspace also satisfies the universal property of tensor products).
(3) If $\dim V \neq 0, 1$, show that not every vector in $V \otimes W$ is of the form $v \otimes w$.

Remark 1.8.7. The construction appearing in the proof of Proposition 1.8.3 appears very complicated: we take a huge vector space (with a basis indexed by elements of $V \times W$), and quotient by a huge subspace. However, things are not as bad as they look! For $V = k^{\oplus n}$ and $W = k^{\oplus m}$, we claim that there is a canonical identification:

$$V \otimes W \simeq k^{\oplus (nm)}.$$

In particular, the tensor product of finite-dimensional vector spaces is again finite-dimensional.

We denote the given bases of V and W by $e_1, \ldots, e_n$ and $e'_1, \ldots, e'_m$ respectively. It suffices to show that the vectors $e_i \otimes e'_j$ form a basis of $V \otimes W$. Given $v = \sum_i \lambda_i e_i$ and $w = \sum_j \eta_j e'_j$, by bilinearity of $V \times W \to V \otimes W$, we have:

$$v \otimes w = \sum_{i,j} \lambda_i \eta_j \cdot e_i \otimes e'_j.$$

By Exercise 1.3, this shows that the vectors $e_i \otimes e'_j$ span $V \otimes W$. In particular, the dimension of $V \otimes W$ is $\leqslant n \cdot m$.

Denote the given basis of $k^{\oplus (nm)}$ by $f_1, \ldots, f_{n \cdot m}$. Define the bilinear map:

$$V \times W \to k^{\oplus (nm)}$$

by the formula:

$$(\sum_i \lambda_i e_i, \sum_j \eta_j e'_j) \mapsto \sum_{i,j} \lambda_i \eta_j \cdot f_{(i-1)\cdot m + j}.$$

(Note that as (i, j) varies subject to the conditions $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$, the value $(i-1) \cdot m + j$ obtains each integral value between 1 and nm exactly once). One immediately verifies that this map actually is bilinear.

The induced map:

$$V \otimes W \to k^{\oplus (nm)}$$

sends $e_i \otimes e'_j$ to $f_{(i-1)\cdot m + j}$, and therefore spans the target. Therefore, since $\dim V \otimes W \leqslant n \cdot m$, this map must be an isomorphism, as desired.

As a corollary, one may try to think of tensor product of vector spaces as an animation of the usual multiplication of non-negative integers, since $\dim(V \otimes W) = \dim(V) \cdot \dim(W)$.

Exercise 1.4. Show that the map:

$$\mathrm{Hom}(V \otimes W, U) \to \mathrm{Hom}(V, \mathrm{Hom}(W, U))$$
$$T \mapsto \Big(v \mapsto \big(w \mapsto T(v, w)\big)\Big)$$

is an isomorphism.

Exercise 1.5. For V and W vector spaces, show that the map:

$$\begin{aligned} V^\vee \otimes W &\to \mathrm{Hom}(V, W) \\ \lambda \otimes w &\mapsto (v \mapsto \lambda(v) \cdot w) \end{aligned}$$

is an isomorphism if V is finite-dimensional.

In the case when W is also finite-dimensional, understand the relationship to the matrix representation of linear transformations in the presence of bases (hint: note that the right hand side is then a vector space of matrices).

Exercise 1.6. Show that the map:

$$V^\vee \otimes W^\vee \to (V \otimes W)^\vee$$

induced by the map:

$$\lambda \otimes \eta \mapsto \Big(v \otimes w \mapsto \lambda(v) \cdot \eta(w)\Big)$$

is an isomorphism for V and W finite-dimensional.

1.9. Suppose that V is a vector space. Define the *tensor algebra* $T(V)$ of V as the direct sum $\oplus_{n \in \mathbb{Z}^{\geqslant 0}} V^{\otimes n}$.

Note that $T(V)$ can indeed be made into a k-algebra in a natural way: the multiplication is given component-wise by the maps:

$$V^{\otimes n} \otimes V^{\otimes m} \xrightarrow{\simeq} V^{\otimes (n+m)}.$$

Proposition 1.9.1. *Let A be a k-algebra. Restriction along the map $V \to T(V)$ induces an isomorphism:*

$$\mathrm{Hom}_{\mathsf{Alg}_{/k}}(T(V), A) \xrightarrow{\simeq} \mathrm{Hom}_k(V, A).$$

Here $\mathrm{Hom}_{\mathsf{Alg}_{/k}}$ *indicates the set of maps of k-algebras.*

Remark 1.9.2. This is a universal property of the tensor algebra, typically summarized by saying that $T(V)$ is the *free associative algebra on the vector space* V.

Proof. Given $T : V \to A$ a map of vector spaces, we obtain a map $V^{\otimes n} \to A$ as the composition:

$$V \otimes \ldots \otimes V \xrightarrow{T \otimes \ldots \otimes T} A \otimes \ldots A \to A$$

where the last map is induced by the n-fold multiplication on A.

One easily verifies that the corresponding map $T(V) = \oplus_n V^{\otimes n} \to A$ is a map of k-algebras and provides an inverse to the restriction map.

□

1.10. Let $n \geqslant 0$ be an integer and let V be a vector space.

We consider quotients Sym^n and Λ^n of V uniquely characterized by the following universal properties:

- A map $T : V^{\otimes n} \to W$ factors as:

$$\begin{tikzcd} V^{\otimes n} \arrow[d, two heads] \arrow[dr, "T"] & \\ \mathrm{Sym}^n(V) \arrow[r, dotted] & W \end{tikzcd}$$

 if and only if T is *symmetric*, i.e., for every collection $v_1, \ldots, v_n$ of vectors of V and every permutation (i.e., bijection) $\sigma : \{1, \ldots, n\} \xrightarrow{\simeq} \{1, \ldots, n\}$, we have:

$$T(v_1 \otimes \ldots \otimes v_n) = T(v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)}).$$

- A map $T : V^{\otimes n} \to W$ factors as:

$$\begin{tikzcd} V^{\otimes n} \arrow[d, two heads] \arrow[dr, "T"] & \\ \Lambda^n(V) \arrow[r, dotted] & W \end{tikzcd}$$

 if and only if T is *alternating*, i.e., for every collection $v_1, \ldots, v_n$ of vectors of V and every permutation (i.e., bijection) $\sigma : \{1, \ldots, n\} \xrightarrow{\simeq} \{1, \ldots, n\}$, we have:

$$T(v_1 \otimes \ldots \otimes v_n) = \mathrm{sgn}(\sigma) \cdot T(v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)}).$$

 Here $\mathrm{sgn}(\sigma)$ is the *sign* of the permutation σ.

Definition 1.10.1. We say that $\mathrm{Sym}^n(V)$ (resp. $\Lambda^n(V)$) is the *nth symmetric power* (resp. *nth alternating power*) of V.

Notation 1.10.2. For $v_1, \ldots, v_n \in V$, we let $v_1 \ldots v_n$ or $v_1 \cdot \ldots \cdot v_n$ denote the corresponding element of $\mathrm{Sym}^n(V)$, i.e., the image of $v_1 \otimes \ldots \otimes v_n$ under the structure map $V^{\otimes n} \to \mathrm{Sym}^n(V)$. We let $v_1 \wedge \ldots \wedge v_n$ denote the corresponding vector of $\Lambda^n(V)$.

Note that e.g. for $n = 2$ we have $v_1 v_2 = v_2 v_1$ and $v_1 \wedge v_2 = -v_2 \wedge v_1$.

Exercise 1.7. Suppose that V is finite-dimensional with basis $e_1, \ldots, e_m$.

(1) Show that $\mathrm{Sym}^n(V)$ has a basis indexed by the set of non-decreasing functions $f : \{1, \ldots, n\} \to \{1, \ldots, m\}$, where the corresponding basis vector is $e_{f(1)} \ldots e_{f(n)}$.

 Deduce that $\dim(\mathrm{Sym}^n(V))$ is the binomial coefficient $\binom{n+m-1}{m}$.

(2) Show that $\Lambda^n(V)$ has a basis indexed by the subsets I of $\{1, \ldots, m\}$ of order n, where the corresponding basis vector is $e_I := \wedge_{i \in I} e_i$ (i.e., the iterated wedge product of the elements e_i for $i \in I$, where we write the wedge product in the increasing order according to size of $i \in [1, n]$).

 Deduce that $\dim(\Lambda^n(V))$ is the binomial coefficient $\binom{m}{n}$ for $0 \leqslant n \leqslant m$, and dimension 0 for $n > m$.

We define the *symmetric algebra* $\mathrm{Sym}(V)$ as $\oplus_{n \geqslant 0} \mathrm{Sym}^n(V)$. Note that $\mathrm{Sym}(V)$ is a commutative algebra in an obvious way. We then have the following analogue of Proposition 1.9.1.

Proposition 1.10.3. *Let A be a commutative k-algebra. Restriction along the map $V \to \mathrm{Sym}(V)$ induces an isomorphism:*

$$\mathrm{Hom}_{\mathsf{Alg}_{/k}}(\mathrm{Sym}(V), A) \xrightarrow{\simeq} \mathrm{Hom}_k(\mathrm{Sym}(V), A).$$

Remark 1.10.4. This is a universal property of the symmetric algebra, typically summarized by saying that $\mathrm{Sym}(V)$ is the *free commutative algebra on the vector space V*.

Construction 1.10.5. Elements of $\mathrm{Sym}^n(V^\vee)$ can be considered as k-valued functions on V, where for $\lambda_1, \ldots, \lambda_n \in V^\vee$ the function on V corresponding to $\lambda_1 \ldots \lambda_n$ is $v \mapsto \lambda_1(v) \ldots \lambda_n(v)$.

Functions on V arising in this manner are by definition *polynomial functions of degree n on V*. Sums of such functions are called polynomial functions on V.

1.11. **Bilinear forms.** Let V be a k-vector space.

Definition 1.11.1. A *(k-valued) bilinear form* on V is a bilinear map $B : V \times V \to k$.

Suppose that V is finite-dimensional.

Remark 1.11.2. Note that $\mathrm{Hom}(V \otimes V, k) = (V^{\otimes 2})^\vee \simeq (V^\vee)^{\otimes 2}$, so that bilinear forms may be considered as elements of $V^{\vee,\otimes 2}$.

Given a bilinear form B, we obtain a second bilinear form B^T by setting $B^T(v, w) := B(w, v)$. We refer to B^T as the *transpose* or *adjoint* bilinear form.

Remark 1.11.3. By Exercise 1.5, we have an isomorphism:

$$V^{\vee,\otimes 2} \xrightarrow{\simeq} \mathrm{Hom}(V, V^\vee)$$

That is, bilinear forms are equivalent to linear transformations $V \to V^\vee$. To normalize the conventions: under this dictionary, the bilinear form B maps to the linear transformation $V \to V^\vee$ associating to a vector v the linear functional $w \mapsto B(v, w)$.

Exercise 1.8. Show that if $T : V \to V^\vee$ corresponds to a bilinear form B, the linear transformation corresponding to B^T is the dual:

$$T^\vee : (V^\vee)^\vee = V \to V^\vee$$

of T.

Definition 1.11.4. A bilinear form B is *non-degenerate* if the corresponding linear transformation $V \to V^\vee$ is an isomorphism.

Exercise 1.9.
1. Show that B is non-degenerate if and only if for every $0 \neq v \in V$ the corresponding functional $B(v, -)$ is non-zero, i.e., there exists $w \in V$ with $B(v, w) \neq 0$.
2. Suppose that $e_1, \ldots, e_n$ is a basis for V and that $B(e_i, -)$ is a non-zero functional for each i. Is B necessarily non-degenerate?
3. Show that B is non-degenerate if and only if B^T is.

1.12. **Symmetric bilinear forms.** We now focus on the special case where the bilinear form is *symmetric.*

Definition 1.12.1. A bilinear form B is *symmetric* if $B = B^T$.

The following result records some equivalent perspectives on the symmetry conditions.

Proposition 1.12.2. *The following are equivalent:*

(1) *The bilinear form B is* symmetric.
(2) $B(v,w) = B(w,v)$ *for all* $v,w \in V$.
(3) $B \in V^{\vee,\otimes 2}$ *is left invariant under the automorphism:*

$$\lambda \otimes \eta \mapsto \eta \otimes \lambda$$

of $V^{\vee,\otimes 2}$.
(4) *The linear transformation* $V \to V^{\vee}$ *corresponding to* B *is self-dual.*

2. Quadratic forms and spaces

2.1. Quadratic spaces.

Let k be a field.

Definition 2.1.1. A *quadratic space* over k is a pair (V,q) where V is a finite-dimensional k-vector space and $q \in \mathrm{Sym}^2(V^{\vee})$.

Remark 2.1.2. Given a quadratic space (V,q), we obtain a function $V \to k$ by Construction 1.10.5. By abuse of notation, we denote this function also by q.

Note that q satisfies the properties:

- $q(\lambda v) = \lambda^2 q(v)$ for $v \in V$ and $\lambda \in k$.
- The map $V \times V \to k$ defined by $(v,w) \mapsto q(v+w) - q(v) - q(w)$ is bilinear.

Definition 2.1.3. A function $q : V \to k$ satisfying the two conditions of Remark 2.1.2 is said to be a *quadratic form on V*.

Exercise 2.1. If $q : V \to k$ is a quadratic form, show that q arises from a unique quadratic space structure on V.

Example 2.1.4. Suppose that (V,q) is a quadratic space with a chosen basis $e_1, \ldots, e_n$ of V. Letting $x_1, \ldots, x_n$ denote the corresponding dual basis of $V^{\vee}$, we see that q is an expression of the form:

$$\sum_i a_{i,i} x_i^2 + \sum_{i<j} a_{i,j} x_i x_j \tag{2.1}$$

for $a_{i,j} \in k$.

The corresponding function $q : k^n \to k$ is given by substitution in the above expression:

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \sum_i a_{i,i} \lambda_i^2 + \sum_{i<j} a_{i,j} \lambda_i \lambda_j.$$

In this case, we will write q as a function of n-variables in the usual way.

Definition 2.1.5. A quadratic space structure on k^n is a *quadratic form in n variables*. That is, a quadratic form in n-variables is an expression (2.1).

Terminology 2.1.6. When referencing quadratic forms without mentioning an ambient vector space, we will typically mean a quadratic form in some variables, i.e., a quadratic form on some $k^{\oplus n}$.

Terminology 2.1.7. We follow classical usage is referring to a quadratic form in one (resp. two, resp. three, resp. four) variables as a *unary* (resp. *binary*, resp. *ternary*, resp. *quaternary*) form.

Remark 2.1.8. By definition, a quadratic form is a quadratic space with chosen coordinates. Historically, quadratic forms were the original objects of study. However, we will see that for some questions (e.g., representability by the form—see below), the perspective of quadratic spaces is much more flexible.

In what follows, we will first develop the theory in the abstract setting of quadratic spaces, and then explain how to translate the questions of the theory into problems of the matrix calculus of linear algebra. This approach will not be ideal for all readers, some of whom will prefer to practice computing before developing the abstract (and simpler) setting. Such readers are advised to skip ahead to §2.6 and refer back as necessary.

Remark 2.1.9. The definition of quadratic space allows to consider the case $V = 0$ (and $q = 0$). We refer to this as the *zero quadratic space*. This should not be confused with the zero quadratic form on a vector space V, which is defined as $q(v) = 0$ for all $v \in V$.

2.2. **Morphisms of quadratic spaces.** A *morphism* $T : (V, q_V) \to (W, q_W)$ of quadratic spaces is a linear transformation $T : V \to W$ such that $q_W(T(v)) = q_V(v)$ for every $v \in V$.

Remark 2.2.1. Such linear transformations are often called *isometries* in the literature (this phrase also sometimes refers more specifically to injective or bijective morphisms of quadratic spaces).

An *isomorphism* of quadratic spaces is a morphism that is a morphism inducing an isomorphism $V \xrightarrow{\simeq} W$ of vector spaces.

We say that two quadratic forms in n-variables are *equivalent* if the corresponding quadratic spaces are isomorphic.

2.3. **Operations with quadratic spaces.** We give some basic constructions of new quadratic spaces from old: direct sum, restriction, and extension of scalars.

Suppose that (V, q_V) and (W, q_W) are quadratic spaces. Then $V \oplus W$ carries a canonical quadratic form $q_V \oplus q_W$ defined by $(q_V \oplus q_W)(v, w) = q_V(v) + q_W(w)$.

Next, suppose that $T : V \to W$ is a linear transformation and q_W is a quadratic form on W. Define q_V by $q_V(v) := q_W(T(v))$. Obviously q_V is a quadratic form on V; we say that it is obtained by *restriction* and sometimes denote it by $q_W|_V$.

$$q(v + w) = q(v) + q(w) + B_q(v, w)$$

Finally, let (V, q) be a quadratic space over k, and let $k \hookrightarrow k'$ be an embedding of fields. Then $V \otimes_k k'$ is naturally a k'-vector space and q extends to a quadratic form $q_{k'}$ on $V \otimes_k k'$. For example, if we fix a basis $V \simeq k^{\oplus n}$ of V, then $V \otimes_k k'$ acquires a natural basis, and the quadratic form (2.1) defines the "same" quadratic form, where we consider elements of k as elements of k' through the embedding.

Remark 2.3.1. It would be better to refer to $(V \oplus W, q_{V \oplus W})$ as the *orthogonal direct sum*, as in [Ser]. Indeed, morphisms $(V \oplus W, q_{V \oplus W}) \to (U, q_U)$ are equivalent to giving a pair of morphisms $T : (V, q_V) \to (U, q_U)$ and $S : (W, q_W) \to (U, q_U)$ with the additional conditions that their images be orthogonal, i.e., $B_{q_U}(T(v), S(w)) = 0$ for all $v \in V$ and $w \in W$.

2.4. **Representability.** Suppose that (V, q) is a quadratic space.

Definition 2.4.1. We say that q *represents* $\lambda \in k$ if there exists a vector $v \in V$ such that $q(v) = \lambda$ (i.e., if λ is in the image of the function q).

Remark 2.4.2. If q represents λ, then q represents $\eta^2\lambda$ for any $\eta \in k$: indeed, if $q(v) = \lambda$ then $q(\eta v) = \eta^2\lambda$. Therefore, we may also speak about q representing classes in $k^\times/(k^\times)^2$.

Example 2.4.3. Suppose that $V = k$ is the trivial one-dimensional vector space. Then by Example 2.1.4, q must be of the form $q(x) = ax^2$ for some $a \in k$, and we assume $a \neq 0$ for conveninence. We see that q represents λ if and only if λ/a is a square in k.

We specialize this to particular fields:

(1) For example, if k is algebraically closed, q represents all values $\lambda \in k$ so long as $a \neq 0$.
(2) If $k = \mathbb{R}$, q represents λ if and only if a and λ have the same sign.
(3) If $k = \mathbb{F}_p$, then q represents λ if and only if λ/a is a quadratic residue in $\mathbb{F}_p$. Indeed, here "quadratic residue" simply means "is a square."
(4) Suppose $k = \mathbb{Q}$. We introduce the following notation: for p a prime number and $n \neq 0$ an integer, let $v_p(n) \in \mathbb{Z}^{\geqslant 0}$ be the order to which p divides n (alias: the *valuation of n at p*), i.e., it is the unique integer such that $\frac{n}{p^{v_p(n)}} \in \mathbb{Z}$ and $\frac{n}{p^{v_p(n)+1}}$ is not. More generally, for $r = \frac{n}{m} \in \mathbb{Q}^\times$, let $v_p(r) := v_p(n) - v_p(m)$. (For example: $v_p(p) = 1$, $v_p(p^2) = 2$, and $v_p(\frac{1}{p}) = -1$).

   Then q as above represents λ if and only if a and λ have the same sign, and moreover, $v_p(a) - v_p(\lambda)$ is even for each prime p (if this is confusing, it is instructive here to work out the case $a = 1$).

   We notice the following feature: q represents a particular value if and only if certain conditions are verified at each prime p, and there is an additional condition involving $\mathbb{R}$ (here that a and λ have the same sign). This is the most elementary instance of the *Hasse principle.*
(5) In general, if a and λ are non-zero (the case where either is zero being trivial), then q obviously represents λ if and only if a and λ project to the same element in the (2-torsion) abelian group $k^\times/(k^\times)^2$. The above analysis amounts to saying that $\mathbb{R}^\times/(\mathbb{R}^\times)^2 = \mathbb{Z}/2\mathbb{Z}$ (according to sign), and $\mathbb{Q}^\times/(\mathbb{Q}^\times)^2$ is isomorphic to the direct sum:
$$\mathbb{Z}/2\mathbb{Z} \oplus \big(\bigoplus_{p \text{ prime}} \mathbb{Z}/2\mathbb{Z}\big)$$
   given by the sign homomorphism on the first factor and the parity of the valuation at p on the factor corresponding to a prime p.

Exercise 2.2. This exercise is meant to flesh out (3) above.

(1) Show that there is a (necessarily unique) isomorphism $\mathbb{F}_p^\times/(\mathbb{F}_p^\times)^2 \simeq \mathbb{Z}/2\mathbb{Z}$ as long as $p \neq 2$.
(2) For $a \in \mathbb{F}_p^\times$, we use the *Legendre symbol* $\left(\frac{a}{p}\right) \in \{1, -1\}$ for the resulting map:
$$\mathbb{F}_p^\times \to \mathbb{F}_p^\times/(\mathbb{F}_p^\times)^2 \simeq \mathbb{Z}/2\mathbb{Z} = \{1, -1\}.$$
   Here we use multiplicative notation for $\mathbb{Z}/2\mathbb{Z}$.

   Show that $\left(\frac{-}{p}\right)$ is a homomorphism $\mathbb{F}_p^\times \to \{1, -1\}$.
(3) Show that $\frac{p-1}{2}$ elements of $\mathbb{F}_p^\times$ are quadratic residues and an equal number are not.

Exercise 2.3. Suppose that $k = \mathbb{Q}$, $V = k^2$ and $q(x, y) = x^2 + y^2$. Show that q represents $\lambda \in \mathbb{Z}$ implies that $\lambda \neq 3$ modulo 4.

Of course, every quadratic form represents 0, since $q(0) = 0$. However, the following condition is less trivial.

Definition 2.4.4. An *isotropic vector* in (V, q) is a nonzero $v \in V$ such that $q(v) = 0$. The quadratic form $q : V \to k$ is *isotropic* if there exists an isotropic vector in (V, q).

We say that q is *anisotropic* if it is not isotropic.

Exercise 2.4. Let k be a subfield of $\mathbb{R}$ (e.g., $k = \mathbb{Q}$ or $k = \mathbb{R}$).

(1) show that $q(x_1, \ldots, x_n) = \sum_{i=1}^n x_i^2$ is anisotropic. Show that for $0 \neq a_i \in k = \mathbb{R}$, the form:
$$q(x_1, \ldots, x_n) = \sum_{i=1}^n a_i x_i^2$$
is anisotropic if and only if all the a_i have the same sign (i.e., are all positive or all negative). Show that for $k = \mathbb{Q}$ this condition is no longer sufficient.

(2) More generally, we say that a form (V, q) is *positive-definite* (resp. *negative-definite*) if $q(v) > 0$ (resp. $q(v) < 0$) for all $0 \neq v \in V \otimes_k \mathbb{R}$, and merely *definite* if it is either positive or negative definite.[1]

Show that any definite form is anisotropic.

(3) Show that a form $q(x, y) = ax^2 + bxy + cy^2$ is definite if and only if the discriminant $\Delta = b^2 - 4ac$ is negative.

Exercise 2.5. Let k be algebraically closed. Show that every quadratic space (V, q) over k with $\dim(V) > 1$ is isotropic.

Note that the property of being isotropic is preserved under extension of scalars, although the property of being anisotropic is not necessarily.

Suppose that $\lambda \in k$ and $q : V \to k$ is a quadratic form. Define $q_\lambda : V \oplus k \to k$ by $q_\lambda(v, \eta) = q(v) - \lambda \cdot \eta^2$: obviously q_λ is a quadratic form.

Proposition 2.4.5. *If q represents λ, then q_λ is isotropic. Moreover, if q is itself anisotropic, then this condition is necessary and sufficient.*

Proof. If $q(v) = \lambda$, then:
$$q_\lambda(v, 1) = q(v) - \lambda \cdot 1 = \lambda - \lambda = 0$$
so that q_λ is isotropic.

Conversely, if $q_\lambda(v, \eta) = 0$, then either $\eta = 0$, in which case $q(v) = 0$, or else:
$$q(\eta^{-1} v) = \eta^{-2} q(v) = \eta^{-2}\big(q_\lambda(v, \eta) + \lambda \cdot \eta^2\big) = \lambda$$
as desired.

□

Remark 2.4.6. In this manner, representability questions can be reduced to the question of a form being isotropic, which can be easier to treat. We will develop these methods further below, proving in particular the strengthening Corollary 2.10.3 of Proposition 2.4.5.

2.5. By Remark 2.1.2, to every quadratic space (V, q) there is an associated bilinear form B_q:
$$B_q : V \times V \to k$$
$$B_q(v, w) := q(v + w) - q(v) - q(w).$$
Obviously B_q is symmetric.

Likewise, given B a bilinear form on V, the function q_B defined by $q_B(v) := B(v, v)$ is a quadratic form, since:
$$q_B(\lambda v) = B(\lambda v, \lambda v) = \lambda^2 B(v, v) = \lambda^2 q_B(v)$$

[1] We remark explicitly that definiteness depends only on the form obtained by extending scalars from k to $\mathbb{R}$.

and:

$$q_B(v+w) - q_B(v) - q_B(w) = B(v+w, v+w) - B(v,v) - B(w,w) =$$
$$B(v,v) + B(v,w) + B(w,v) + B(w,w) - B(v,v) - B(w,w) = B(v,w) + B(w,v) = (B + B^T)(v,w)$$

and the latter expression is manifestly bilinear as the sum of two bilinear forms.

Observe that these constructions are almost, but not quite, inverses to each other: the quadratic form q_{B_q} (associated with the bilinear form associated with the quadratic from q) is $2 \cdot q$:

$$q_{B_q}(v) = B_q(v,v) = q(v+v) - q(v) - q(v) = q(2 \cdot v) - 2 \cdot q(v) = 4q(v) - 2q(v) = 2q(v).$$

Likewise, $B_{q_B} = 2 \cdot B$ if B is symmetric bilinear (generally, it is $B + B^T$, as we have already seen).

Therefore, we deduce that there is a bijection between quadratic forms and symmetric bilinear forms as long as the characteristic of k is not 2, with each of the maps above being bijections (though not quite mutually inverse bijections due to the factor of 2).

Exercise 2.6. Show that the bilinear form associated with the quadratic form:

$$\sum_{i=1}^{n} a_{i,i} x_i^2 + \sum_{i<j} a_{i,j} x_i x_j$$

is given by the bilinear form:

$$\left(\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}, \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \right) \mapsto \sum_{i=1}^{n} 2a_{i,i} \lambda_i \eta_j + \sum_{i<j}^{n} a_{i,j}(\lambda_i \eta_j + \lambda_j \eta_i).$$

Deduce that the assignment of a symmetric bilinear form to a quadratic form is neither injective nor surjective in characteristic 2.

Exercise 2.7. If $T : (V, q_V) \to (W, q_W)$ is a morphism of quadratic spaces, show that for every $v_1, v_2 \in V$ we have:

$$B_{q_V}(v_1, v_2) = B_{q_W}(T(v_1), T(v_2)).$$

Remark 2.5.1. The equivalence between quadratic forms and symmetric bilinear forms for $\text{char}(k) \neq 2$ may be considered as an instance of Maschke's theorem from the representation theory of finite groups, considering the representation of $\mathbb{Z}/2\mathbb{Z}$ on $V^{\otimes 2}$ given by switching factors.

2.6. **Matrices.** Suppose that V is equipped with a bilinear form $B : V \otimes V \to k$ and a basis $e_1, \ldots, e_n$. As in §1.11, we associate to B a linear transformation $V \to V^\vee$, and using our basis we may view this as a linear transformation from $k^{\oplus n}$ to itself, i.e., as a matrix. This evidently defines a bijection between $(n \times n)$-matrices and bilinear forms on $k^{\oplus n}$.

Exercise 2.8. (1) Show that the bilinear form associated to a $(n \times n)$-matrix A is given by:

$$(v, w) \mapsto w^T A v$$

where we consider v and w as column vectors, and the superscript T indicates the transpose of a matrix (so w^T is the row vector associated to w); note that the result of the above matrix multiplication is a (1×1)-matrix, i.e., an element of k.

(2) Show that the matrix associated to the bilinear form B above is:

$$A = \begin{pmatrix} B(e_1,e_1) & B(e_2,e_1) & \dots & B(e_n,e_1) \\ B(e_1,e_2) & B(e_2,e_2) & \dots & B(e_n,e_2) \\ \vdots & \vdots & \vdots & \vdots \\ B(e_1,e_n) & B(e_2,e_n) & \dots & B(e_n,e_n) \end{pmatrix}.$$

In particular, the bilinear form is symmetric if and only if its associated matrix is symmetric (meaning that it is equal to its own transpose).

Exercise 2.9. Suppose that $f_1,\dots,f_n$ is a second choice of basis. Let S be the (invertible) matrix whose ith column is the vector $f_i \in k^{\oplus n}$ (written in coordinates relative to the basis $\{e_j\}$). Suppose that we find that matrix A' of a bilinear form B with respect to the basis $f_1,\dots,f_n$. Show that A' is obtained from A as:

$$A' = S^T AS.$$

Conclude that if $\mathrm{char}(k) \neq 2$, then equivalence classes of quadratic spaces over k are in bijection with the quotient set of symmetric matrices modulo the action of $GL_n(k)$ by $S \cdot A := S^T AS$.

We will also speak of the matrix of a quadratic form, meaning the matrix of the associated symmetric bilinear form.

2.7. **A remark regarding characteristic 2.** Quadratic forms in characteristic 2 behave different from in other characteristics. In many respects, they can be regarded as pathological.

We note that they do arise in nature as the reduction modulo 2 of integral quadratic forms, and for some purposes they cannot be avoided. However, for the purposes of these notes, we will not need the theory.

Still, the author has opted to include some finer material regarding characteristic 2 than is needed. The reader may safely ignore all of it, ignoring in particular the difference between the radical and the reduced radical, and the difference between various notions of non-degeneracy. The only loss to the reader would be a few fun exercises.

2.8. **Radical.** Let (V,q) be a quadratic space. We define the *radical* $\mathrm{Rad}(V)$ of V as the subset of vectors $v \in V$ such that $B_q(v,w) = 0$ for every $w \in V$. We define the *reduced radical* $\overline{\mathrm{Rad}}(V)$ as the subset of isotropic vectors in the radical of V.

Clearly $\mathrm{Rad}(V)$ is a subspace, and one immediately finds that $\overline{\mathrm{Rad}}(V)$ is as well: indeed, if $v,w \in \overline{\mathrm{Rad}}(V)$, then $q(v) = q(w) = 0$ and $B_q(v,w) = 0$ so that $q(v+w) = q(v)+q(w)+B_q(v,w) = 0$.

If $\mathrm{char}(k) \neq 2$, then $\mathrm{Rad}(V) = \overline{\mathrm{Rad}}(V)$: indeed, for any $v \in \mathrm{Rad}(V)$ we have $0 = B_q(v,v) = 2 \cdot q(v)$.

Construction 2.8.1. For every $v \in \overline{\mathrm{Rad}}(V)$ and every $w \in V$, we have:

$$q(v+w) = q(v) + q(w) + B_q(v,w) = q(w)$$

and therefore we obtain an induced quadratic form on the quotient space $V/\overline{\mathrm{Rad}}(V)$. Of course, more generally, we obtain such a form on the quotient by any subspace of the reduced radical.

We define the *rank* $\mathrm{rank}(q)$ of a quadratic space (V,q) as $\dim(V) - \dim(\mathrm{Rad}(V))$ and the *reduced rank* of a quadratic space as $\dim(V) - \dim(\overline{\mathrm{Rad}}(V))$.

Exercise 2.10. Let k be a perfect[2] field of characteristic 2. Show that $\dim(\mathrm{Rad}(V)) - \dim(\overline{\mathrm{Rad}}(V))$ must equal 0 or 1.

2.9. **Non-degenerate quadratic forms.** We say that (V, q) is *non-degenerate* if $\overline{\mathrm{Rad}}(V) = 0$, and we say that (V, q) is *strongly non-degenerate* if $\mathrm{Rad}(V) = 0$ (equivalently: B_q is non-degenerate). These notions coincide if $\mathrm{char}(k) \neq 2$.

Remark 2.9.1. Unwinding the definitions: (V, q) is strongly non-degenerate if for every $0 \neq v \in V$ there exists $w \in V$ with $B_q(v, w) \neq 0$, and non-degenerate if for every $0 \neq v \in V$, either there exists such a $w \in V$ or else $q(v) \neq 0$.

Remark 2.9.2. Clearly a form is (resp. strongly) non-degenerate if and only if its rank (resp. reduced rank) is equal to its dimension.

Example 2.9.3. The 1-variable quadratic form $q(x) = ax^2$ for $a \in k$ is non-degenerate in the above sense if and only if $a \neq 0$ (for k of arbitrary characteristic). However, for $\mathrm{char}(k) = 2$, this form is not strongly non-degenerate.

Exercise 2.11. Suppose that $T : (V, q_V) \to (W, q_W)$ is a morphism of quadratic spaces with (V, q_V) non-degenerate. Show that T is injective.

Exercise 2.12. Show that $V/\overline{\mathrm{Rad}}(V)$ is non-degenerate when equipped with the quadratic space structure of Construction 2.8.1.

Definition 2.9.4. Suppose that (V, q) is a quadratic space and W is a subspace of V. The *orthogonal subspace* $W^\perp$ to W is the subspace of vectors $v \in V$ such that $B_q(v, w) = 0$ for all $w \in W$.

Remark 2.9.5. If (V, q) is strongly non-degenerate, then $\dim W + \dim W^\perp = \dim V$, since $W^\perp := \mathrm{Ker}(V \to W^\vee)$ and the map $V \to W^\vee$ is surjective by strong non-degeneracy.

Remark 2.9.6. If $\mathrm{char}(k) \neq 2$, then any vector in $W \cap W^\perp$ is isotropic.

Proposition 2.9.7. *Suppose that (V, q) is a quadratic space and $W \subseteq V$ with $(W, q|_W)$ strongly non-degenerate. Then $(V, q) \simeq (W, q|_W) \oplus (W^\perp, q|_{W^\perp})$.*

$(W^\perp, q|_{W^\perp})$ is strongly non-degenerate if and only if (V, q) is.

Proof. First, we claim that $W \cap W^\perp = 0$. Suppose that $0 \neq w \in W$. Then, because $q|_W$ is strongly non-degenerate, there exists $w' \in W$ with $B_q(w, w') \neq 0$, meaning that $w \notin W^\perp$.

Next, for any $v \in V$, we claim that there exists $w \in W$ with $B_q(v, w') = B_q(w, w')$ for every $w' \in W$. Indeed, $B_q(v, -) : W \to k$ is a linear functional, and therefore, by strong non-degeneracy of W, it is of the form $B_q(w, -)$ for a unique choice of $w \in W$. Observing that $v - w$ is obviously in $W^\perp$, we obtain that $V = W \oplus W^\perp$.

We then note that for any $w \in W$, $w' \in W^\perp$, we have:

$$q(w + w') = q(w) + q(w') + B_q(w, w') = q(w) + q(w')$$

so that the quadratic form q is obtained as the direct sum of its restriction to these subspaces.

The second part follows from Exercise 2.13 below.

□

Exercise 2.13. Show that the direct sum of two quadratic spaces is non-degenerate if and only if each of the summands is.

[2] Recall that a field of characteristic p is said to be perfect if the map $x \mapsto x^p$ is an isomorphism.

The following lemma is useful for verifying non-degeneracy computationally.

Lemma 2.9.8. *Suppose that q is a quadratic form in n-variables with associated bilinear form having matrix A, as in §2.6.*

Then q is strongly non-degenerate if and only if $\det(A) \neq 0$.

Proof. Indeed, for $V = k^n$, A is the matrix associated to the map $k^n = V \to V^\vee = k^n$, so it is an equivalence if and only if its determinant is non-zero.

□

Remark 2.9.9. Some authors use the term "regular" instead of non-degenerate.

Remark 2.9.10. The reader may safely skip this remark.

A more sophisticated variant of the above in characteristic 2: we say that (V, q) is *geometrically non-degenerate* if $\overline{\mathrm{Rad}}(V \otimes_k k') = 0$ for every field extension $k \hookrightarrow k'$. Note that formation of the usual radical Rad commutes with extension of scalars, so this definition still gives the same answer in characteristic not equal to 2. Note that by Exercise 2.10 this implies that the dimension of the radical is $\leqslant 1$. We note that if $\lambda \in k$ is not a square, then $x^2 + \lambda y^2$ is an example of a non-degenerate but not geometrically non-degenerate quadratic form.

The reader may take as an exercise to show that geometric non-degeneracy is equivalent to strong non-degeneracy if $\dim(V)$ is even, and equivalent to $\dim(\overline{\mathrm{Rad}}) = 1$ if $\dim(V)$ is odd.

We remark that geometric non-degeneracy is equivalent to algebro-geometric conditions: essentially just smoothness of the associated (projective) quadric hypersurface $q = 0$.

2.10. **The hyperbolic plane.** The most important example of an isotropic quadratic space is the *hyperbolic plane.*[3]

By definition, this is the quadratic space $k^{\oplus 2}$ with the quadratic form $q(x, y) = xy$. We denote the hyperbolic plane as (H, q_H). The matrix for the associated symmetric bilinear form is:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Proposition 2.10.1. *For every non-degenerate isotropic quadratic space (V, q) is isomorphic to $(V', q') \oplus (H, q_H)$ for an appropriate choice of (V', q'). The quadratic space (V', q') is necessarily non-degenerate as well.*

Proof. Let v be an isotropic vector in V. We claim that there is an isotropic vector $w \in V$ such that $B_q(v, w) = 1$.

Indeed, by non-degeneracy of q, there exists $w_0 \in V$ with $B_q(v, w_0) \neq 0$. Therefore, we can define:

$$w := \frac{1}{B_q(v, w_0)} \cdot \big(w_0 - \frac{q(w_0)}{B_q(v, w_0)} \cdot v\big).$$

We then readily verify:

$$q(w) = \frac{1}{B_q(v, w_0)^2} \cdot q(w_0 - \frac{q(w_0)}{B_q(v, w_0)} \cdot v) = \frac{1}{B_q(v, w_0)^2} \cdot \big(q(w_0) - B_q(\frac{q(w_0)}{B_q(v, w_0)} v, w_0)\big) = 0$$

and:

[3] We note that there's no substantive relationship here to the hyperbolic plane in geometry. Rather, H is a plane because it is 2-dimensional, and hyperbolic because its non-zero level sets are hyperbolas.

$$B_q(v,w) = \frac{1}{B_q(v,w_0)} \cdot B_q(v, w_0 - \frac{q(w_0)}{B(v,w_0)}v) = \frac{1}{B_q(v,w_0)} B_q(v,w_0) = 1$$

as desired.

This gives an embedding of the hyperbolic plane into (V,q), and then we obtain the result by Proposition 2.9.7.

□

Corollary 2.10.2. *If q is non-degenerate and isotropic, then every $\lambda \in k$ is represented by q.*

Proof. This follows from Proposition 2.10.1 upon noting that the hyperbolic plane represents every $\lambda \in k$.

□

Corollary 2.10.3. *A non-degenerate quadratic space (V,q) represents $\lambda \in k$ if and only if $(V \oplus k, q_\lambda)$ is isotropic (recall that $q_\lambda(x,\eta) := q(x) - \lambda\eta^2$).*

Proof. The case that (V,q) is anisotropic is treated in Proposition 2.4.5. If (V,q) is isotropic, then by Corollary 2.10.2 (V,q) represents every value.

□

Exercise 2.14. Suppose in (slightly) more generality that (V,q) is an isotropic non-degenerate quadratic space. Does (V,q) necessarily contain a copy of the hyperbolic plane?

2.11. **Diagonalization.** Let (V,q) be a quadratic space.

Definition 2.11.1. A *diagonalization* of (V,q) is a decomposition $(V,q) = (V_1,q_1) \oplus \ldots \oplus (V_n,q_n)$ with $\dim(V_i) = 1$ for all i.

Remark 2.11.2. In terms of quadratic forms in n-variables, a diagonalization of a quadratic form $q(x_1,\ldots,x_n)$ is an equivalence of q with a form q' of the type:

$$q'(y_1,\ldots,y_n) = \sum_{i=1}^{n} a_i y_i^2.$$

Proposition 2.11.3. *Let (V,q) be a quadratic space over k with* $\mathrm{char}(k) \neq 2$. *Then there exists a diagonalization of (V,q).*

Proof. First, suppose that V is non-degenerate. We proceed by induction: the result is tautological if $\dim(V) = 0$.

If $\dim(V) > 0$ there exists $v \in V$ with $q(v) \neq 0$, since otherwise q is identically zero and therefore B_q could not be non-degenerate.

Let $W \subseteq V$ be the subspace $k \cdot v$ spanned by v. Then $W^\perp \cap W = 0$ by assumption on v, and we obviously have a decomposition direct sum decomposition $V = W \oplus W^\perp$. Moreover, for $w \in W$ and $w' \in W^\perp$, we compute:

$$q(w + w') = q(w) + q(w') + B_q(w,w') = q(w) + q(w')$$

and therefore we obtain:

$$V = (W, q|_W) \oplus (W^\perp, q|_{W^\perp})$$

and obtain the result in this case by induction.

In general, let $V' \subseteq V$ be a subspace so that $\mathrm{Rad}(V) \oplus V' \simeq V$. Note that $(V', q|_{V'})$ maps isomorphically onto the quotient $V/\mathrm{Rad}(V)$ equipped with its canonical quadratic space structure, and therefore is non-degenerate by Exercise 2.12. Moreover, we see that $(V, q) \simeq (V', q|_{V'}) \oplus (\mathrm{Rad}(V), 0)$, each of which can be diagonalized.

□

Corollary 2.11.4. *Over an algebraically closed field k with* $\mathrm{char}(k) \neq 2$*, all quadratic spaces of the same rank and dimension are equivalent.*

Exercise 2.15. Show that the hyperbolic plane cannot be diagonalized in characteristic 2. Deduce that Proposition 2.11.3 fails for strongly non-degenerate quadratic spaces in characteristic 2. Where does the argument in Proposition 2.11.3 fail?

Exercise 2.16. Deduce Corollary 2.11.4 from Proposition 2.10.1 and Exercise 2.5.

2.12. **Transposes.** Let (V, q) be a strongly non-degenerate quadratic space and let $S : V \to V$ be any linear transformation (not necessarily a morphism of quadratic spaces).

We define the *q-transpose $S^{T_q} : V \to V$ of S* as the unique linear transformation such that the diagram:

$$\begin{tikzcd}
V \arrow[r, "S^{T_q}"] \arrow[d, "\simeq"] & V \arrow[d, "\simeq"] \\
V^\vee \arrow[r, "S^\vee"] & V^\vee .
\end{tikzcd} \tag{2.2}$$

commutes. Here the vertical maps are defined by B_q. Note that S^{T_q} is uniquely defined because these vertical maps are isomorphisms by the assumption of strong non-degeneracy.

Proposition 2.12.1. *S^{T_q} is the unique linear transformation $V \to V$ satisfying:*

$$B_q(S^{T_q}(v), w) = B_q(v, S(w))$$

for all $v, w \in V$.

Proof. The commutation of the diagram (2.2) says that for every $v \in V$, two certain functionals $V \to k$ should be equal. Evaluating these two functionals on $w \in V$, one finds that the bottom leg of the diagram gives $B_q(v, S(w))$, while the top leg of the diagram gives $B_q(S^{T_q}(v), w)$.

□

Lemma 2.12.2. *In the above setting,* $\det(S) = \det(S^{T_q})$.

Proof. Immediate from the commutation of the diagram (2.3), since this gives $\det(S^{T_q}) = \det(S^\vee)$, and $\det(S^\vee) = \det(S)$.

□

Exercise 2.17. Suppose that $V = k^n$, so that B_q is given by a symmetric matrix A (that is invertible by assumption) and S can be considered as a matrix.

Show that the matrix S^{T_q} is computed as:

$$S^{T_q} = A^{-1} S^T A.$$

Deduce that the q-transpose associated with the quadratic form $\sum_{i=1}^n x_i^2$ is the usual transpose.

2.13. **Orthogonal groups.** Let (V,q) be a quadratic space. We define $O(q)$ as the group of automorphisms of the quadratic space (V,q), i.e., elements of $O(q)$ are isomorphisms $(V,q) \xrightarrow{\simeq} (V,q)$ with multiplication in $O(q)$ defined by composition.

Note that $O(q)$ acts on V.[4]

Example 2.13.1. Let $v \in V$ with $q(v) \neq 0$. Define $s_v \in O(q)$ the *reflection* through v by the formula:

$$s_v(w) = w - \frac{B_q(v,w)}{q(v)} \cdot v.$$

Clearly s_v is a linear transformation. To see that it lies in $O(q)$, we compute:

$$q(s_v(w)) = q(w - \frac{B_q(v,w)}{q(v)} \cdot v) = q(w) + (\frac{B_q(v,w)}{q(v)})^2 q(v) - \frac{B_q(v,w)}{q(v)} B_q(v,w) = q(w).$$

Note that $s_v(v) = -v$, and if $B_q(v,w) = 0$ then $s_v(w) = 0$. Therefore, if $\text{char}(k) \neq 2$ we see that s_v is a non-identity linear transformation with eigenvalue -1 of multiplicity 1 and eigenvalue 1 of multiplicity $\dim(V) - 1$.

Lemma 2.13.2. *Let (V,q) be a strongly non-degenerate quadratic space.*

1. *A linear transformation $g : V \to V$ lies in $O(q)$ satisfies $g^{T_q} g = \text{id}_V$. If $\text{char}(k) \neq 2$, then the converse holds as well.*
2. *For every $g \in O(q)$, $\det(g) = \pm 1$.*

Proof. For (1), we use Proposition 2.12.1 to compute:

$$B_q(g^{T_q} g(v), w) = B_q(g(v), g(w)).$$

We see that the identity $B_q(g(v), g(w)) = B_q(v,w)$ is equivalent to $g^{T_q} g = \text{id}_V$, giving the desired result.

For (2), we compute:

$$1 = \det(\text{id}_V) = \det(g^{T_q} g)) = \det(g^{T_q}) \cdot \det(g) = \det(g)^2$$

using Lemma 2.12.2.

□

Remark 2.13.3. By Example 2.13.1, if $\text{char}(k) \neq 2$ then the map $O(q) \xrightarrow{\det} \{1, -1\}$ is surjective. We let $SO(q)$ denote its kernel, the *special orthogonal group*.

Exercise 2.18. The purpose of this exercise is to show that Lemma 2.13.2 (2) holds even when q is merely assumed to be non-degenerate.

Let (V,q) be a quadratic space.

1. Show that any $g \in O(q)$ fixes $\text{Rad}(V)$ and $\overline{\text{Rad}}(V)$.
2. Suppose that $\det(g|_{\text{Rad}(V)}) \in \{1, -1\}$.[5] Show that $\det(g) \in \{1, -1\}$.
3. Suppose that (V,q) is a non-degenerate quadratic space with $V = \text{Rad}(V)$. Show that $O(q)$ consists only of the identity element.
4. Deduce that (2) holds for q non-degenerate.

[4] In fact, $O(q)$ acts in a way that commutes appropriately with addition and scaling operations in V, i.e., its action on V arises through a homomorphism $O(q) \to GL(V)$. In this situation, one says that V is a *representation* of $O(q)$.

[5] The notation is a bit lazy. For clarity: in characteristic 2 we regard $\{1, -1\}$ as consisting of one element.

2.14. **Spheres.** Let (V, q) be a quadratic space and fix $r \in k$.

We define the associated *sphere* $\mathbb{S}_r(q)$ as $\{x \in V \mid q(x) = r\}$.

Example 2.14.1. Let $k = \mathbb{R}$ and let $q = \sum_{i=1}^{n+1} x_i^2$. Then $\mathbb{S}_r(q)$ is the usual n-sphere of radius $\sqrt{r}$.

Note that the action of $O(q)$ on V preserves $\mathbb{S}_r(q)$.

Proposition 2.14.2. *Let (V, q) be a non-degenerate quadratic space and suppose* $\mathrm{char}(k) \neq 2$. *Then for every $r \neq 0$, the action of $O(q)$ on $\mathbb{S}_r(q)$ is transitive.*

Proof. Let $v, w \in \mathbb{S}_r(q)$.

We claim that either $v + w$ or $v - w$ is not isotropic. Indeed, we otherwise have:

$$0 = q(v + w) - q(v - w) = q(v) + q(w) + B_q(v, w) - q(v) - q(w) + B_q(v, w) = 2B_q(v, w)$$

so that $B_q(v, w) = 0$. Then $q(v + w) = q(v) + q(w) = 2r \neq 0$.

Suppose that $v - w$ is not isotropic. Then we can make sense of the reflection $s_{v-w} \in O(q)$. We claim that $s_{v-w}(v) = w$.

First, note that $B_q(v + w, v - w) = 0$ since:

$$B_q(v + w, v - w) = 2q(v) - 2q(w) + B_q(v, w) - B_q(v, w) = 0.$$

Therefore, we have:

$$\begin{aligned} v + w &= s_{v-w}(v + w) = s_{v-w}(v) + s_{v-w}(w) \\ -v + w &= s_{v-w}(v - w) = s_{v-w}(v) - s_{v-w}(w). \end{aligned}$$

Adding these equations, we obtain $s_{v-w}(v) = w$ as desired.

Otherwise, $v + w$ is anisotropic, and the above gives that $s_{v+w}(v) = -w$; composing s_{v+w} with $-\mathrm{id}_V \in O(q)$, we obtain the result.

□

Exercise 2.19. In the above setting, suppose that $\mathrm{rank}(q) > 1$. Does $SO(q)$ act transitively on $\mathbb{S}_r(q)$?

2.15. **Witt's cancellation theorem.** Suppose that we have an identification:

$$(V, q_V) \oplus (W_1, q_{W_1}) \simeq (V, q_V) \oplus (W_2, q_{W_2})$$

of quadratic spaces. What can we deduce about the relationship between (W_1, q_{W_1}) and (W_2, q_{W_2})?

Witt's theorem addresses this problem completely.

Theorem 2.15.1 (Witt)**.** *Suppose that* $\mathrm{char}(k) \neq 2$. *Then in the above situation, (W_1, q_{W_1}) and (W_2, q_{W_2}) are isomorphic.*

Proof. We proceed by steps.

Step 1. First, we reduce to the case that q_V is non-degenerate.

Indeed, choosing a complementary subspace in V to $\mathrm{Rad}(V)$, we see as in the proof of Proposition 2.11.3 that V is the direct sum of its radical and a non-degenerate quadratic space, giving the reduction.

Step 2. Next, we reduce to the case where q_{W_1} and q_{W_2} are non-degenerate.

One immediately finds (supposing that q_V is non-degenerate) that $\mathrm{Rad}(V \oplus W_1) = \mathrm{Rad}(W_1)$ and similarly for W_2. Therefore, quotienting out by the radical of $V \oplus W_1 = V \oplus W_2$, we find that:

$$(V, q_V) \oplus (W_1/\operatorname{Rad}(W_1), q_{W_1/\operatorname{Rad}(W_1)}) \simeq (V, q_V) \oplus (W_2/\operatorname{Rad}(W_2), q_{W_2/\operatorname{Rad}(W_2)})$$

Therefore, if we know the result for triples of non-degenerate spaces, we can deduce that:

$$(W_1/\operatorname{Rad}(W_1), q_{W_1/\operatorname{Rad}(W_1)}) \simeq (W_2/\operatorname{Rad}(W_2), q_{W_2/\operatorname{Rad}(W_2)}).$$

Since e.g (W_1, q_{W_1}) is the direct sum of its radical and its quotient by its radical, and since the radicals of W_1 and W_2 have been identified, we obtain the result in the general case.

Step 3. Choose $v \in V$ with $q(v) \neq 0$, and let V' be the orthogonal complement to $k \cdot v$, so that $V = V' \oplus \operatorname{span}(v)$ as a quadratic space.

Under the isomorphism $T : V \oplus W_1 \xrightarrow{\simeq} V \oplus W_2$, we have $q_{V\oplus W_2}(T(v, 0)) = q_V(v) = q_{V\oplus W_2}(v, 0)$. Therefore, by Proposition 2.14.2, there is an automorphism of $(V \oplus W_2, q_{V\oplus W_2})$ taking $T(v)$ to $(v, 0)$.

This automorphism induces an equivalence between the orthogonal complement to (the span of) $T(v)$ and $(v, 0)$. The latter is obviously $V' \oplus W_2$. Using T, we see that the former is isomorphic to $V' \oplus W_1$.

We now obtain the result by induction.

□

Corollary 2.15.2 (Sylvester's theorem)**.** *Every quadratic form $q(x_1, \ldots, x_n)$ over $\mathbb{R}$ of rank r is equivalent to $\sum_{i=1}^{j} x_i^2 - \sum_{i=j+1}^{r} x_i^2$ for a unique choice of $1 \leqslant j \leqslant r$.*

Proof. Diagonalizing q, it is clear that q is equivalent to such a form. Uniqueness then follows by induction from Witt's theorem.

□

Definition 2.15.3. For q a quadratic form over $\mathbb{R}$ equivalent to $\sum_{i=1}^{j} x_i^2 - \sum_{i=j+1}^{r} x_i^2$, the pair $(j, r - j) \in \mathbb{Z}^2$ is called the *signature* of q.

2.16. **Tensor products.** We suppose in §2.16-2.18 that $\operatorname{char}(k) \neq 2$.

Construction 2.16.1. Suppose that (V, q_V) and (W, q_W) are quadratic spaces. Recall that $q_{\frac{1}{2}B_{q_V}} = q_V$, and similarly for q_W.

We define a quadratic space structure $q_{V\otimes W}$ on $V \otimes W$ by associating it to the bilinear form:

$$V \otimes W \to (V \otimes W)^\vee = V^\vee \otimes W^\vee$$

obtained by tensor product of the maps $V \to V^\vee$ and $W \to W^\vee$ defined by $\frac{1}{2}B_{q_V}$ and $\frac{1}{2}B_{q_W}$.

Exercise 2.20. Show that the bilinear form $B_{q_{V\otimes W}}$ satisfies:

$$B_{q_{V\otimes W}}(v_1 \otimes w_1, v_2 \otimes w_2) = \frac{1}{2} B_{q_V}(v_1, v_2) B_{q_W}(w_1, w_2).$$

Deduce that $q_{V\otimes W}(v \otimes w) = q_V(v) q_W(w)$.

Exercise 2.21. Show that $(V, q_V) \otimes (k, q(x) := x^2) \simeq (V, q_V)$.

Exercise 2.22. Go to the Wikipedia page for "Kronecker product" and figure out the relation to this construction.

2.17. **Alternating and symmetric powers.** Let $n \geqslant 0$ be an integer and let (V, q_V) be a quadratic space. We claim that there are natural quadratic space structures on $\mathrm{Sym}^n(V)$ and $\Lambda^n(V)$. (We remind that we have assumed that $\mathrm{char}(k) \neq 2$).

First, note that (in any characteristic) we have canonical "norm" maps:

$$\mathrm{Sym}^n(V) \to V^{\otimes n} \text{ and } \Lambda^n(V) \to V^{\otimes n}.$$

The former is induced by the (unnormalized) symmetrization map:

$$\begin{aligned} V^{\otimes n} &\to V^{\otimes n} \\ v_1 \otimes \ldots \otimes v_n &\mapsto \sum_{\sigma \in S_n} v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)} \end{aligned}$$

and the latter is induced similarly by:

$$\begin{aligned} V^{\otimes n} &\to V^{\otimes n} \\ v_1 \otimes \ldots \otimes v_n &\mapsto \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)}. \end{aligned}$$

We remark that each composition:

$$\begin{gathered} \mathrm{Sym}^n(V) \to V^{\otimes n} \to \mathrm{Sym}^n(V) \\ \Lambda^n(V) \to V^{\otimes n} \to \Lambda^n(V) \end{gathered}$$

is multiplication by $n! = |S_n|$.

Exercise 2.23. What is the relationship between the above constructions and the connection between symmetric bilinear forms and quadratic forms?

We now obtain a quadratic form $q_{\mathrm{Sym}^n(V)}$ (resp. $q_{\Lambda^n(V)}$) on $\mathrm{Sym}^n(V)$ and $\Lambda^n(V)$ by restriction of $q_{V^{\otimes n}}$ (as defined by Construction 2.16.1) along these homomorphisms.

The following important exercise gives the interaction between symmetric and alternating powers and duality, and then describes the relation to the above constructions.

Exercise 2.24. The following exercises (except (3)) work for both Sym^n and Λ^n: we write them for Sym^n for definiteness.

(1) Show that the map $\mathrm{Sym}^n(V)^\vee \to V^{\otimes,\vee}$ dual to the structure map $V^{\otimes n} \to \mathrm{Sym}^n(V)$ is injective and has image the subspace of functionals on $V^{\otimes n}$ invariant under the action of the symmetric group.
(2) Show that the norm map $\mathrm{Sym}^n(V^\vee) \to V^{\vee,\otimes n}$ factors through the subspace $\mathrm{Sym}^n(V)^\vee$.
(3) Show that the maps $\mathrm{Sym}^n(V^\vee) \to \mathrm{Sym}^n(V)^\vee$ are isomorphisms if and only if $\mathrm{char}(k) = 0$ or $n < \mathrm{char}(k)$, but that the maps $\Lambda^n(V^\vee) \to \Lambda^n(V)^\vee$ are always isomorphisms whenever $\mathrm{char}(k) \neq 2$.
(4) Suppose now that $V \to V^\vee$ is a symmetric bilinear form. Consider the induced map:

$$\mathrm{Sym}^n(V) \to \mathrm{Sym}^n(V)^\vee \tag{2.3}$$

defined as the composition:

$$\mathrm{Sym}^n(V) \to \mathrm{Sym}^n(V^\vee) \to \mathrm{Sym}^n(V)^\vee$$

with the first map given by functoriality[6] of Sym^n and the second map is the one constructed in (2) above.

Show that if q_V is the induced quadratic form on V, then $q_{\mathrm{Sym}^n(V)}$ is induced by (2.3).

2.18. **The discriminant.** Suppose that (V, q_V) is a quadratic space and $\mathrm{char}(k) \neq 2$.

Recall that $\det(V) := \Lambda^{\dim(V)}(V)$ is 1-dimensional. By §2.17, we obtain a quadratic space structure on this 1-dimensional vector space.

Lemma 2.18.1. *The quadratic form on* $\det(V)$ *is non-zero if and only if* (V, q) *is non-degenerate.*

Proof. Let $B : V \to V^\vee$ be defined by the symmetric bilinear structure on V giving rise to q_V. By Exercise 2.24, the induced map $\det(B) : \det(V) \to \det(V^\vee) = \det(V)^\vee$ is the symmetric bilinear form giving rise to $q_{\det(V)}$. But B is an isomorphism if and only if $\det(B)$ is, and since $B = \frac{1}{2}B_q$ we obtain the result.

□

Suppose now that (V, q) is a non-degenerate quadratic space. By Lemma 2.18.1, we obtain that $(\det(V), q_{\det(V)})$ is a non-degenerate 1-dimensional quadratic space.

Obviously isomorphism classes of non-degenerate 1-dimensional quadratic forms are in bijection with the set $k^\times/(k^\times)^2$, where the quadratic form $q(x) = ax^2, a \neq 0$ has invariant the class of a in $k^\times/(k^\times)^2$.

Definition 2.18.2. For (V, q) non-degenerate as above, the *discriminant* $\mathrm{disc}(q) \in k^\times/(k^\times)^2$ of V is the invariant of $(\det(V), q_{\det(V)})$ as defined above.

Now let us explain how to concretely compute the discriminant.

Proposition 2.18.3. *Suppose that A is an invertible symmetric $(n \times n)$-matrix and let q be the associated quadratic form $q(x) = x^T Ax$ in n-variables.*

Then $\mathrm{disc}(q) = \det(A) \bmod (k^\times)^2$.

Proof. By construction the map $\det(V) \to \det(V)^\vee$ defined by q is given as multiplication by $\det(A)$. Therefore, the induced quadratic form in 1-variable is $x \mapsto \det(A) \cdot x^2$ as desired.

□

Remark 2.18.4. Recall from Exercise 2.9 that matrices A and A' define equivalent quadratic forms if and only if $A' = S^T AS$, in which case:

$$\det(A') = \det(S^T AS) = \det(S^T)\det(A)\det(S) = \det(S)^2\det(A)$$

showing directly that $\det(A)$ considered modulo $(k^\times)^2$ is a well-defined invariant of the underlying quadratic space.

Remark 2.18.5. Suppose that q has been diagonalized and we realize q as equivalent to a form $\sum_{i=1}^n a_i x_i^2$. We see that $\mathrm{disc}(q) = \prod a_i \bmod (k^\times)^2$.

Exercise 2.25. Let $k = \mathbb{F}_p$ with $p \neq 2$ and let q_1 and q_2 be the binary forms $q_1(x, y) = x^2 + y^2$ and $q_2(x, y) = x^2 + ry^2$ for $r \in \mathbb{F}_p^\times$ a quadratic non-residue (i.e., $\left(\frac{r}{p}\right) = -1$).

(1) Show that each of q_1 and q_2 represents every element of $\mathbb{F}_p$.
(2) Show that q_1 and q_2 are inequivalent quadratic forms.

[6] *Functorial* is a scientific word that here refers to the obvious map $\mathrm{Sym}^n(V) \to \mathrm{Sym}^n(W)$ induced by a given linear transformation $T : V \to W$ (similarly, Λ^n is functorial.)

Exercise 2.26. Use discriminants to prove Witt's theorem Theorem 2.15.1 in the case that (V, q) is non-degenerate and $\dim W_1 = \dim W_2 = 1$.

Exercise 2.27. (1) Show that the discriminant of the hyperbolic plane is -1.
(2) Show that if (V, q) is a non-degenerate quadratic space with $\dim(V) = 2$ and $\mathrm{disc}(q) = -1$, then V is equivalent to the hyperbolic plane.

3. p-adic fields

3.1. Recall from §2.4 that 1-dimensional quadratic forms over $\mathbb{Q}$ are of the form $q(x) = ax^2$, $a \in \mathbb{Q}$, where the equivalence class of q is determined by the sign of a and the parity of $v_p(a)$ for each prime p, where $v_p(a)$ measures the degree to which p divides a.

We want to formulate this information in a more systematic way so that we can generalize it to forms q of higher rank. To generalize the sign of a, we use $q_{\mathbb{R}}$ the extension of scalars of q along $\mathbb{Q} \hookrightarrow \mathbb{R}$.

For each prime p, we will introduce the field $\mathbb{Q}_p \supset \mathbb{Q}$ of *p-adic numbers* such that the extension of scalars $q_{\mathbb{Q}_p}$ generalizes $v_p(a)$ from the rank 1 case.

3.2. We will give several perspectives on $\mathbb{Q}_p$ below: as a metric completion of $\mathbb{Q}$, as a (field of fractions of) a kind of limit of the rings $\mathbb{Z}/p^n\mathbb{Z}$, and as infinite base p exansions of numbers.

3.3. **Analytic approach.** Recall that for $a \in \mathbb{Q}^\times$, we have defined in §2.4 the *p-adic valuation* $v_p(a)$ as the unique integer such that $a = p^{v_p(a)}\frac{n}{m}$ with n and m not divisible by p. We extend this definition to $\mathbb{Q}$ by setting $v_p(0) = \infty$.

Lemma 3.3.1. *(1) For $a, b \in \mathbb{Q}$, we have $v_p(ab) = v_p(a) + v_p(b)$.*
(2) For $a, b \in \mathbb{Q}$, we have $\min\{v_p(a), v_p(b)\} \leqslant v_p(a + b)$.

Proof. The first part is obvious. If $a, b \in \mathbb{Z}$ with $a = p^{v_p(a)}a'$, $b = p^{v_p(b)}b'$, then clearly $p^{\min\{a,b\}}$ divides $a + b$ giving the second part in this case. We can deduce it in general by either the same method, or by clearing denominators and applying the first part.

□

Example 3.3.2. We have $v_p\big(1 + (p-1)\big) = 1 > 0 = \max\{v_p(1), v_p(p-1)\}$, giving an example in which the inequality in Lemma 3.3.1 (2) is strict.

We then define the *p-adic absolute norm* $|a|_p \in \mathbb{R}$ of $a \in \mathbb{Q}$ as $p^{-v_p(a)}$, where for $a = 0$ we interpret $p^{-\infty}$ as 0.

Remark 3.3.3. The function $|\cdot|_p : \mathbb{Q} \to \mathbb{R}$ can only take the values in $\{p^{\mathbb{Z}}\} \cup \{0\} = \{\ldots, \frac{1}{p^2}, \frac{1}{p}, 1, p, p^2, \ldots\} \cup \{0\}$.

Warning 3.3.4. We have $|p|_p = p^{-1}$.

From Lemma 3.3.1, we immediately deduce:

Lemma 3.3.5. *We have $|ab|_p = |a|_p|b|_p$ and $|a + b|_p \leqslant \max\{|a|_p, |b|_p\}$.*

Note that $d(x, y) := |x - y|_p$ defines a metric on $\mathbb{Q}$. Indeed, by the lemma we have:

$$|x - y|_p + |y - z|_p \geqslant \max\{|x - y|_p, |y - z|_p\} \geqslant |x - z|_p$$

so that $d(x, y)$ satisfies a strong version of the triangle inequality. We therefore refer to the inequality $|a + b|_p \leqslant \max\{|a|_p, |b|_p\}$ as the *ultrametric inequality*.

We define $\mathbb{Q}_p$ to be the completion of $\mathbb{Q}$ with respect to this metric.

Exercise 3.1.
(1) Show that addition and multiplication on $\mathbb{Q}$ are continuous with respect to the p-adic metric. Deduce that $\mathbb{Q}_p$ is a commutative ring with $\mathbb{Q}$ a dense subring.
(2) Show that the inversion map $\mathbb{Q}^\times \to \mathbb{Q}^\times$, $x \mapsto x^{-1}$ is a continuous homomorphism with respect to the p-adic metric. Deduce that $\mathbb{Q}_p$ is a field.

We see that v_p and $|\cdot|_p$ extend to $\mathbb{Q}_p$ by continuity; we denote these extensions in the same manner. Since on $\mathbb{Q}$ these functions took values in the closed subsets $\mathbb{Z} \subseteq \mathbb{R}$ and $p^{\mathbb{Z}} \cup \{0\} \subseteq \mathbb{R}$ respectively, their extensions to $\mathbb{Q}_p$ do as well.

We will also use the notation $p^n | a$ to say that $v_p(a) \geqslant n$.

Notation 3.3.6. By Lemma 3.3.5, we see that $\{a \in \mathbb{Q}_p \mid |a|_p \leqslant 1\}$ forms a subring of $\mathbb{Q}_p$. We deduce this subring by $\mathbb{Z}_p$ and call it the *ring of p-adic integers*.

Note that $x \in \mathbb{Z}_p$ if and only if $v_p(x) \geqslant 0$.

Exercise 3.2.
(1) Show that $\mathbb{Z}_p$ is open and closed in $\mathbb{Q}_p$.
(2) Show that a rational number $\frac{n}{m} \in \mathbb{Q} \subseteq \mathbb{Q}_p$ lies in $\mathbb{Z}_p$ if and only if when n and m are chosen relatively prime p does not divide m.
(3) Show that the image of $\mathbb{Z} \hookrightarrow \mathbb{Z}_p$ has dense image.
(4) Show that the field of fractions of $\mathbb{Z}_p$ is $\mathbb{Q}_p$. More precisely, show that $\mathbb{Q}_p$ is obtained from $\mathbb{Z}_p$ by inverting the single element p.

Exercise 3.3. Prove the *product formula*, which states that for $a \in \mathbb{Q}$ we have:

$$|a|_\infty \cdot \prod_{p \text{ a prime}} |a|_p = 1$$

where $|\cdot|_\infty$ is the "usual" absolute value on $\mathbb{Q}$, whose completion is $\mathbb{R}$.

Exercise 3.4. Show that if $x \in \mathbb{Q}_p$ is a root of a monic polynomial:

$$x^n + a_1 x^{n-1} + \ldots + a_n$$

with $a_i \in \mathbb{Z}_p$, then $x \in \mathbb{Z}_p$.

3.4. Infinite summation in $\mathbb{Q}_p$.

We briefly digress to discuss infinite sums in $\mathbb{Q}_p$.

We say that a sequence $x_1, x_2, \ldots$ has *convergent sum* if the sequence $S_n := \sum_{i=1}^n x_i$ converges in $\mathbb{Q}_p$. In this case, we write $\sum_{i=1}^\infty$ for the limit of the sequence S_n. We will summarize the situation informally by saying that the sum $\sum_{i=1}^\infty x_i$ converges.

Lemma 3.4.1. *The sum $\sum_{i=1}^\infty x_i$ converges if and only if $v_p(x_i) \to \infty$ as $i \to \infty$.*

Proof. To check if the sequence S_n of partial sums is convergent, it suffices to check if it is Cauchy, which in turn translates to verifying that for every $\varepsilon > 0$ there exists N such that for every $m > N$ we have:

$$|\sum_{i=N}^m x_i|_p < \varepsilon.$$

By the ultrametric inequality, we have $|\sum_{i=N}^m x_i|_p \leqslant \max\{|x_i|_p\}_{i=N}^m$, giving the result. □

Remark 3.4.2. This result stands in stark contrast to the case of $\mathbb{R}$, where convergence of an infinite series is a much more subtle question.

3.5. **Some computations.** We'll now do some computations to get a feeling for working with $\mathbb{Q}_p$.

A first observation is that $1, p, p^2, p^3, \ldots$ converges to 0 in $\mathbb{Q}_p$. Indeed, we have $|p^n|_p = p^{-n}$, giving the result.

More generally, we see that a sequence $\{x_n\}_{n \geqslant 0}$ converges to 0 if and only if for every $N > 0$ there exists M such that for $n > M$ we have $x_n \in p^N \mathbb{Z}_p$. Indeed, $p^N \mathbb{Z}_p$ is the open (and closed) neighborhood of the identity consisting of all $x \in \mathbb{Q}_p$ with $|x|_p \leqslant p^{-N}$.

Example 3.5.1. Note that the sequence $x_n := 1 + p + p^2 + \ldots + p^n$ is Cauchy and therefore converges in $\mathbb{Q}_p$. The usual argument shows that this element is the inverse to $1 - p$ in $\mathbb{Q}_p$. Clearly each of $1 - p$ and $1 + p + p^2 + \ldots$ lie in $\mathbb{Z}_p$.

Next, observe that $p^n \mathbb{Z}_p$ is an ideal of $\mathbb{Z}_p$ with:

$$\mathbb{Z}/p^n\mathbb{Z} \xrightarrow{\simeq} \mathbb{Z}_p/p^n\mathbb{Z}_p.$$

Indeed, this follows since $\mathbb{Z}$ is dense in $\mathbb{Z}_p$ and $p^n\mathbb{Z}_p$ is an open ideal.

Proposition 3.5.2. *The topological space $\mathbb{Z}_p$ is compact.*

Remark 3.5.3. Note that this proposition does not hold for $\mathbb{Q}_p$, since the sequence p^{-n} $(n \geqslant 0)$ does not converge to a limit.

Proof of Proposition 3.5.2. We need to show that every sequence x_n contains a convergent subsequence y_m.

There must be some residue class a_1 in $\mathbb{Z}/p\mathbb{Z}$ such that $x_n \bmod p\mathbb{Z}_p$ takes that value infinitely often. Choose $y_1 = x_{n_1}$ to be some element in our sequence that reduces to this element.

Of the classes in $\mathbb{Z}/p^2\mathbb{Z}$, there must be at least one that reduces to a_1 infinitely often. Call one such as a_2. Then there must be an index $n_2 > n_1$ such that $y_2 := x_{n_2}$ reduces to a_2 modulo $p^2\mathbb{Z}_p$.

Repeating this for $\mathbb{Z}/p^3\mathbb{Z}$, etc., we obtain a sequence y_n that is obviously Cauchy and therefore convergent.

□

3.6. **The p-adic numbers as a limit.** Suppose we have a sequence:

$$\ldots \to A_{n+1} \xrightarrow{\varphi_n} A_n \xrightarrow{\varphi_{n-1}} \ldots \xrightarrow{\varphi_1} A_1$$

of commutative rings A_n. We call such a datum a *projective system.*

Define the *projective limit* $\lim_n A_n$ as the subset of the product $\prod_n A_n$ consisting of elements $(a_n) \in \prod_n A_n$ such that for each $i > 1$ we have $\varphi_{n-1}(a_n) = a_{n-1}$. That is, elements of $\lim_n A_n$ are elements of A_1 with a chosen lift to A_2, and a chosen lift of that element of A_2 to A_3, etc. The commutative algebra structure on $\lim_n A_n$ is defined by termwise addition and multiplication.

Example 3.6.1. We have a projective system with $A_n = \mathbb{Z}/p^n\mathbb{Z}$, where the structure map $\mathbb{Z}/p^{n+1}\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ is given by modding out by the ideal $p^n\mathbb{Z}/p^{n+1}\mathbb{Z}$.

Remark 3.6.2. We can also make sense of projective limits of groups, projective limit of abelian groups, projective limit of rings, etc.

Proposition 3.6.3. *The canonical map:*

$$\mathbb{Z}_p \to \lim_n \mathbb{Z}_p/p^n\mathbb{Z}_p = \lim_n \mathbb{Z}/p^n\mathbb{Z}$$

is an isomorphism.

Proof. First, note that the resulting map is injective: an element $a \in \mathbb{Z}_p$ goes to zero if and only if it lies in $p^n\mathbb{Z}_p$ for each n, but this implies that the $|a|_p \leqslant p^{-n}$ for each n, so $|a|_p = 0$ which implies that $a = 0$.

To see that this map is surjective, let (a_n) be an element of the right hand side, we $a_n \in \mathbb{Z}/p^n\mathbb{Z}$ a compatible sequence. Choose $x_n \in \mathbb{Z}$ reducing to a_n modulo p^n. Note that for $n, m \geqslant N$, we have $p^N \mid x_n - x_m$, so that $|x_n - x_m|_p \leqslant p^{-N}$. Therefore, this sequence is Cauchy and therefore converges to some $x \in \mathbb{Z}_p$.

Note that $x - x_n \in p^n\mathbb{Z}_p$ since $x_m - x_n \in p^n\mathbb{Z}$ for all $m \geqslant n$. Therefore, $x - x_n \in p^n\mathbb{Z}_p$, so that $x = a_n \bmod p^n\mathbb{Z}_p$, as desired.

□

Remark 3.6.4. One upshot of this construction of $\mathbb{Z}_p$ is that it is *algebraic*, i.e., it makes no reference to the real numbers (whereas the very notion of metric space does). Note that we can obtain $\mathbb{Q}_p$ from $\mathbb{Z}_p$ by inverting p, so this construction gives an algebraic perspective on $\mathbb{Q}_p$ as well

3.7. **Infinite base p expansion.** We have the following lemma that gives a concrete way of thinking about p-adic numbers.

Proposition 3.7.1. *For every $x \in \mathbb{Q}_p$, there exist unique integers $0 \leqslant a_i < p$ defined for $i \geqslant v_p(x)$ such that:*

$$x = \sum_{i=v_p(x)} a_i p^i.$$

The coefficient $a_{v_p(x)}$ is non-zero. We have $a_i = 0$ for all i sufficiently large if and only if $p^{v_p(x)} x \in \mathbb{Z} \subseteq \mathbb{Z}_p$.

Proof. First, by multiplying by $p^{v_p(x)}$, we reduce to the case when $x \in \mathbb{Z}_p$.

Reduce x modulo p^n; there is a unique integer $0 \leqslant x_n < p^n$ with $x_n = x \bmod p^n\mathbb{Z}_p$. Take the base p expansion of x_n:

$$x_n = \sum_{i=0}^{n-1} a_i p^i.$$

It is clear that the coefficients $a_i, 0 \leqslant i < n$ are the same when we work with x_m for $m \geqslant 0$ instead. Moreover, the infinite sum obviously converges to x, giving the existence of such an expression.

The remaining properties are easily verified.

□

Remark 3.7.2. The proof shows that we could replace the condition that a_i are integers between 0 and $p - 1$ by the condition that the $a_i \in S$ where $S \subseteq \mathbb{Z}_p$ is a subset of coset representatives of $\mathbb{Z}/p\mathbb{Z}$.

A common choice one finds in the literature instead of $S = \{0, 1, \ldots, p-1\}$ is that S is the set of *Teichmuller lifts* of elements of $\mathbb{F}_p$: these are defined in Remark 3.10.6. The Teichmuller normalization is important in the theory of Witt vectors.

3.8. **Invertibility in $\mathbb{Z}_p$.** Here we will give several perspectives on the following important result.

Proposition 3.8.1. *An element $x \in \mathbb{Z}_p$ is invertible in $\mathbb{Z}_p$ if and only if $p \nmid x$.*

First proof of Proposition 3.8.1. Suppose $x \in \mathbb{Z}_p$ is non-zero, so we can make sense of $x^{-1} \in \mathbb{Q}_p$. Then we have:

$$0 = v_p(1) = v_p(xx^{-1}) = v_p(x) + v_p(x^{-1}).$$

Since by hypothesis $v_p(x) \geqslant 0$, we see that $x^{-1} \in \mathbb{Z}_p$ if and only if:

$$v_p(x) = v_p(x^{-1}) = 0$$

as desired.

□

We will give two additional and more explicit proofs below modeled on Example 3.5.1. Though one argument will be analytic and the other algebraic, we note from the onset that they are two essentially identical arguments packaged in different ways.

Second proof of Proposition 3.8.1. First, suppose that $x = 1 \bmod p\mathbb{Z}_p$.

Write $x = 1 - (1 - x)$ so that we heuristically expect:

$$\frac{1}{x} = \frac{1}{1-(1-x)} = \sum_{i=0}^{\infty}(1-x)^i$$

Now observe that $p^i \mid (1-x)^i$ by assumption on x, so that the series on the right converges. We then easily compute:

$$(1-x) \cdot \sum_{i=0}^{\infty}(1-x)^i = \sum_{i=1}^{\infty}(1-x)^i = \big(\sum_{i=0}^{\infty}(1-x)^i\big) - 1$$

so that on subtracting $\sum_{i=0}^{\infty}(1-x)^i$ from each side we find:

$$x \cdot \sum_{i=0}^{\infty}(1-x)^i = 1$$

as desired.

In general, choose $y \in \mathbb{Z}$ such that $xy = 1 \bmod p\mathbb{Z}_p$: such y exists because $x \neq 0 \bmod p\mathbb{Z}_p$. Then $\frac{1}{y} \in \mathbb{Q} \subseteq \mathbb{Q}_p$ obviously lies in $\mathbb{Z}_p$, so we see that y is invertible on the one hand and xy is too, giving the result.

□

Remark 3.8.2. Here's another presentation of the above argument: as in the proof, we reduce to the case where $x = 1 \bmod p\mathbb{Z}_p$. Then the operator $T : \mathbb{Z}_p \to \mathbb{Z}_p$ defined as $T(y) = (1-x)y + 1$ multiplication by $1-x$ is contracting, since $|T(y_1) - T(y_2)|_p \leqslant \frac{1}{p} \cdot |y_1 - y_2|_p$. Therefore, by Banach's fixed point theorem, there exists $x_0 \in \mathbb{Z}_p$ with $x_0 = T(x_0) = x_0(1-x) + 1$ meaning that $x_0 x = 1$ as desired.

For the third proof, we use the following result.

Lemma 3.8.3. *(1) Let A be a commutative ring, let $x \in A$ be invertible, and let $y \in A$ be nilpotent (i.e., $y^N = 0$ for N large enough).*

Then $x + y$ is invertible.

(2) Let A be a commutative ring and let I be an ideal consisting only of nilpotent elements of A.

Then $x \in A$ is invertible if and only if its reduction $\overline{x} \in A/I$ is invertible.

Proof. For (1), we explicitly compute:

$$\frac{1}{x+y} = \frac{x^{-1}}{1+\frac{y}{x}} = \frac{1}{x}\sum_{i=0}^{\infty}(-1)^i\big(\frac{y}{x}\big)^i$$

where the infinite sum makes sense because y is nilpotent, i.e., it is really only a finite sum.

For (2): first, note that if x is invertible in A, then it certainly is modulo I as well.

For the converse, choose some $y \in A$ such that $xy = 1 \bmod I$. Then xy is invertible by (1), since it is of the form $1+$(an element of I), and every element of I is nilpotent. We we find that $(xy)^{-1} \cdot y$ is the inverse to x.

□

Third proof of Proposition 3.8.1. To show that x is invertible, it suffices to show that $x \bmod p^n\mathbb{Z}_p$ is invertible for each n: indeed, the inverses clearly form a compatible system and therefore define an element of $\mathbb{Z}_p = \lim \mathbb{Z}/p^n\mathbb{Z}$.

But for each $n > 1$, the ideal $p^{n-1}\mathbb{Z}/p^n\mathbb{Z}$ consists only of nilpotent elements, so to test if an element in $\mathbb{Z}/p^n\mathbb{Z}$ is invertible, it suffices by Lemma 3.8.3 (2) to check that its reduction modulo p is.

□

3.9. **Algebraic structure of the unit group.** We want to describe $\mathbb{Q}_p^\times$ in more detail. First, note that we have an isomorphism:

$$\begin{aligned} \mathbb{Q}_p^\times &\xrightarrow{\simeq} \mathbb{Z} \times \mathbb{Z}_p^\times \\ x &\mapsto \big(v_p(x), p^{-v_p(x)} \cdot x\big). \end{aligned} \tag{3.1}$$

Therefore, it suffices to understand $\mathbb{Z}_p^\times$.

For each $n > 0$, note that:

$$\operatorname{Ker}\big(\mathbb{Z}_p^\times \to (\mathbb{Z}_p/p^n\mathbb{Z}_p)^\times\big)$$

is the subgroup $1 + p^n\mathbb{Z}_p$, noting that subgroup indeed lies in $\mathbb{Z}_p^\times$ by Proposition 3.8.1. Moreover, the resulting map is surjective because

Proposition 3.9.1. *The map:*

$$\mathbb{Z}_p^\times \to \lim_n\, (\mathbb{Z}_p/p^n\mathbb{Z}_p)^\times$$

is an isomorphism.

Proof. If $x \in \mathbb{Z}_p^\times$ is in the kernel of this map, then $x \in 1 + p^n\mathbb{Z}_p$ for all $n > 0$, meaning that x is arbitrarily close to 1 and therefore equals 1 as desired. Surjectivity is clear from Proposition 3.6.3: any element in the projective limit arises from an element of $\mathbb{Z}_p$ that obviously lies in $\mathbb{Z}_p^\times$.

□

Remark 3.9.2. The p-adic exponential and logarithm functions introduced below shed further light on $\mathbb{Z}_p^\times$.

3.10. **Hensel's lemma.** Hensel's lemma is a generalization Proposition 3.8.1, giving solutions to equations in $\mathbb{Z}_p$ by checking for their reduction modulo p.

Definition 3.10.1. Let A be a ring and let $f(t) = \sum_{i=0}^{n} a_i \cdot t^i \in A[t]$ be a polynomial. We define the *formal derivative* of f as:

$$f'(t) = \sum_{i=1}^{n} i a_i t^{i-1} \in A[t]$$

Proposition 3.10.2 (Hensel's lemma). *Let $f(t) \in \mathbb{Z}_p[t]$ be a polynomial with coefficients in $\mathbb{Z}_p$. Let $\overline{f}(t) \in \mathbb{F}_p[t]$ be the polynomial induced by reducing the coefficients of f modulo p.*

Suppose that $\overline{f}(t)$ has a root $\overline{x}$ with $\overline{f}'(t)(\overline{x}) \neq 0 \in \mathbb{F}_p$, where $\overline{f}'$ is the formal derivative of $\overline{f}$.[7]

Then there exists a unique root $x \in \mathbb{Z}_p$ of f that reduces to $\overline{x}$.

Remark 3.10.3. The hypothesis on the derivative of f is necessary: otherwise, consider $f(t) = t^2 - p$. Since $\overline{f}(t) = t^2$, there exists a solution modulo p, but $\mathbb{Q}_p$ does not contain a square root of p: indeed, we would have $v_p(\sqrt{p}) = \frac{1}{2}v_p(p) = \frac{1}{2}$, and v_p can only take integral values.

Proof of Proposition 3.10.2. We use the notation x_1 in place of $\overline{x}$. For each $n > 0$, we let $f_n(t) \in \mathbb{Z}/p^n\mathbb{Z}[t]$ denote the corresponding reduction of f.

We will prove the following statement by induction:

$(*)_n$: There exists a unique root $x_n \in \mathbb{Z}/p^n\mathbb{Z}$ of f_n lifting x_1.

Clearly this statement suffices, since the x_n obviously form a compatible system and therefore define the desired x in $\mathbb{Z}_p$.

Observe that the inductive hypothesis $(*)_1$ is obvious; therefore, it suffices to perform the inductive step, deducing $(*)_{n+1}$ from $(*)_n$.

Let $\widetilde{x}_{n+1} \in \mathbb{Z}/p^{n+1}\mathbb{Z}$ be some element lifting x_n. Define the polynomial $g(t)$ as $f_{n+1}(\widetilde{x}_{n+1} + p^n \cdot t)$. It suffices to show that $g(t)$ has a exactly p^n roots y, since then $\widetilde{x}_{n+1} + p^n y = x_{n+1}$ obviously is well-defined and is the unique solution to f_{n+1} lifting x_n.

Note that for each $r \geqslant 0$, $(\widetilde{x}_{n+1} + p^n \cdot t)^r = \widetilde{x}_{n+1}^r + rp^n\widetilde{x}_{n+1} \cdot t$ since $(p^n)^2 = 0 \in \mathbb{Z}/p^{n+1}\mathbb{Z}$. We deduce by linearity that $g(t) = f_{n+1}(\widetilde{x}_{n+1}) + p^n f'_{n+1}(\widetilde{x}_{n+1}) \cdot t$.

We see that $f'_{n+1}(\widetilde{x}_{n+1})$ is a unit in $\mathbb{Z}/p^{n+1}\mathbb{Z}$ by assumption on f and by Lemma 3.8.3 (2). Therefore, $g(t)$ is a linear polynomial over $\mathbb{Z}/p^{n+1}\mathbb{Z}$ with leading and constant terms lying in $p^n\mathbb{Z}/p^{n+1}\mathbb{Z}$, giving the claim.

□

Corollary 3.10.4. *Let p be an odd prime. Then $x \in \mathbb{Z}_p^\times$ is a square if and only if $x \bmod p\mathbb{Z}_p$ is a quadratic residue.*

More generally, $x \in \mathbb{Q}_p$ is a square if and only if $v_p(x)$ is even and $p^{-v_p(x)}x$ is a quadratic residue modulo $p\mathbb{Z}_p$.

Proof. For $x \in \mathbb{Z}_p$, the polynomial $f(t) = t^2 - x$ has derivative $2t$, which is non-vanishing derivative at any element of $\mathbb{F}_p^\times$ as long as p is prime. We deduce the first part then from Hensel's lemma.

For the more general case, we appeal to the isomorphism (3.1).

□

Corollary 3.10.5. *For p odd, the map:*

[7] This condition is equivalent to saying that $\overline{x}$ is a simple root of $\overline{f}$.

$$\mathbb{Q}_p^\times \to \{1,-1\} \times \{1,-1\}$$
$$x \mapsto \Big((-1)^{v_p(x)}, \Big(\frac{\overline{p^{-v_p(x)}x}}{p}\Big)\Big)$$

induces a group isomorphism:

$$\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \xrightarrow{\simeq} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

Here the second coordinate of this map indicates the Legendre symbol of $\overline{p^{-v_p(x)}x} \in \mathbb{F}_p^\times$, this element being the reduction modulo p of $p^{-v_p(x)}x \in \mathbb{Z}_p^\times$.

Exercise 3.5. (1) Show that $\mathbb{Q}_p$ contains all $(p-1)$st roots of unity.

(2) Show that there is a unique multiplicative (but not additive) map $\mathbb{F}_p \to \mathbb{Z}_p$ such that the induced map $\mathbb{F}_p \to \mathbb{Z}_p \to \mathbb{F}_p$ is the identity.

Remark 3.10.6. The map $\mathbb{F}_p \to \mathbb{Z}_p$ is called "Teichmuller lifting."

3.11. ***p*-adic exponential and logarithm.** We now introduce p-adic analogues of the usual exponential and logarithm functions in order to better understand the structure of the unit groups $\mathbb{Q}_p^\times$. Note that this is in analogy with the case of the real and complex numbers, where these functions are crucial for giving a complete description of the relevant multiplicative groups.

We will define exp and log in Notation 3.11.3 by the usual power series definitions after discussing the convergence of these series.

Lemma 3.11.1. *(1) Suppose that $x \in \mathbb{Q}_p$ with $v_p(x) > \frac{1}{p-1}$. Then the series:*

$$\sum_{i=0}^{\infty} \frac{1}{i!} \cdot x^i \tag{3.2}$$

converges in $\mathbb{Q}_p$.

(2) For $|x|_p < 1$, the series:

$$\sum_{i=1}^{\infty} -\frac{1}{i} x^i \tag{3.3}$$

converges.

Remark 3.11.2. We note that in (1), $v_p(x) > \frac{1}{p-1}$ is equivalent to $x \in p\mathbb{Z}_p$ for p odd, and is equivalent to $x \in p^2\mathbb{Z}_p$ for $p = 2$. We formulate the result in terms of the apparently too precise estimate of $\frac{1}{p-1}$ since it is what emerges from the proof, and is relevant for generalizations of this lemma to finite extensions of $\mathbb{Q}_p$. Therefore, we will continue to use this estimate below to indicate the radius of convergence of the p-adic exponential function, silly though it might appear given the above.

Proof of Lemma 3.11.1. For (1): it suffices to see that $v_p(\frac{1}{i!}x^i) \to \infty$. A standard sieving argument gives:

$$v_p(i!) = \sum_{i=1}^{\infty} \lfloor \frac{i}{p^i} \rfloor$$

(note that this sum is actually finite). Therefore, we estimate:

$$v_p(i!) < \sum_{i=1}^{\infty} \frac{i}{p^i} = i \cdot \frac{1}{p} \cdot \frac{1}{1-\frac{1}{p}} = i \cdot \frac{1}{p-1}.$$

On the other hand, $v_p(x^i) = i \cdot v_p(x)$, so we find:

$$v_p(\frac{x^i}{i!}) = v_p(x^i) - v_p(i!) > i \cdot (v_p(x) - \frac{1}{p-1})$$

which goes to ∞ with i if $v_p(x) > \frac{1}{p-1}$.

For (2): we expand $v_p(\frac{1}{i}x^i)$ as $i \cdot v_p(x) - v_p(i)$. As $p^{v_p(i)} \leqslant i$, we can bound $v_p(i)$ by $\frac{\log(i)}{\log(p)}$, so that:

$$v_p(\frac{1}{i}x^i) \geqslant iv_p(x) - \frac{\log(i)}{\log(p)}.$$

As long as $|x|_p < 1$ (equivalently: $v_p(x) > 0$), the linear leading term grows faster than the logarithmic second term, and we obtain the result.

□

Notation 3.11.3. For $v_p(x) > \frac{1}{p-1}$, we let $\exp(x)$ denote the result of evaluating the series (3.2). For $|x|_p < 1$, let $\log(1-x)$ denote the result of evaluating the series (3.3) at x.

Note that $\exp(x)$ lies in the subset $1 + p\mathbb{Z}_p$ of $\mathbb{Q}_p$ (for $v_p(x) > \frac{1}{p-1}$), since the first term in the series (3.2) is 1 and the higher order terms are divisible by p (we see that for $p = 2$, we even have $\exp(x) \in 1 + 4\mathbb{Z}_2$).

Note that $x \in 1 + p\mathbb{Z}_p$ is equivalent to requiring $|1-x|_p < 1$, so these are exactly the elements for which we can make sense of the p-adic logarithm $\log(x)$.

Proposition 3.11.4. *(1) For x and y of valuation greater than $\frac{1}{p-1}$, we have:*

$$\exp(x+y) = \exp(x) \cdot \exp(y).$$

(2) For $x, y \in 1 + p\mathbb{Z}_p$, we have:

$$\log(xy) = \log(x) + \log(y).$$

(3) For $v_p(x) > \frac{1}{p-1}$, we have $\log(\exp(x)) = x$, and for $v_p(1-x) > \frac{1}{p-1}$ we have $v_p(\log(x)) > \frac{1}{p-1}$ and $\exp(\log(x)) = x$.

Proof. These all follow from standard series manipulations. For (1):

$$\exp(x+y) := \sum_{i=0}^{\infty} \frac{1}{i!}(x+y)^i = \sum_{i=0}^{\infty}\sum_{j=0}^{i} \frac{1}{i!}\frac{i!}{j!(i-j)!}x^j y^{i-j} = \sum_{i=0}^{\infty}\sum_{j=0}^{i} \frac{1}{j!}\frac{1}{(i-j)!}x^j y^{i-j} =$$

$$\sum_{i=0}^{\infty} \sum_{\substack{j,k \geqslant 0 \\ j+k=i}} \frac{1}{j!}\frac{1}{k!}x^j y^k = \Big(\sum_{j=0}^{\infty} \frac{1}{j!}x^j\Big) \cdot \Big(\sum_{k=0}^{\infty} \frac{1}{k!}y^k\Big) = \exp(x) \cdot \exp(y).$$

For (2), we first claim that for a general field k of characteristic 0, in the ring $k[[t,s]]$ of formal power series in two variables we have an identity:

$$\sum_{i=1}^{\infty} -\frac{1}{i}(s+t-st)^i = \sum_{i=1}^{\infty} -\frac{1}{i}(s^i + t^i). \tag{3.4}$$

Note that in the left hand side the coefficient of $s^n t^m$ obviously has only finitely many contributions, so this is a well-defined formal power series. To verify this identity, we first apply the formal partial derivative $\frac{\partial}{\partial t}$ to the left hand side and compute:

$$\frac{\partial}{\partial t}\Big(\sum_{i=1}^{\infty} -\frac{1}{i}(s+t-st)^i\Big) = \sum_{i=1}^{\infty} -(1-s)(s+t-st)^{i-1} = -\frac{1-s}{(1-s-t+st)} =$$

$$-\frac{1-s}{(1-s)(1-t)} = -\frac{1}{1-t} = \frac{\partial}{\partial t}\Big(-\sum_{i=1}^{\infty}\frac{1}{i}\big(t^i+s^i\big)\Big).$$

Therefore, it suffices to verify the equality (3.4) after setting $t=0$, since the formal derivative computation implies that the coefficients of $s^n t^m$ of the left and right hand sides of (3.4) are equal whenever $n>0$. But we obviously have the desired equality after setting $t=0$.

We then immediately deduce (2) from the equality of the formal power series (3.4), since by convergence we can substitute $1-x$ for s and $1-y$ for t into (3.4).

Finally, for (3), we similarly use formal power series in a variable t. First, note that:

$$\sum_{j=0}^{\infty}\frac{1}{j!}\Big(\sum_{i=1}^{\infty} -\frac{1}{i}t^i\Big)^j$$

is defined in $k[[t]]$ (for k as above), since one easily sees that for each coefficient t^i only finitely many terms contribute. Let $g(t)$ denote the resulting power series; we wish to show that $g(t) = 1-t$. One computes:

$$\frac{d}{dt}g(t) = \frac{d}{dt}\Big(\sum_{j=0}^{\infty}\frac{1}{j!}\Big(\sum_{i=1}^{\infty} -\frac{1}{i}t^i\Big)^j\Big) = \sum_{j=1}^{\infty}\frac{1}{(j-1)!}\Big(\sum_{i=1}^{\infty} -\frac{1}{i}t^i\Big)^{j-1}\cdot\frac{-1}{1-t} = \frac{-1}{1-t}\cdot g(t).$$

Multiplying by $1-t$, we see that this equation defines a recursion on the coefficients of $g(t)$, and one finds that it characterizes $g(t)$ up to scaling. Moreover, $1-t$ is a solution to this equation. Setting $t=0$, we then deduce that $g(t)=1-t$. One argues similarly that:

$$\sum_{i=0}^{\infty}\frac{-1}{i}\big(-\sum_{j=1}^{\infty}\frac{1}{j!}t^j)^i\big) = t.$$

Next, we claim that for $y \in \mathbb{Q}_p$ with $v_p(y) > \frac{1}{p-1}$, we have:

$$i\cdot v_p(y) - v_p(i) > \frac{1}{p-1}. \tag{3.5}$$

Supposing this inequality for a moment, then applying this to $y = 1-x$ we see that $\log(x) = \sum_{i=1}^{\infty} -\frac{1}{i}(1-x)^i$ is a sum of terms of valuation at least $\frac{1}{p-1}$, giving that $\log(x)$ has the same property, as desired. From here, the rest of (3) follows from convergence and the properties of formal power series noted above.

It remains to verify the claim (3.5). First, observe that it is true for $i<p$: in this case, we have $v_p(i)=0$, so the result is clear.

To treat $i \geqslant p$, we first bound the left hand side as:

$$i\cdot v_p(y) - v_p(i) > \frac{i}{p-1} - \frac{\log(i)}{\log(p)}. \tag{3.6}$$

We can now regard i as a continuous variable in $\mathbb{R}$. The derivative of this function is then:

$$\frac{1}{p-1} - \frac{1}{i\log(p)}.$$

This derivative is zero exactly for $i = \frac{p-1}{\log(p)}$, which is obviously the minimum for the function.

Note that $\frac{p-1}{\log(p)} < p$ for all p. Indeed, for $p \neq 2$ (i.e., $p > e :=$ the base of the natural logarithm) this is clear since then the left hand side is less than $p-1$. For $p = 2$, this says that $\frac{1}{2} < \log(2)$, i.e., $\sqrt{e} < 2$, which is clear. Moreover, substituting the value $i = p$ into the right hand side of (3.6) above gives:

$$\frac{p}{p-1} - 1 = \frac{1}{p-1}.$$

Since this p is past the minimum of the function, this gives the desired result for all $i \geqslant p$.

□

Corollary 3.11.5. *The exponential and logarithm maps define mutually inverse equivalences of abelian groups between $\{x \in \mathbb{Z}_p \mid v_p(x) > \frac{1}{p-1}\}$ and $\{x \in \mathbb{Z}_p \mid v_p(1-x) > \frac{1}{p-1}\}$, the former being considered as an abelian group under addition, and the latter as an abelian group under multiplication.*

3.12. **Squares in $\mathbb{Q}_2$.** We now discuss the consequences of Corollary 3.11.5 for the structure of squares in $\mathbb{Q}_p$, generalizing Corollary 3.10.4.

Proposition 3.12.1. *$x \in \mathbb{Z}_2^\times$ is a square if and only if $x = 1 \bmod 8$.*

Proof. To see necessity, note that invertibility of x implies that $x \in \{1,3,5,7\} \bmod 8$, and the only square in $\mathbb{Z}/8\mathbb{Z}$ among these is 1 (since $(2k+1)^2 = 4(k^2+k)+1$ and k^2+k is always even).

For the converse, first note that for $p = 2$, the inequality $v_p(x) > \frac{1}{p-1}$ translates to $v_p(x) \geqslant 2$. Therefore, by Corollary 3.11.5 we have an isomorphism $1 + 4\mathbb{Z}_2 \overset{\log}{\simeq} 4\mathbb{Z}_2$, the former group being considered with multiplication and the latter with addition. Since the image of multiplication by 2 in $4\mathbb{Z}_2$ is $8\mathbb{Z}_2$, it suffices to show that this equivalence identifies $8\mathbb{Z}_2$ with $1 + 8\mathbb{Z}_2$.

Recall from the proof of Lemma 3.11.1 that we proved in general that:

$$v_p(\frac{x^i}{i!}) > i \cdot (v_p(x) - \frac{1}{p-1}).$$

For $p = 2$, we obtain $v_2(\frac{x^i}{i!}) > i \cdot (v_2(x) - 1)$. For $v_2(x) \geqslant 2$, this implies that for $i \geqslant 2$ we have:

$$v_2(\frac{x^i}{i!}) > 2(v_2(x) - 1) = v_2(x) + (v_2(x) - 2) \geqslant v_2(x)$$

so that:

$$|1 - \exp(x)|_p = |x|_p$$

for all x with $v_p(x) \geqslant 2$, verifying the claim.

□

Corollary 3.12.2. *We have an isomorphism:*

$$\mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2 \xrightarrow{\simeq} \{1,-1\} \times (\mathbb{Z}/8\mathbb{Z})^\times$$
$$x \mapsto \Big((-1)^{v_2(x)}, \overline{2^{-v_2(x)}x}\Big)$$

where $\overline{2^{-v_2(x)}x} \in (\mathbb{Z}/8\mathbb{Z})^\times$ *denotes the reduction of* $2^{-v_2(x)} \cdot x \in \mathbb{Z}_2^\times$*.*

Remark 3.12.3. In particular, we see that $|\mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2| = 8$.

Exercise 3.6. Show that $\mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2$ is generated by the images of 2, 3 and 5.

3.13. **Quadratic forms over finite fields.** Before proceeding to discuss quadratic forms over $\mathbb{Q}_p$, we need to digress to treat quadratic forms over finite fields.

The main result on the former subject is the following.

Proposition 3.13.1. *Let p be an odd prime.*

(1) Any form of rank $\geqslant 2$ represents every value in $\mathbb{F}_p$.

(2) Two non-degenerate forms over $\mathbb{F}_p$ are equivalent if and only if their discriminants are equal.

(3) Let r be a quadratic non-residue modulo p. Then every non-degenerate rank n form q over $\mathbb{F}_p$ is equivalent to exactly one of the forms:

$$x_1^2 + \ldots + x_n^2 \text{ or}$$
$$x_1^2 + \ldots + rx_n^2.$$

Proof. For (1), by diagonalization it suffices to show any form $ax^2 + by^2$ with $a, b \neq 0$ represents every value $\lambda \in \mathbb{F}_p$.

Note that x^2 represents exactly $\frac{p+1}{2}$ values in $\mathbb{F}_p$: the $\frac{p-1}{2}$ quadratic residues and 0. Therefore, the same counts hold for ax^2 and by^2, since $a, b \neq 0$. Therefore, the intersection:

$$\{a\alpha^2 \mid \alpha \in \mathbb{F}_p\} \cap \{\lambda - b\beta^2 \mid \beta \in \mathbb{F}_p\}$$

is non-empty, since each of these sets has order greater than $\frac{p}{2}$. Choosing $\alpha, \beta \in \mathbb{F}_p$ defining an element of the intersection, this gives $a\alpha^2 = \lambda - b\beta^2$ as desired.

For (3), we proceed by induction on n. The result is clear for $n = 1$. Otherwise, by (1), q represents $1 \in \mathbb{F}_p$. Choosing a vector v in the underlying quadratic space V with $q(v) = 1$ and decomposing according to orthogonal complements, we obtain the result by induction.

Finally, (2) obviously follows from (2).

□

3.14. **Hilbert symbol.** Let $k = \mathbb{Q}_p$ for some p or $k = \mathbb{R}$.

For $a, b \in k^\times$, we define the *Hilbert symbol* of a and b as 1 if the binary quadratic form $ax^2 + by^2$ represents 1, and -1 otherwise. We denote the Hilbert symbol of a and b by:

$$(a, b)_p \in \{1, -1\} \text{ if } k = \mathbb{Q}_p, \text{ and}$$
$$(a, b)_\infty \in \{1, -1\} \text{ if } k = \mathbb{R}.$$

Remark 3.14.1. Of course, this definition makes sense over any field, but it does not have good properties: Theorem 3.14.5 will use specific facts about the choice $k = \mathbb{Q}_p$ and $k = \mathbb{R}$. This failure represents the fact that our definition is somewhat ad hoc.

Remark 3.14.2. The role the Hilbert symbol plays for us is that it is clearly an invariant of the quadratic form $q(x, y) = ax^2 + by^2$, i.e., it only depends on the underlying quadratic space, not on a and b. Eventually, we will denote it by $\varepsilon(q)$, call it the *Hasse-Minkowski invariant* of q, and we will generalize it to higher rank quadratic forms.

Here we record the most obvious properties of the Hilbert symbol.

Proposition 3.14.3. *The following identities are satisfied by the Hilbert symbol for p a prime or $p = \infty$.*

(1) $(a, b)_p = (b, a)_p$.
(2) $(a, 1 - a)_p = 1$.
(3) $(1, a)_p = 1$.
(4) $(a, -a)_p = 1$.
(5) $(a\lambda^2, b)_p = (a, b)_p$ *for all non-zero* λ.
(6) $(a, b)_p = (a, -ab)_p$.

Proof. In words, (1) says that $ax^2 + by^2$ represents 1 if and only if $bx^2 + ay^2$ does, which is clear.

Then (2) says that $ax^2 + (1 - a)y^2$ represents 1, which it does with $(x, y) = (1, 1)$, while (3) says that $x^2 + ay^2$ represents 1, which it does with $(1, 0)$.

Next, (4) says that $ax^2 - ay^2$ represents 1, and this is clear because the form is isotropic and non-degenerate, therefore equivalent to the hyberbolic plane, and therefore represents all values.

(5) follows because $(a, b)_p$ is an invariant of the equivalence class of the quadratic form $ax^2 + by^2$.

Finally, for (6), we note that $(a, b)_p = 1$ if and only if $q_1(x, y, z) := ax^2 + by^2 - z^2$ is isotropic by Corollary 2.10.3. Similarly, $(a, -ab)_p = 1$ if and only if:

$$q_2(x, y, z) := \frac{-1}{a}(ax^2 - aby^2 - z^2) = -x^2 + by^2 + \frac{1}{a}z^2$$

is isotropic. Therefore, it suffices to see that q_1 and q_2 are equivalent quadratic forms. But this is clear: swap the variables x and z and change $\frac{1}{a}$ to a by multiplication by a^2.

□

Remark 3.14.4. By (5), we the Hilbert symbol defines a symmetric pairing:

$$\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \to \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \to \{1, -1\} \tag{3.7}$$

and similarly for $\mathbb{R}$.

We will prove the following less trivial results below. It is here that we will really need that our field is $\mathbb{Q}_p$ or $\mathbb{R}$.

Theorem 3.14.5. *Let p be a prime or ∞.*

(1) $(a, bc)_p = (a, b)_p \cdot (a, c)_p$.
(2) If $(a, b)_p = 1$ *for all invertible b, then a is a square.*

Remark 3.14.6. We may view $k^\times/(k^\times)^2$ as an $\mathbb{F}_2$-vector space, where we change from multiplicative notation in $k^\times/(k^\times)^2$ to additive notation for vector addition.

Then Theorem 3.14.5 exactly says that for $k = \mathbb{Q}_p$ or $k = \mathbb{R}$, $(a, b) \mapsto (a, b)_p \in \{1, -1\} \simeq \mathbb{F}_2$ defines a non-degenerate symmetric bilinear form on the vector space $k^\times/(k^\times)^2$ (which we note is finite-dimensional over $\mathbb{F}_2$, of dimension 1 for $k = \mathbb{R}$, 2 for $k = \mathbb{Q}_p$ and $p \neq 2$, and 3 for $\mathbb{Q}_2$).

3.15. **Consequences for quadratic forms.**

For this subsection, we will assume Theorem 3.14.5, leaving its proof (and a detailed description of how to compute with the Hilbert symbol) to §3.16.

Let $k = \mathbb{Q}_p$ or $k = \mathbb{R}$; in the latter case, the reader should understand p below as a stand-in for ∞. As in Remark 3.14.2 binary[8] quadratic form q, we let $\varepsilon(q) \in \{1, -1\}$ be 1 if q represents 1 and -1 otherwise. So if $q = ax^2 + by^2$, then $\varepsilon(q) = (a, b)_p$.

[8] We remind that this means that its underlying quadratic space has dimension 2.

Proposition 3.15.1. *Let q be a non-degenerate binary quadratic form with discriminant* $\operatorname{disc}(q)$*. Then q represents $\lambda \in k^\times$ if and only if:*

$$(-\operatorname{disc}(q), \lambda)_p = \varepsilon(q).$$

Proof. That q represents λ is equivalent to saying that $\frac{1}{\lambda}q$ represents 1. Diagonalizing q so that $q(x, y) = ax^2 + by^2$, we see that this is equivalent to having $(\frac{a}{\lambda}, \frac{b}{\lambda})_p = 1$. For convenience, we replace this expression by $(\lambda a, \lambda b)_p$, which is justified by the equality $\lambda = \frac{1}{\lambda} \in k^\times/(k^\times)^2$.

Applying the bilinearity Theorem 3.14.5 (1) of the Hilbert symbol repeatedly, we compute:

$$(\lambda a, \lambda b)_p = (\lambda, \lambda)_p(\lambda, b)_p(a, \lambda)_p(a, b)_p = (-1, \lambda)_p^2(\lambda, \lambda)_p(b, \lambda)_p(a, \lambda)_p(a, b)_p =$$
$$\Big((-1, \lambda)_p(\lambda, \lambda)_p\Big)\Big((-1, \lambda)_p(b, \lambda)_p(a, \lambda)_p\Big)(a, b)_p = (-\lambda, \lambda)_p(-ab, \lambda)_p(a, b)_p.$$

Observing that $(-\lambda, \lambda)_p = 1$ by Proposition 3.14.3 (4), we see that:

$$(\lambda a, \lambda b)_p = (-\operatorname{disc}(q), \lambda)_p \cdot \varepsilon(q).$$

This exactly means that $\varepsilon(q) = (-\operatorname{disc}(q), \lambda)_p$ if and only if q represents λ, as desired.

□

Remark 3.15.2. Here are some sanity tests to make sure that Proposition 3.15.1 accords with reality.

First of all, for $\lambda = 1$, this formula says that q represents 1 if and only if $(-\operatorname{disc}(q), 1)_p = \varepsilon(q)$, but since 1 is a square, we see that $(-\operatorname{disc}(q), 1)_p = 1$, so this is equivalent to saying $\varepsilon(q) = 1$, as desired.

Next, if q is the hyperbolic plane, we have $\operatorname{disc}(q) = -1$ and $\varepsilon(q) = 1$, so Proposition 3.15.1 says that q represents λ if and only if $(1, \lambda)_p = 1$. As above $(1, \lambda)_p = 1$ for all λ, so this says that the hyperbolic plane represents every value, which is obviously true.

Corollary 3.15.3. *Non-degenerate binary quadratic forms q_1 and q_2 are equivalent if and only if* $\operatorname{disc}(q_1) = \operatorname{disc}(q_2)$ *and* $\varepsilon(q_1) = \varepsilon(q_2)$*.*

Proof. Diagonalize q_1 as $ax^2 + by^2$. By Proposition 2.12.1, q_1 and q_2 represent the same values, and therefore q_2 represents a as well. By Proposition 2.9.7, q_2 can be diagonalized as $ax^2 + b'y^2$. We now see that:

$$ab' = \operatorname{disc}(q_2) = \operatorname{disc}(q_1) = ab \bmod (k')^2$$

so that changing b' by a square we can take $b' = b$ as desired.[9]

□

Corollary 3.15.4. *If q is a non-degenerate binary quadratic form other than the hyperbolic plane, then q represents exactly half of the elements of the finite set $k^\times/(k^\times)^2$.*

Proof. By Exercise 2.27, $-\operatorname{disc}(q) \in (k^\times)^2$ if and only q is the hyperbolic plane. Therefore, if q is not the hyperbolic plane, then $(-\operatorname{disc}(q), -) : (k^\times/(k^\times)^2) \to \{1, -1\}$ is a non-trivial homomorphism by Theorem 3.14.5 (2), and λ is represented by q if and only if it lies in the appropriate coset of this homomorphism.

□

[9] This argument implies over a general field of characteristic $\neq 2$ that two non-degenerate binary quadratic forms of the same discriminant and representing the same values are equivalent. Note that this is not true in rank $\geqslant 3$.

Exercise 3.7. For $k = \mathbb{R}$, verify each of the above statements by hand (note that the Hilbert symbol for $\mathbb{R}$ is explicitly described in Proposition 3.16.1 below).

Corollary 3.15.5. *Any non-degenerate ternary quadratic form q represents every $\lambda \in k^\times/(k^\times)^2$ except possibly $\lambda = -\operatorname{disc}(q)$.*

Proof. We diagonalize q as $q(x, y, z) = ax^2 + by^2 + cz^2$ for $a, b, c \in k^\times$.

Fix $\lambda \neq -\operatorname{disc}(q)$. By Corollary 2.10.3, it suffices to show that $q_\lambda := q_1 - q_2 = ax^2 + by^2 + cz^2 - \lambda w^2$ is isotropic.

Observe that the forms $q_1(x, y) := ax^2 + by^2$ and $q_2(z, w) = -cz^2 + \lambda w^2$ have different discriminants by assumption on λ. Note that we can assume that $\operatorname{disc}(q_1)$ and $\operatorname{disc}(q_2)$ are each not -1, since otherwise one of these forms is the hyperbolic plane, which obviously would imply that q_λ is isotropic.

Therefore, $-\operatorname{disc}(q_1)$ and $-\operatorname{disc}(q_2)$ are linearly independent when regarded as elements of the $\mathbb{F}_2$-vector space $k^\times/(k^\times)^2$, since they are distinct non-zero vectors in a $\mathbb{F}_2$-vector space. Therefore, by Theorem 3.14.5 (2), there exists $\eta \in k^\times$ with $(-\operatorname{disc}(q_1), \eta)_p = \varepsilon(q_1)$ and $(-\operatorname{disc}(q_2), \eta)_p = \varepsilon(q_2)$. Indeed, in the perspective of Remark 3.14.6, we are given linearly independent vectors and a non-degenerate bilinear form, so we can find a vector that pairs to these vectors in any specified way.

But then $q_1(\eta) - q_2(\eta) = 0$, so the form $ax^2 + by^2 + cz^2 - \lambda w^2$ is isotropic, giving the claim.

□

In the special case that $k = \mathbb{Q}_p$ (i.e., $k \neq \mathbb{R}$), we moreover obtain the following results.

Corollary 3.15.6. *Any quadratic form q over $k = \mathbb{Q}_p$ of rank $\geqslant 5$ is isotropic.*

Proof. We can write $q = q_1 + q_2 + q_3$ where q_1 has rank 3, q_2 has rank 2 and q_3 is otherwise arbitrary. By Corollary 3.15.5, q_1 represents every value in $k^\times/(k^\times)^2$ except possibly 1 (namely, $-\operatorname{disc}(q_1)$).

Therefore, since $k \neq \mathbb{R}$ so that $|k^\times/(k^\times)^2| > 2$, by Corollary 3.15.4 there exists η in $k^\times$ that is represented by both q_1 and $-q_2$, meaning that $q_1 + q_2$–and therefore q itself–is isotropic.

□

Corollary 3.15.7. *For $k = \mathbb{Q}_p$, any form of rank $\geqslant 4$ represents every $\lambda \in k$.*

Proof. Immediate from Corollary 3.15.6 and Corollary 2.10.3.

□

3.16. We now will give explicit formulae for the Hilbert symbol. We will deduce Theorem 3.14.5 from these explicit formulae. We separate into cases of increasing difficulty: $k = \mathbb{R}$, $k = \mathbb{Q}_p$ with $p \neq 2$, and $\mathbb{Q}_2$.

Proposition 3.16.1. *For $a, b \in \mathbb{R}^\times$, $(a, b)_\infty = 1$ unless a and b are both negative.*

Proof. This says that $ax^2 + by^2$ represents 1 unless a and b are both negative, which is clear.

□

In order to compute the Hilbert symbol for $\mathbb{Q}_p$, it is convenient to pass through the *tame symbol*. This is a rule that takes $a, b \in \mathbb{Q}_p^\times$ and produces $\mathrm{Tame}_p(a, b) \in \mathbb{F}_p^\times$.

To define the tame symbol, note that for $a, b \in \mathbb{Q}_p^\times$, we have $\frac{a^{v_p(b)}}{b^{v_p(a)}} \in \mathbb{Z}_p^\times$, since its valuation is 0. Therefore, we may reduce it to $\overline{(\frac{a^{v_p(b)}}{b^{v_p(a)}})} \in \mathbb{F}_p^\times$. Then we define:

$$\mathrm{Tame}_p(a, b) := (-1)^{v_p(a) \cdot v_p(b)} \cdot \overline{(\frac{a^{v_p(b)}}{b^{v_p(a)}})} \in \mathbb{F}_p^\times.$$

Observe that the tame symbol is bilinear, i.e., we have:

$$\mathrm{Tame}_p(aa', b) = \mathrm{Tame}_p(a, b) \cdot \mathrm{Tame}_p(a', b) \in \mathbb{F}_p^\times \text{ and}$$
$$\mathrm{Tame}_p(a, bb') = \mathrm{Tame}_p(a, b) \cdot \mathrm{Tame}_p(a, b') \in \mathbb{F}_p^\times.$$

Proposition 3.16.2. *If p is an odd prime, then:*

$$(a,b)_p = \left(\frac{\mathrm{Tame}_p(a,b)}{p}\right). \tag{3.8}$$

Proof. First, we analyze how the tame symbol of a and b changes when a is multiplied by p^2. By definition, this is computed as:

$$\left(\overline{(\frac{p^{v_p(b)}}{b})}\right)^2 \cdot \mathrm{Tame}_p(a,b).$$

Therefore, when we apply the Legendre symbol, this does not change the resulting value, i.e., the right hand side of (3.8) does not change under this operation.

By Proposition 3.14.3 (5), the right hand side does not change under this operation either.

Now observe that the right hand side of (3.8) is symmetric under switching a and b: indeed, switching a and b has the effect of inverting the tame symbol in $\mathbb{F}_p^\times$, and the Legendre symbol is immune to this change.

Therefore, we reduce to studying three cases: 1) $v_p(a) = v_p(b) = 0$, 2) $v_p(a) = 0$, $v_p(b) = 1$, and 3) $v_p(a) = v_p(b) = 1$.

Case 1. $v_p(a) = v_p(b) = 0$.

In this case, we see that the tame symbol of a and b is 1. Therefore, we need to show that their Hilbert symbol is as well, i.e., that $q(x, y) = ax^2 + by^2$ represents 1.

By Proposition 3.13.1, we can solve $\overline{a}x^2 + \overline{b}y^2 = 1$ over $\mathbb{F}_p$, say with $(x, y) = (\overline{\alpha}, \overline{\beta})$. Choose a lift $\alpha \in \mathbb{Z}_p$ of $\overline{\alpha}$. We see that the polynomial $bt^2 + (a\alpha^2 - 1)$ satisfies the hypotheses of Hensel's lemma and therefore there exists a unique lift β of $\overline{\beta}$ that is a solution to the equation $b\beta^2 + a\alpha^2 - 1 = 0$, as desired.

Case 2. $v_p(a) = 0$, $v_p(b) = 1$.

We find $\mathrm{Tame}_p(a, b) = \overline{a}$, so the right hand side of (3.8) is the Legendre symbol of $\overline{a}$.

To compare with the left hand side: note that if $(x, y) \in \mathbb{Q}_p^2$ then $v_p(ax^2+by^2) = \min\{v_p(ax^2), v_p(bx^2)\}$. Indeed, since $v_p(ax^2) \in 2\mathbb{Z}$ and $v_p(bx^2) \notin 2\mathbb{Z}$, so these numbers are not equal. Therefore, if $ax^2 + by^2 = 1$, then $0 = \min\{v_p(ax^2), v_p(by^2)\} = \min\{2v_p(x), 1 + 2v_p(y)\}$, meaning that x and y are p-adic integers with x a p-adic unit.

Therefore, we can reduce an equation $ax^2 + by^2 = 1$ modulo p to see that a must be a quadratic residue. The converse direction follows from Corollary 3.10.4.

Case 3. $v_p(a) = v_p(b) = 1$.

By Proposition 3.14.3 (6), we have $(a, b)_p = (a, -ab)_p$. Noting that $v_p(-ab) \in 2\mathbb{Z}$, Case (2) applies and we see that:

$$(a, -ab)_p = \left(\frac{\mathrm{Tame}_p(a, -ab)}{p}\right).$$

We then compute that:

$$\mathrm{Tame}_p(a,-ab) = \frac{a^2}{-ab} = -\frac{a}{b} = \mathrm{Tame}_p(a,b)$$

giving the desired result.

□

It remains to treat the case of $\mathbb{Q}_2$.

For $a \in \mathbb{Q}_2^\times$, let $\varpi(a) := \frac{(2^{-v_2(a)}a)^2-1}{8} \in \mathbb{Z}/2\mathbb{Z}$. That is, for $a \in \mathbb{Z}_2^\times$ we have:

$$\begin{cases} \varpi(a) = 0 & \text{if } a = 1, -1 \bmod 8\mathbb{Z}_2 \\ \varpi(a) = 1 & \text{if } a = 3, 5 \bmod 8\mathbb{Z}_2 \end{cases}$$

and for general $a \in \mathbb{Q}_2^\times$ we have $\varpi(a) = \varpi(2^{-v_2(a)}a)$.

We also set $\vartheta(a) = \frac{2^{-v_2(a)}a-1}{2} \in \mathbb{Z}/2\mathbb{Z}$. So again, $\vartheta(a) = \vartheta(2^{-v_2(a)}a)$, and for $a \in \mathbb{Z}_2^\times$ we have $\vartheta(a) = 0$ if $a = 1 \bmod 4$ and $\vartheta(a) = 1$ if $a = 3 \bmod 4$.

Proposition 3.16.3. *For all $a, b \in \mathbb{Q}_2^\times$, we have:*

$$(a,b)_2 = (-1)^{\varpi(a)v_2(b)+v_2(a)\varpi(b)+\vartheta(a)\vartheta(b)}. \tag{3.9}$$

Proof. As in the proof of Proposition 3.16.2, we reduce to the cases 1) $v_2(a) = v_2(b) = 0$, 2) $v_2(a) = 0$, $v_2(b) = 1$, and 3) $v_2(a) = v_2(b) = 1$.

Case 1. $v_2(a) = v_2(b) = 0$.

In this case, we see that the right hand side of (3.9) is 1 unless $a = b = 3 \bmod 4\mathbb{Z}_2$.

We treat separately the cases i) where a (or b) is 1 modulo 8, ii) a (or b) is 5 modulo 8, and iii) a and b are both 3 modulo 4.

Subcase 1. $a = 1 \bmod 8\mathbb{Z}_2$.

Then by Corollary 3.12.2, a is a square in $\mathbb{Q}_2$, so $ax^2 + by^2$ represents 1.

Subcase 2. $a = 5 \bmod 8\mathbb{Z}_2$.

In this case, $a + 4b = 1 \bmod 8$ and therefore is a square in $\mathbb{Q}_2$, meaning that $ax^2 + by^2$ represents a square and therefore represents 1.

Subcase 3. $a = b = 3 \bmod 4\mathbb{Z}_2$.

Suppose $ax^2 + by^2 = 1$, $x, y \in \mathbb{Q}_2$. Let $n := \max\{0, -v_2(x), -v_2(y)\}$.

Then we have:

$$a(2^n x)^2 + b(2^n y)^2 = 4^n$$

as an equation in $\mathbb{Z}_2$.

If $n = 0$ so that $x, y \in \mathbb{Z}_2$, this equation gives $3x^2 + 3y^2 = 1 \bmod 4$, which isn't possible since $\lambda^2 \in \{0, 1\}$ for $\lambda \in \mathbb{Z}/4\mathbb{Z}$.

If $n > 0$, then up to change of variables we can assume $n = -v_2(x)$, in which case $2^n x \in \mathbb{Z}_2^\times$, so $a(2^n x)^2 + b(2^n y)^2 = 3 + 3(2^n y)^2 = 0 \bmod 4$. Now observe that $(2^n y)^2$ can only reduce to 0 or 1 modulo $4\mathbb{Z}_2$, and neither of these provides a solution to the given equation.

Case 2. $v_2(a) = 0$, $v_2(b) = 1$.

Let $b = 2c$ with $c \in \mathbb{Z}_2^\times$. Then we need to show that $(a,b)_2 = (-1)^{\varpi(a)+\vartheta(a)\vartheta(c)}$.

We note that, as in Case 2 in the proof of Proposition 3.16.2, any solution to $ax^2 + by^2 = 1$ must have $x, y \in \mathbb{Z}_2$.

Therefore, if $\overline{a}x^2 + 2\overline{c}y^2 = 1$ does not have a solution in $\mathbb{Z}/8\mathbb{Z}$, then $(a,b)_2 = -1$.

Note that any solution to this equation in $\mathbb{Z}/8\mathbb{Z}$ must have $\overline{x} = 1 \bmod 2$, so the first term is $\overline{a} + 2\overline{c}y^2$. From here, one readily sees that there are no solutions in the cases:

- $a = 3 \bmod 8$ and $c = 1 \bmod 4$
- $a = 5 \bmod 8$
- $a = 7 \bmod 8$ and $c = 3 \bmod 4$

We compute the left hand side of (3.9) in these cases as:

- $(-1)^{1+1\cdot 0} = -1$
- $(-1)^{1+0\cdot 1} = -1$
- $(-1)^{0+1\cdot 1} = -1$

as desired.

We now treat the remaining cases:

Subcase 1. If $a = 1 \bmod 8$, then a is a square so $(a,b)_2 = 1 = (-1)^{0+0\cdot\vartheta(c)}$.

Subcase 2. If $a = 3 \bmod 8$ and $c = 3 \bmod 4$, then we need to show that $(a,b)_2 = 1 = (-1)^{1+1\cdot 1}$.

We note that:

$$a + b = a + 2c = 3 + 6 = 1 \bmod 8$$

and therefore $a+b$ is a square in $\mathbb{Z}_2$, so that $ax^2 + by^2$ represents a square and therefore 1.

Subcase 3. If $a = 7 \bmod 8$ and $c = 1 \bmod 4$, then we need to show that $(a,b)_2 = 1 = (-1)^{0+1\cdot 0}$.

We see that $a + b = 1 \bmod 8$, so as above, we deduce that $(a,b)_2 = 1$.

Case 3. $v_2(a) = v_2(b) = 1$.

As in Case 3 for the proof of Proposition 3.16.2, we have $(a,b)_2 = (a,-ab)_2$, and we note that $v_2(-ab)$ is even so we can compute this Hilbert symbol using Case 2 as:

$$(a,-ab)_2 = (-1)^{\varpi(-ab)+\vartheta(a)\vartheta(-ab)}.$$

Observe that $(-1)^{\vartheta(a)\vartheta(-ab)} = (-1)^{\vartheta(a)\vartheta(b)}$. Moreover, it is easy to check that we have:

$$(-1)^{\varpi(-ab)} = (-1)^{\varpi(a)} \cdot (-1)^{\varpi(b)}.$$

Therefore, we obtain:

$$(a,b)_2 = (a,-ab)_2 = (-1)^{\varpi(a)+\varpi(b)+\vartheta(a)\vartheta(b)}.$$

as desired.

□

Exercise 3.8. Deduce Theorem 3.14.5 from Propositions 3.16.1, 3.16.2 and 3.16.3.

3.17. Hasse-Minkowski invariant.

Let $k = \mathbb{Q}_p$ or $k = \mathbb{R}$. As always, if $k = \mathbb{R}$, e.g. a Hilbert symbol $(-,-)_p$ should be interpreted with $p = \infty$.

Theorem 3.17.1. *There is a unique invariant $\varepsilon_p(q) \in \{1,-1\}$ defined for every non-degenerate quadratic form q over k such that:*

- *$\varepsilon_p(q) = 1$ if q has rank 1.*
- *$\varepsilon_p(q_1 + q_2) = \varepsilon_p(q_1)\varepsilon_p(q_2) \cdot (\operatorname{disc}(q_1), \operatorname{disc}(q_2))_p$.*

Remark 3.17.2. This result is proved in elementary ways in [Ser] and [Cas]. We choose not to prove it here because a proper argument would require too much of a digression on Brauer groups. This is justified because Corollary 3.17.4 below gives an algorithm on how to compute it; the only question is well-definedness.

Remark 3.17.3. We call this invariant the *Hasse-Minkowski* invariant of q.

Corollary 3.17.4. *For q diagonalized as $q(x_1, \dots, x_n) = \sum_i a_i x_i^2$, we have:*

$$\varepsilon_p(q) = \prod_{i<j} (a_i, a_j)_p.$$

Exercise 3.9. Deduce the corollary from Theorem 3.17.1.

3.18. Using the Hasse-Minkowski invariant, we can complete the description of §3.15 describing when a quadratic form q over $k = \mathbb{Q}_p$ represents some $\lambda \in \mathbb{Q}_p$.

Proposition 3.18.1. *A quadratic form q over $\mathbb{Q}_p$ represents $0 \neq \lambda$ if and only if one of the following results holds:*

- $\operatorname{rank}(q) = 1$ *and* $\lambda = \operatorname{disc}(q) \in k^\times/(k^\times)^2$.
- $\operatorname{rank}(q) = 2$ *and* $(-\operatorname{disc}(q), \lambda) = \varepsilon_p(q)$.
- $\operatorname{rank}(q) = 3$ *and either* $\lambda \neq -\operatorname{disc}(q)$ *or* $\lambda = -\operatorname{disc}(q)$ *and* $(-\operatorname{disc}(q), -1)_p = \varepsilon_p(q)$.
- $\operatorname{rank}(q) \geqslant 4$.

Proof. The case of rank 1 is obvious: it is true over any field. The case of rank 2 is treated in Corollary 3.15.3, and the case of rank 4 is treated in Corollary 3.15.7. The case of $\operatorname{rank}(q) = 3$ and $\lambda \neq -\operatorname{disc}(q)$ is treated in Corollary 3.15.5.

This leaves us to show that a non-degenerate ternary form q respresents $-\operatorname{disc}(q)$ if and only if $(-\operatorname{disc}(q), -1)_p = \varepsilon_p(q)$.

Let $q(x, y, z) = ax^2 + by^2 + cz^2$. That q represents $-\operatorname{disc}(q) = -abc$ is equivalent to saying that $q(x, y, z, w) = ax^2 + by^2 + cz^2 + abcw^2$ is isotropic, which in turn is equivalent to saying that $q_1(x, y) = ax^2 + by^2$ and $q_2(z, w) = -cz^2 - abcw^2$ represent a common value.

Note that $\operatorname{disc}(q_1) = \operatorname{disc}(q_2)$. Therefore, by Corollary 3.15.3, saying that q_1 and q_2 represent a common value is equivalent to:

$$(a, b)_p = \varepsilon_p(q_1) = \varepsilon_p(q_2) = (-abc, -c)_p.$$

Using Theorem 3.14.5, we find:

$$(-abc, -c)_p = (-abc, c)_p \cdot (-abc, -1)_p.$$

Moreover, by Proposition 3.14.3 (6), we have $(-abc, c) = (ab, c) = (a, c) \cdot (b, c)$, so that the above is equivalent to:

$$\varepsilon_p(q) = (a, b)_p (a, c)_p (b, c)_p = (-abc, -1)_p = (-\operatorname{disc}(q), -1)_p$$

as desired.

□

Corollary 3.18.2. *For $k = \mathbb{Q}_p$, two non-degenerate quadratic forms q_1 and q_2 over k are equivalent if and only if* $\operatorname{disc}(q_1) = \operatorname{disc}(q_2)$ *and* $\varepsilon_p(q_1) = \varepsilon_p(q_2)$.

Proof. By Proposition 3.18.1, two quadratic forms q_1 and q_2 with the same discriminant and Hasse-Minkowski invariant represent the same values. Choose some λ that they both represent.

Then $q_1 = \lambda x^2 + q_1'$ and $q_2 = \lambda x^2 + q_2'$ by our usual method. Then we have:

$$\begin{aligned} \operatorname{disc}(q_1') = \lambda \operatorname{disc}(q_1) = \lambda \operatorname{disc}(q_2) = \operatorname{disc}(q_2') \in k^\times/(k^\times)^2 \\ \text{and } \varepsilon_p(q_1') = \varepsilon_p(q_1)(\operatorname{disc}(q_1'), \lambda)_p = \varepsilon_p(q_2)(\operatorname{disc}(q_2'), \lambda)_p = \varepsilon_p(q_2') \end{aligned}$$

so we obtain the result by induction.

□

Remark 3.18.3. As a consequence, we obtain the following nice interpretation of the Hasse-Minkowski invariant: q has $\varepsilon_p(q) = 1$ if and only if q is equivalent to a form:

$$x_1^2 + \ldots + x_{n-1}^2 + dx_n^2.$$

Indeed, this is clearly only possible with $d = \operatorname{disc}(q)$, and in that case, we see that this form has the same discriminant as q, and it obviously has $\varepsilon_p = 1$.

4. The Hasse principle

4.1. The goal for this section is to prove the following theorem.

Theorem 4.1.1 (Hasse principle). *A non-degenerate quadratic form q over $\mathbb{Q}$ represents a value $\lambda \in \mathbb{Q}$ if and only if the extension of scalars $q_{\mathbb{Q}_p}$ represents $\lambda \in \mathbb{Q} \subseteq \mathbb{Q}_p$ for every prime p, and $q_{\mathbb{R}}$ represents λ.*

4.2. We will prove Theorem 4.1.1 by treating different ranks n of q separately.

Exercise 4.1. Deduce the $n = 1$ case of Theorem 4.1.1 from the discussion of Example 2.4.3 (4).

We will give two proofs of the $n = 2$ case below, one from [Cas] using Minkowski's *geometry of numbers*, and a second from [Ser] that is more direct.

4.3. **Geometry of numbers.** We now give a quick crash course in Minkowski's geometry of numbers, following [Cas] Chapter 5.

Definition 4.3.1. A *lattice* is an abelian group Λ isomorphic to $\mathbb{Z}^n$ for some integer n, i.e., it is a finitely generated torsion-free abelian group. The integer n is called the *rank* of Λ.

A *metrized lattice* is a pair $(\Lambda, q_{\mathbb{R}})$ of a lattice Λ and a positive-definite quadratic form $q_{\mathbb{R}}$ defined on $\Lambda_{\mathbb{R}} \subseteq \Lambda \underset{\mathbb{Z}}{\otimes} \mathbb{R}$.

Remark 4.3.2 (Discriminants). Let $(\Lambda, q_{\mathbb{R}})$ be a metrized lattice. Let $e_1, \ldots, e_n$ be a basis of Λ. Then we can associate to this datum the matrix:

$$A = \begin{pmatrix} \frac{1}{2}B_{q_{\mathbb{R}}}(e_1, e_1) & \frac{1}{2}B_{q_{\mathbb{R}}}(e_2, e_1) & \ldots & \frac{1}{2}B_{q_{\mathbb{R}}}(e_n, e_1) \\ \frac{1}{2}B_{q_{\mathbb{R}}}(e_1, e_2) & \frac{1}{2}B_{q_{\mathbb{R}}}(e_2, e_2) & \ldots & \frac{1}{2}B_{q_{\mathbb{R}}}(e_n, e_2) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{2}B_{q_{\mathbb{R}}}(e_1, e_n) & \frac{1}{2}B_{q_{\mathbb{R}}}(e_2, e_n) & \ldots & \frac{1}{2}B_{q_{\mathbb{R}}}(e_n, e_n) \end{pmatrix}.$$

As in Exercise 2.9, a change of basis corresponding to a matrix $S \in GL_n(\mathbb{Z})$ changes A by:

$$A \mapsto S^T AS.$$

Because $\det(S) \in \mathbb{Z}^\times = \{1, -1\}$, we see that $\operatorname{disc}\big((\Lambda, q_{\mathbb{R}})\big) := \det(A) \in \mathbb{R}^\times$ is well-defined.

Remark 4.3.3. Because $q_{\mathbb{R}}$ is assumed to be a positive-definite quadratic space, $(\Lambda_{\mathbb{R}}, q_{\mathbb{R}})$ is isomorphic to $(\mathbb{R}^n, x_1^2 + \ldots + x_n^2)$. We can therefore talk about open sets in $\Lambda_{\mathbb{R}}$ and volumes of such open sets, since these notions do not depend on the given isomorphism (i.e., they are invariants of $\mathbb{R}^n$ under the action of the orthogonal group for the form $x_1^2 + \ldots + x_n^2$).

Remark 4.3.4. The discriminant in the sense above refines $\operatorname{disc}(q_{\mathbb{R}}) \in \mathbb{R}^\times/(\mathbb{R}^\times)^2 = \{1, -1\}$, and therefore we see that (because $q_{\mathbb{R}}$ is assumed positive definite) that $\operatorname{disc}\big((\Lambda, q_{\mathbb{R}})\big) > 0$.

Remark 4.3.5. If V is a finite-dimensional vector space over $\mathbb{R}$, we say that a subset U of V is *convex* if $x, y \in U$ implies that $tx + (1-t)y \in U$ for all $t \in [0,1]$. We say that U is *symmetric* if $x \in U$ implies that $-x \in U$.

Theorem 4.3.6 (Minkowski's theorem)**.** *Let $(\Lambda, q_{\mathbb{R}})$ be a metrized lattice of rank n and let $U \subseteq \Lambda_{\mathbb{R}}$ be a convex symmetric open set with:*

$$\operatorname{vol}(U) > 2^n \cdot \sqrt{\operatorname{disc}\big((\Lambda, q_{\mathbb{R}})\big)}.$$

Then U contains a non-zero point of Λ.

Proof. For convenience, we use the language of measure theory.

Note that $T := \Lambda_{\mathbb{R}}/\Lambda$ inherits a canonical measure μ_T, where the measure of a "small" connected subset T is by definition the measure of a connected component of the inverse image in $\Lambda_{\mathbb{R}}$.

Indeed, one can choose a *fundamental domain* F in $\Lambda_{\mathbb{R}}$ for the action of Λ and restrict the usual measure on $\Lambda_{\mathbb{R}}$. E.g., choosing a basis $e_1, \ldots, e_n$ for Λ, we can choose F to be:

$$\{x = \sum x_i e_i \in \Lambda_{\mathbb{R}}, x_i \in \mathbb{R} \mid 0 \leqslant x_i < 1\}.$$

Clearly μ is an additive measure on the torus T, i.e., $\mu(x + V) = \mu(V)$ for every open set $V \subseteq T$ and $x \in T$.

We claim:

$$\mu(T) = \sqrt{\operatorname{disc}\big((\Lambda, q_{\mathbb{R}})\big)}. \tag{4.1}$$

Indeed, choose an orthonormal basis $f_1, \ldots, f_n$ of $\Lambda_{\mathbb{R}}$, and define the matrix $B \in M_n(\mathbb{R})$ to have ith column the vectors e_i written in the basis $f_1, \ldots, f_n$. By the usual characterization of volumes in terms of determinants, we have:

$$|\det(B)| = \operatorname{vol}(F) = \mu(T).$$

But for A as in Remark 4.3.2, we obviously have $B^T B = A$, and therefore:

$$\operatorname{disc}\big((\Lambda, q_{\mathbb{R}})\big) = \det(A) = \det(B^T B) = \det(B)^2$$

as was claimed in (4.1).

Let p denote the quotient map $\Lambda_{\mathbb{R}} \to T$. Define the function:

$$\chi : T \to \mathbb{R} \cup \{\infty\}$$

by letting $\chi(x)$ be the number of points in $\frac{1}{2}U \cap p^{-1}(x)$.[10] Note that χ is a measurable function.

Then we have:

[10] One can prevent this function from possibly taking the value ∞ by intersecting U with a large disc around 0, but it's not a big deal.

$$\int_T \chi(x)d\mu(x) = \mathrm{vol}(\frac{1}{2}U) = \frac{1}{2^n}\mathrm{vol}(U) > \mathrm{disc}\big((\Lambda, q_{\mathbb{R}})\big) = \mu(T).$$

Therefore, $\chi(x) > 1$ for some $x \in T$. Since χ takes integer values, we must have $\chi(x) \geqslant 2$, and therefore there exist distinct points $y, z \in \frac{1}{2}U$ with $p(y) = p(z) = x$. Then we have:

$$0 \neq y - z = \frac{1}{2}(2y) + \frac{1}{2}(-2z) \in U$$

by symmetry and convexity of U, and also $y - z \in \Lambda$ since $p(y - z) = p(y) - p(z) = 0$.

Therefore, $y - z$ is a non-zero point of $\Lambda \cap U$, as claimed.

□

Exercise 4.2. For $(\Lambda, q_{\mathbb{R}})$ a metrized lattice and $\Lambda' \subseteq \Lambda$ a subgroup of index $n < \infty$, prove that:

$$\mathrm{disc}\big((\Lambda', q_{\mathbb{R}})\big) = n^2\,\mathrm{disc}\big((\Lambda, q_{\mathbb{R}})\big).$$

4.4. **Application of geometry of numbers to the Hasse principle.** The above theorem can be used to prove Theorem 4.1.1 in the $n = 2$ case in general: we refer to [Cas] §6.4 where this is effected. However, to simplify the explanation given in [Cas], we will only give the argument in a special case (that at least shows how things are done, and we'll treat the general $n = 2$ case by a simpler method following [Ser] in §4.6).

First, note that any binary quadratic form $q(x, y) = ax^2 + by^2$ is equivalent to one where $v_p(a), v_p(b) \in \{0, 1\}$ for every prime p. Indeed, multiplying a and b by squares, we easily put them into this form.

Let $q(x, y) = ax^2 + by^2$ be a binary quadratic form with a and b square-free as above, and, moreover assuming: (1) $a, b \in \mathbb{Z}$ relatively prime, (2) a, b odd, and (3) $a + b = 0 \bmod 4$. We will prove that if q represents 1 in every $\mathbb{Q}_p$ then it does so in $\mathbb{Q}$.

For each prime p dividing a, the fact that $ax^2 + by^2$ represents 1 means that we can choose some $r_p \in \mathbb{Z}$ with $br_p^2 = 1$, and similarly for each prime p dividing b, we can choose some $r_p \in \mathbb{Z}$ with $ar_p^2 = 1$. Indeed, this follows e.g. from Proposition 3.16.2.

Then define $\Lambda \subseteq \mathbb{Z}^3$ as consisting of the triples $(x, y, z) \in \mathbb{Z}^3$ with:

$$\begin{cases} y = r_p z & \text{if } p \mid a \\ x = r_p z & \text{if } p \mid b \\ x = y \bmod 2 \\ 2 \mid z. \end{cases}$$

We consider Λ as a metrized lattice by embedding $\mathbb{Z}^3$ into $\mathbb{R}^3$ in the usual way. By Exercise 4.2, we have $\mathrm{disc}(\Lambda) = (2ab)^4$.

Observe that for $(x, y, z) \in \Lambda$ and p a prime dividing a, we have:

$$ax^2 + by^2 = br_p^2 z^2 = z^2 \bmod p$$

and similarly for primes dividing b. Moreover, we have:

$$ax^2 + by^2 = 0 = z^2 \bmod 4$$

where the first equality follows because $b = -a \bmod 4$, so $ax^2 + by^2 = ax^2 - ay^2 = a(x-y)(x+y) = 0 \bmod 4$ because $x - y = 0 \bmod 4$.

Therefore, by the Chinese remainder theorem and the square-free assumption of a and b, we have:

$$ax^2 + by^2 = z^2 \bmod 4|ab|. \tag{4.2}$$

Now define:

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid |a|x^2 + |b|y^2 + z^2 < 4|ab|.\}$$

Then U is obviously convex and symmetric, and moreover, we find:

$$\operatorname{vol}(U) = \frac{4}{3}\pi \frac{(4|ab|)^{\frac{3}{2}}}{\sqrt{|a|}\sqrt{|b|}} = 2^3 \cdot \frac{\pi}{3} 4|ab| > 2^3 \cdot 4|ab| = 2^3 \operatorname{disc}(\Lambda).$$

Therefore, Minkowski's theorem guarantees the existence of non-zero (x, y, z) in Λ and U. Then we have:

$$ax^2 + by^2 - z^2 \in 4|ab|\mathbb{Z}$$

by (4.2), but lying in U implies that:

$$|ax^2 + by^2 - z^2| \leqslant |a|x^2 + |b|y^2 + z^2 < 4|ab|$$

and therefore we must have $ax^2 + by^2 = z^2$ as desired.

4.5. **Digression: Fermat and Legendre's theorems.** We now digress temporarily to prove two classical results.

Theorem 4.5.1 (Fermat's theorem)**.** *An odd prime p can be written as the sum of two squares $p = x^2 + y^2$ $(x, y \in \mathbb{Z})$ if and only if $p = 1 \bmod 4$.*

Proof. Necessity follows from congruences mod 4. For sufficiency, suppose $p = 1 \bmod 4$. In this case, we can choose an integer r with $r^2 = -1 \bmod p$.

Define the lattice $\Lambda \subseteq \mathbb{Z}^2$ to consist of those (x, y) with $y = rx \bmod p$. Note that for $(x, y) \in \Lambda$, we have:

$$x^2 + y^2 = x^2 + r^2x^2 = 0 \bmod p.$$

Moreover, observe that the discriminant of the metrized latticed Λ is p^2.

Define $U \subseteq \mathbb{R}^2$ to be the disc:

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 2p.\}$$

Then the area of U is $2p \cdot \pi$.

Now observing that:

$$2p \cdot \pi > 4p$$

because $\pi > 2$, Minkowski's theorem applies, and we deduce the existence of $(x, y) \neq (0, 0)$ in $U \cap \Lambda$.

But then $p \mid x^2 + y^2$ and $0 < x^2 + y^2 < 2p$, so we must have $x^2 + y^2 = p$.

□

Theorem 4.5.2 (Legendre's theorem)**.** *Any integer $n \geqslant 0$ can be written as the sum of four squares:*

$$n = x^2 + y^2 + z^2 + w^2, \quad x, y, z, w \in \mathbb{Z}.$$

Proof. First, observe that it suffices to treat the case where p is a prime. Indeed, it is a easy, elementary identity to verify that the product of numbers that are each the sum of four squares is again of the same form (one uses the norm form on the quaternions to actually reconstruct this identity in practice).

By Proposition 3.13.1, we can find integers r and s with $r^2 + s^2 = -1 \bmod p$.

We then define the lattice:

$$\Lambda = \{(x, y, z, w) \in \mathbb{Z}^4 \mid z = (rx + sy) \bmod p, w = (rx - sy) \bmod p\}.$$

Then for $(x, y, z, w) \in \Lambda$, observe that:

$$x^2 + y^2 + z^2 + w^2 = x^2 + y^2 + (rx + sy)^2 + (rx - sy)^2 = x^2 + y^2 + (r^2 + s^2)(x^2 + y^2) = 0 \bmod p.$$

Observe that Λ has discriminant p^4.

Now let $U \subseteq \mathbb{R}^4$ be the disc $\{(x, y, z, w) \in \mathbb{R} \mid x^2 + y^2 + z^2 + w^2 < 2p\}$ of radius $\sqrt{2p}$. We recall that in $\mathbb{R}^4$, the disc of radius 1 has volume $\frac{\pi^2}{2}$. Therefore, U has volume $2\pi^2 p^2$.

Now, because $\pi^2 > 8$, we have:

$$2\pi^2 p^2 > 2^4 p^2$$

and therefore Minkowski's theorem applies. Clearly any non-zero vector in $\Lambda \cap U$ solves our equation, giving the desired result.

□

4.6. **Hasse-Minkowski for $n = 2$.** We now prove the Hasse principle for binary quadratic forms following [Ser].

Proof of Theorem 4.1.1 for binary forms. Let $q(x, y) = ax^2 + by^2$. Up to rescaling, it suffices to show that q represents 1 if and only if $(a, b)_p = 1$ for all p and $(a, b)_\infty = 1$. As in §4.4, we can assume a and b are (possibly negative) square-free integers. Without loss of generality, we can assume $|a| \geqslant |b|$.

We proceed by induction on $|a| + |b|$.

The base case is $|a| + |b| = 2$, i.e., $|a| = |b| = 1$. Then we must have $a = 1$ or $b = 1$, since otherwise $a = b = -1$ and $(a, b)_\infty = -1$, and the result is clear.

We now perform the inductive step. We can assume $|a| > 1$.

We first claim that b is a square in $\mathbb{Z}/a\mathbb{Z}$.

By the Chinese remainder theorem and by the assumption that a is square-free, it suffices to show that if p is a prime dividing a, then b is a square mod p. If $p = 2$, this is clear, since everything is a square in $\mathbb{Z}/2\mathbb{Z}$.

Therefore, let p be an odd prime dividing a. If p divides b, then $b = 0 \bmod p$ and we're done. Otherwise, b is a unit in $\mathbb{Z}_p$, and therefore $(a, b)_p = 1$ implies that b is a square by Proposition 3.16.2.

That b is a square in $\mathbb{Z}/a\mathbb{Z}$ means that we can find $s, t \in \mathbb{Z}$ such that:

$$t^2 - b = sa.$$

Certainly we can choose t with $0 \leqslant t \leqslant \frac{|a|}{2}$.

We claim that for $p = \infty$ or p a prime number,[11] we have:

[11] Not to be confused with our earlier use of p.

$$(b,s)_p = 1. \tag{4.3}$$

Indeed, first observe that $(b, t^2 - b)_p = 1$ because:

$$b \cdot (\frac{1}{t})^2 + (t^2 - b) \cdot (\frac{1}{t})^2 = 1.$$

Then we observe that:

$$(b,s)_p = (b, \frac{t^2 - b}{a})_p = (b, a(t^2 - b))_p = (b,a)_p (b, t^2 - b)_p = 1 \cdot 1 = 1.$$

Moreover, we have:

$$|s| = |\frac{t^2 - b}{a}| \leqslant \frac{t^2}{|a|} + \frac{|b|}{|a|} \leqslant \frac{|a|}{4} + 1 < |a|.$$

where the last inequality holds because $|a| \geqslant 2$, and the first inequality holds because $t \leqslant \frac{|a|}{2}$ and $|b| \leqslant |a|$.

Then $|a| + |b| > |s| + |b|$, and therefore, since $(b,s)_p = 1$ for all p, by induction there exist rational numbers x and y such that:

$$sx^2 + by^2 = 1. \tag{4.4}$$

In this case, we claim:

$$a(sx)^2 + b(1 - ty)^2 = (t - by)^2 \tag{4.5}$$

meaning that q represents a square, giving the desired result. More precisely: if $t - by \neq 0$, we're clearly done, and if $t - by = 0$, then we have shown that q is a hyperbolic and therefore it represents 1.[12]

To verify this algebra, we compute:

$$(as)(sx^2) + b(1 - ty)^2 = (t^2 - b)(1 - by^2) + b - 2byt + bt^2y^2 =$$
$$t^2 - b - bt^2y^2 + b^2y^2 + b - 2byt + bt^2y^2 = t^2 + b^2y^2 - 2byt = (t - by)^2.$$

□

Remark 4.6.1. The end of the proof could be explained in the following more conceptual way. First, for a field k and fixed $b \in k^\times$ not a square, to say that $ax^2 + by^2 = 1$ is equivalent to $a = (\frac{1}{x})^2 - b(\frac{y}{x})^2$, and from here one easily finds that the solvability of this equation is equivalent to say that *a is a norm for the extension $k[\sqrt{b}]$ of k*. In this way, using the multiplicativity of the norm map, one sees that the set of a for which this equation is solvable forms a group under multiplication.

Now, in the notation of the proof, we could have instead said that the identity $as = t^2 - b$ says that as is a norm, and the identity $sx^2 + by^2 = 1$ says that s is a norm, which in turn implies that s^{-1} is a norm, so this implies that $a = as \cdot s^{-1}$ is a norm, as desired. Indeed, the elementary but random identity verified at the end of the argument translates to this argument.

By the same method, one could verify (4.3) more directly.

[12] To deduce that q is hyperbolic, we need to be careful that in (4.5) we have sx and $1 - ty$ non-zero. But (4.4) and the square-free assumption on b force sx to be non-zero, giving the claim.

4.7. **Quadratic reciprocity.** We will need to appeal to Gauss's quadratic reciprocity law below. We recall the statement and proof here.

Theorem 4.7.1. *(1) For p and q distinct odd primes, we have:*

$$\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\left(\frac{q}{p}\right).$$

(2) For p an odd prime, we have:

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}.$$

Remark 4.7.2. We note that the complicated-looking expression $(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}$ produces the value 1 unless $p = q = 3 \bmod 4$, in which case it gives the value -1.

Proof of Theorem 4.7.1 (1). We follow the elementary proof of [Ser].

Let $\overline{\mathbb{F}}_p$ denote an algebraic closure of $\mathbb{F}_p$ in what follows.

Step 1. Here is the strategy of the proof below. We will write down the Gauss sum $\mathcal{G} \in \overline{\mathbb{F}}_p$, and verify first $\mathcal{G}$ is a square root of $(-1)^{\frac{q-1}{2}} \cdot q$, and second that $\mathcal{G}$ lies in $\mathbb{F}_p$ exactly when $\left(\frac{p}{q}\right) = 1$.

This means that $(-1)^{\frac{q-1}{2}} \cdot q$ is a quadratic residue modulo p exactly when $\left(\frac{p}{q}\right) = 1$, so that:

$$1 = \left(\frac{(-1)^{\frac{q-1}{2}}q}{p}\right)\cdot\left(\frac{p}{q}\right) = \left(\frac{(-1)^{\frac{q-1}{2}}}{p}\right)\left(\frac{q}{p}\right)\left(\frac{p}{q}\right).$$

Then noting that[13] $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$, we obtain the result.

Step 2. Let $\zeta_q \in \overline{\mathbb{F}}_p$ be a solution to $x^q = 1$ other than 1; such a ζ_q exists because $x^q - 1$ is a separable polynomial over $\mathbb{F}_p$.

We define the *Gauss sum* $\mathcal{G} \in \overline{\mathbb{F}}_p$ as:

$$\mathcal{G} = \sum_{x \in \mathbb{F}_q^\times} \left(\frac{x}{q}\right)\cdot \zeta_q^x.$$

Step 3. We claim that $\mathcal{G}^2 = (-1)^{\frac{q-1}{2}}q$.

Indeed, we compute:

$$\mathcal{G}^2 = \sum_{x,y \in \mathbb{F}_q^\times}\left(\frac{x}{q}\right)\left(\frac{y}{q}\right)\zeta_q^{x+y} = \sum_{x,y\in\mathbb{F}_q^\times}\left(\frac{xy}{q}\right)\zeta_q^{x+y} = \sum_{z\in\mathbb{F}_q^\times}\left(\frac{z}{q}\right)\sum_{x\in\mathbb{F}_q^\times}\zeta_q^{x+xz}.$$

Here we have applied the change of variables $z = \frac{y}{x}$, noting that:

$$\left(\frac{\frac{y}{x}}{p}\right) = \left(\frac{\frac{x}{y}}{p}\right) = \left(\frac{xy}{p}\right).$$

[13] Proof: Under an isomorphism $\mathbb{F}_q^\times \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, clearly -1 corresponds to $\frac{p-1}{2}$, and this element is a multiple of 2 if and only if $4 \mid p-1$.

Now observe that for $z \neq -1$, $\zeta_q^{x+xz} = (\zeta_q^{z+1})^x$ ranges exactly over the primitive qth roots of unity in $\overline{\mathbb{F}}_p$. Because the sum of all the qth roots of unity is 0, this means that for $z \neq 1$ the corresponding summand is -1. Therefore, the above sum is:

$$\left(\frac{-1}{q}\right) \sum_{x \in \mathbb{F}_q^\times} 1 + \sum_{z \in \mathbb{F}_q \setminus \{0,-1\}} - \left(\frac{z}{q}\right).$$

where the first term corresponds to $z = -1$ and the rest corresponds to $z \neq -1$. We note that $\sum_{z \in \mathbb{F}_q^\times} \left(\frac{z}{q}\right) = 0$, since there are equal numbers of quadratic residues and non-residues, so the second term above is $\left(\frac{-1}{q}\right)$. Clearly the first term sums to $(q-1)\left(\frac{-1}{q}\right)$, giving the result.

Step 4. Next, we recall a general technique for checking if an element of $\overline{\mathbb{F}}_p$ lies in $\mathbb{F}_p$.

For k a field of characteristic p, note that by the binomial formula that $\mathrm{Fr}_p(x) := x^p$ is a homomorphism.

Then we claim that $x \in \overline{\mathbb{F}}_p$ lies in $\mathbb{F}_p$ if and only if $x^p = x$. Indeed, this is obviously satisfied for elements of $\mathbb{F}_p$, and for degree reasons there are at most p roots of this polynomial.

Step 5. It remains to show that $\mathcal{G}$ lies in $\mathbb{F}_p$ exactly when $\left(\frac{p}{q}\right) = 1$. It suffices to check when $\mathcal{G}$ is fixed by the Frobenius in $\overline{\mathbb{F}}_p$.

We have:

$$\mathrm{Fr}_p(\mathcal{G}) = \sum_{x \in \mathbb{F}_q^\times} \left(\frac{x}{q}\right)^p \cdot \zeta_q^{px} = \sum_{x \in \mathbb{F}_q^\times} \left(\frac{x}{q}\right) \cdot \zeta_q^{px} = \sum_{x \in \mathbb{F}_q^\times} \left(\frac{p^{-1}x}{q}\right) \cdot \zeta_q^{x} = \left(\frac{p}{q}\right) \mathcal{G}$$

as desired.

□

Proof of Theorem 4.7.1 (2). The strategy is similar to that of Theorem 4.7.1 (1): we find a convenient expression for $\sqrt{2} \in \overline{\mathbb{F}}_p$ and then test by the Frobenius when it lies in $\mathbb{F}_p$.

Let $\zeta_8 \in \overline{\mathbb{F}}_p$ be a primitive 8th root of unity. We claim that $\zeta_8 + \zeta_8^{-1}$ is a square root of 2.[14]

Indeed, we compute:

$$(\zeta_8 + \zeta_8^{-1})^2 = (\zeta_8^2 + 2 + \zeta_8^{-2}).$$

Observing that ζ_8^2 and ζ_8^{-2} are distinct square roots of -1, their sum must be zero, giving the claim.

Now we observe that:

$$\mathrm{Fr}_p(\zeta_8 + \zeta_8^{-1}) = \zeta_8^p + \zeta_8^{-p}$$

which equals $\zeta_8 + \zeta_8^{-1}$ if and only if $p = 1, -1 \bmod 8$. This is equivalent to requiring that $(-1)^{\frac{p^2-1}{8}} = 1$.

□

[14] As a heuristic for this expression, note that over the complex numbers $\frac{\sqrt{2}}{2} + i \cdot \frac{\sqrt{2}}{2}$ is visibly an 8th root of unity ζ_8, and we have $\zeta_8 + \zeta_8^{-1} = \zeta_8 + \overline{\zeta}_8 = 2\,\mathrm{Re}(\zeta_8) = \sqrt{2}$.

4.8. **Global properties of the Hasse-Minkowski invariant.** We have the following result, which we will see is essentially a repackaging of the quadratic reciprocity law.

Proposition 4.8.1 (Hilbert reciprocity)**.** *Let $a, b \in \mathbb{Q}^\times$.*

For almost every[15] *prime p, we have:*

$$(a,b)_p = 1.$$

Moreover, we have:

$$(a,b)_\infty \cdot \prod_{p \text{ } prime} (a,b)_p = 1.$$

Proof. For the first claim, note that if p is an odd prime with $v_p(a) = v_p(b) = 0$, then Proposition 3.16.2 implies that $(a,b)_p = 1$.

For the second part, we proceed by cases, applying Propositions 3.16.2 and 3.16.3 freely.

Case 1. $a = p$, $b = q$ are distinct odd prime numbers.

Then we have:

$$\begin{cases} (p,q)_2 = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} \\ (p,q)_p = \left(\frac{q}{p}\right) \\ (p,q)_q = \left(\frac{p}{q}\right) \\ (p,q)_\ell = 1 & \text{if } \ell \neq 2, p, q \\ (p,q)_\infty = 1. \end{cases}$$

Then Theorem 4.7.1 (1) immediately gives the claim.

Case 2. $a = b = p$ an odd prime number.

Then we have:

$$\begin{cases} (p,p)_2 = (-1)^{(\frac{p-1}{2})^2} = (-1)^{\frac{p-1}{2}} \\ (p,p)_p = \left(\frac{-1}{p}\right) \\ (p,p)_\ell = 1 & \text{if} \ell \neq p \\ (p,p)_\infty = 1 \end{cases}$$

giving the claim.

Case 3. $a = p$ is an odd prime and $b = 2$.

Then we have:

$$\begin{cases} (p,2)_2 = (-1)^{\frac{p^2-1}{8}} \\ (p,2)_p = \left(\frac{2}{p}\right) \\ (p,2)_\ell = 1 & \text{if } \ell \neq 2, p. (p,2)_\infty = 1. \end{cases}$$

Then Theorem 4.7.1 (2) gives the claim.

[15] This phrase means "all but finitely many."

Case 4. $a = b = 2$.

Then we have $(2, 2)_p = 1$ for p prime and for $p = \infty$ because:

$$2(\frac{1}{2})^2 + 2(\frac{1}{2})^2 = 1.$$

Case 5. $a = p$ is an odd prime and $b = -1$.

Then we have:

$$\begin{cases} (p, -1)_2 = (-1)^{\frac{p-1}{2}} & \\ (p, -1)_p = \left(\frac{-1}{p}\right) & \\ (p, -1)_\ell = 1 & \text{if } \ell \neq 2, p \\ (p, -1)_\infty = 1 & \end{cases}$$

giving the result (one can alternatively reduce this case to Case 2).

Case 6. $a = 2$ and $b = -1$.

Then we have $(2, -1)_p = 1$ for all p prime and $p = \infty$ because:

$$2 \cdot (1)^2 - (1^2) = 1.$$

Case 7. $a = -1$ and $b = -1$.

Then we have:

$$\begin{cases} (-1, -1)_2 = -1 & \\ (-1, -1)_\ell = 1 & \text{if } \ell \neq 2 \\ (-1, -1)_\infty = -1 & \end{cases}$$

giving the result.

Case 8. General case.

By the bimultiplicativity of the Hilbert symbol (Theorem 3.14.5 (1)), we reduce to the above cases.

□

Remark 4.8.2. Hilbert's form of the quadratic reciprocity law should be viewed as analogous to the conclusion of Exercise 3.3.

Corollary 4.8.3. *For any quadratic form q over $\mathbb{Q}$, we have:*

$$\varepsilon_\infty(q) \cdot \prod_{p \text{ } prime} \varepsilon_p(q) = 1.$$

Corollary 4.8.4. *If q is a non-degenerate binary quadratic form over $\mathbb{Q}$, q represents $\lambda \in \mathbb{Q}$ if and only if it represents it in each $\mathbb{Q}_p$ and in $\mathbb{R}$* with possibly one exception.

Proof. Indeed, we have seen in Proposition 3.15.1 that q represents λ in $\mathbb{Q}_p$ (allowing $p = \infty$ to correspond to $\mathbb{R}$) if and only if $(-\operatorname{disc}(q), \lambda)_p = \varepsilon(q)$.

Therefore, Hilbert reciprocity implies that if q represents λ locally with one exception, then q represents λ locally with no exceptions, and therefore, the Hasse principle for $n = 2$ implies that q represents λ in $\mathbb{Q}$.

□

4.9. We now conclude the proof of the Hasse principle.

Lemma 4.9.1. *Let S be a finite set consisting of primes and possibly the symbol ∞.*

Suppose that we are given $x_p \in \mathbb{Q}_p^\times$ for every $p \in S$ (where $\mathbb{Q}_\infty$ is understood as $\mathbb{R}$).

Then there exists a prime ℓ and $x \in \mathbb{Q}$ such that:

$$\begin{aligned} x \in x_p(\mathbb{Q}_p^\times)^2 &\textit{ for all } p \in S \\ x \in \mathbb{Z}_p^\times &\textit{ for all primes } p \notin S \cup \{\ell\}. \end{aligned} \tag{4.6}$$

Proof. We suppose $\infty \in S$ for convenience (certainly it does not harm either our hypothesis or our conclusion!). Let $\varepsilon \in \{1, -1\}$ be the sign of x_∞.

By Dirichlet's theorem on primes in arithmetic progressions, there exists a prime ℓ such that:

$$\ell = \varepsilon \cdot (p^{-v_p(x)} x_p) \cdot \prod_{\substack{q \in S \text{ prime} \\ q \neq p}} q^{-v_q(x_q)} \bmod p$$

for every odd prime p in S, and for $p = 2$, we impose the same congruence but require it to hold mod $p^3 = 8$.

We now define:

$$x = \varepsilon \cdot \ell \cdot \prod_{q \in S \text{ prime}} q^{v_q(x_q)} \in \mathbb{Q}.$$

Clearly x has no primes other than those in S and ℓ divides the numerator or denominator of x, so the second requirement of (4.6) holds for it.

We see that x has the same sign as x_∞. Moreover, for $p \in S$ an odd prime, we have:

$$x x_p^{-1} \in 1 + p\mathbb{Z}_p$$

and therefore is a square. For $p = 2$, we similarly find $x x_2^{-1} \in 1 + 8\mathbb{Z}_2$, and again, therefore is a square.

□

Completion of the proof of Theorem 4.1.1.

Case 1. To prove the $n = 3$ case of the Hasse principle, it suffices to show that any non-degenerate form q of rank 4 that is locally isotropic (i.e., isotropic over each $\mathbb{Q}_p$ and $\mathbb{R}$) is globally isotropic (i.e., isotropic over $\mathbb{R}$).

Let q be such a form. Diagonalizing q we see that we can write q as the difference of two binary quadratic forms $q_1 - q_2$.

Let S be the set consisting of 2, ∞, and the (finite) set of primes p with at least one of $v_p(\text{disc}(q_1))$ or $v_p(\text{disc}(q_2))$ odd.

Because q is locally isotropic, for every $p \in S$, we can find $\alpha_p, \beta_p \in \mathbb{Q}_p \times \mathbb{Q}_p$ (with $\mathbb{Q}_p$ meaning $\mathbb{R}$ for $p = \infty$) not both the zero vector such that $q_1(\alpha_p) = q_2(\beta_p)$. We define x_p to be the common value $q_1(\alpha_p) = q_2(\beta_p)$.

We can take x_p to be non-zero: indeed, e.g., if $\alpha_p \neq 0$ and $q_1(\alpha_p) = 0$, then q_1 is the hyperbolic plane and therefore represents any non-zero value represented by $q_2(\beta_p)$.

Therefore, we can find $x \in \mathbb{Q}^\times$ associated with these x_p as in Lemma 4.9.1. Let ℓ be defined as in *loc. cit.*

Observe that q_1 represents x in $\mathbb{R}$ and in each $\mathbb{Q}_p$, $p \neq \ell$. Indeed, for $p \in S$ this follows because q_1 represents x_p and therefore every element of $x_p(\mathbb{Q}_p^\times)^2$. Then for $p \notin S$, $p \neq \ell$: by Proposition 3.15.1, we needs to show that $(-\operatorname{disc}(q_1), x_p) = 1$. Because $p \notin S$, $v_p(\operatorname{disc}(q_1))$ is even, and therefore the coset $\operatorname{disc}(q_1) \in \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2$ is represented by an element of $\mathbb{Z}_p^\times$. Then because p is odd (by virtue of not being in S) and $x_p \in \mathbb{Z}_p^\times$, we obtain the required identity of Hilbert symbols from Proposition 3.16.2.

Now by Corollary 4.8.4, q_1 represents x over $\mathbb{Q}$, say, $q_1(\alpha) = x$. By the same logic, we can find $\beta \in \mathbb{Q} \times \mathbb{Q}$ with $q_2(\beta) = x$. Then $q_1(\alpha) - q_2(\beta) = 0$, proving that q was isotropic.

Case 2. We now treat the $n \geqslant 4$ case of the Hasse principle.

Again, it suffices to show a non-degenerate rational form q of rank $\geqslant 5$ that is locally isotropic is globally isotropic.

We again write $q = q_1 - q_2$, where q_1 is a non-degenerate binary form and q_2 is a form of rank $\geqslant 3$.

Let S be set consisting of ∞, 2, and that finite set of primes for which $v_p(\operatorname{disc}(q_2))$ is odd or $\varepsilon_p(q_2) = -1$. Because q is locally isotropic, we can find x_p and y_p vectors over $\mathbb{Q}_p$ with $q(x_p) = q(y_p) \neq 0$ for every $p \in S$.

By the Chinese remainder theorem, we can find $x \in \mathbb{Q}$ such that $q_1(x)q_1(x_p)^{-1} \in (\mathbb{Q}_p^\times)^2$ for every $p \in S$.[16]

As before, it suffices to show that q_2 represents $q_1(x)$ over $\mathbb{Q}$. By induction, it suffices to show that q_2 represents $q_1(x)$ locally.

This is clear for $p \in S$, in particular, for $p = 2$ or $p = \infty$. For all other primes p, we have:

$$\operatorname{disc}(q_2) \in \mathbb{Z}_p^\times \cdot (\mathbb{Q}_p^\times)^2 \subseteq \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2$$

so that $(-\operatorname{disc}(q_2), -1)_p = 1$ (by Proposition 3.16.2) and $\varepsilon_p(q_2) = 1$, and therefore we obtain that q_2 represents every value over $\mathbb{Q}_p$ by Proposition 3.18.1.

□

4.10. **Three squares theorem.** As an application, we now deduce the following result of Gauss.

Theorem 4.10.1. *Any integer $n \in \mathbb{Z}^{\geqslant 0}$ can be written as the sum $x_1^2 + x_2^2 + x_3^2$, $x_i \in \mathbb{Z}$, if and only if x is of the form $4^n(8k+7)$.*

Exercise 4.3.
1. Show that $q(x,y,z) = x^2 + y^2 + z^2$ represents every $\lambda \in \mathbb{Q}_p$ for p odd, and represents $\lambda \in \mathbb{Q}_2$ if and only if $-\lambda \in 2^{2\mathbb{Z}} \cdot (1 + 8\mathbb{Z}_2)$.
2. Deduce the rational version of Theorem 4.10.1 from the Hasse principle, i.e., the version in which x_1, x_2, x_3 are taken to be rational numbers instead of integers.

By Exercise 4.3, it suffices to pass from rational solutions to integral solutions. The following result gives a way to do this in some cases.

Proposition 4.10.2 (Hasse-Davenport). *Let Λ be a finite rank free abelian group equipped with an integral quadratic form*[17] *$q : \Lambda \to \mathbb{Z}$.*

Let $\Lambda_\mathbb{Q} = \Lambda \otimes \mathbb{Q}$, and we abuse notation in letting $q : \Lambda_\mathbb{Q} \to \mathbb{Q}$ denote the induced quadratic form.

Suppose that the quadratic form q is anisotropic as a quadratic form over $\mathbb{Q}$. Suppose moreover that for every $v \in \Lambda_\mathbb{Q} := \Lambda \otimes \mathbb{Q}$, there exists $w \in \Lambda$ with:

$$|q(v-w)| < 1.$$

[16] We emphasize that we do not need Lemma 4.9.1 here: we are not imposing any conditions away from the primes in S.

[17] Since we've only spoken about quadratic forms on vector spaces before, this warrants an explanation: all this means is that $q(nv) = n^2 q(v)$ for every $n \in \mathbb{Z}$ and $v \in \Lambda$.

Then q represents $\lambda \in \mathbb{Z}$ by a vector in Λ if and only if q represents λ with a vector in $\Lambda_{\mathbb{Q}}$.

Proof of Theorem 4.10.1. By Exercise 4.3, it suffices to see that $q(x, y, z) = x^2 + y^2 + z^2$ satisfies the hypotheses of Proposition 4.10.2.

Observe that q is anisotropic over $\mathbb{Q}$ because it is so over $\mathbb{R}$ (by positive-definiteness).

Then for every $(x, y, z) \in \mathbb{Q}^3$, we can find $(x', y', z') \in \mathbb{Z}^3$ with $|x - x'| \leqslant \frac{1}{2}$, and similarly for y and z. Then:

$$q(x - x', y - y', z - z') \leqslant (x - x')^2 + (y - y')^2 + (z - z')^2 \leqslant \frac{1}{4} + \frac{1}{4} + \frac{1}{4} < 1$$

as desired.

□

Proof of Proposition 4.10.2. Here is the strategy of the proof: we start we some $v \in \Lambda_{\mathbb{Q}}$ representing $\lambda \neq 0$, and we need to find some $v' \in \Lambda \subseteq \Lambda_{\mathbb{Q}}$ representing λ. We know from (the proof of) Proposition 2.14.2 that we can obtain *any* such $v' \in \Lambda_{\mathbb{Q}}$ with $q(v') = q(v)$ by acting on v by reflections. So we expect to find our v' by such a process.

On to the actual argument:

Choose $w \in \Lambda$ with $|q(v - w)| < 1$. If $q(v - w) = 0$, then by the anisotropic assumption, we have $v = w$ and we are done.

Otherwise, since $q(v - w) \neq 0$ we have the reflection s_{v-w} as in Example 2.13.1. So we define v' as:

$$v' = s_{v-w}(v) := v - \frac{B_q(v, v - w)}{q(v - w)} \cdot (v - w).$$

By Example 2.13.1, we have $q(v') = q(v)$.

We want to show that this process produces a vector in Λ after some number of iterations. To this end, let n be a positive integer such that $nv \in \Lambda \subseteq \Lambda_{\mathbb{Q}}$. It suffices to show that:

$$q(v - w) \cdot n \in \mathbb{Z} \tag{4.7}$$

and:

$$(q(v - w) \cdot n)v' \in \Lambda \tag{4.8}$$

since by assumption, $|q(v - w)n| < n$. Indeed, we then have a sequence of vectors $v, v', \ldots$ with $q(v) = q(v') = \ldots$, and a strictly decreasing sequence of positive integers $n, n' := |q(v - w)n|, \ldots$ with $nv \in \Lambda$, $n'v' \in \Lambda$, etc., and eventually we must obtain an honest element of Λ.

For (4.7), we compute:

$$q(v - w)n = (q(v) + q(w) - B_q(v, w))n = q(v)n + q(w)n - B_q(nv, w).$$

Then $q(v) = \lambda \in \mathbb{Z}$, $q(w) \in \mathbb{Z}$ because $w \in \Lambda$, and $B(nv, w) \in \mathbb{Z}$ because $nv, w \in \Lambda$.

Then for (4.8), we expand:

$$(q(v - w) \cdot n)v' = (q(v - w) \cdot n)(v - \frac{B_q(v, v - w)}{q(v - w)} \cdot (v - w)) =$$
$$\big(q(v - w) - B_q(v, v - w)\big)nv + B_q(nv, v - w)w.$$

Because nv and w lie in Λ, it suffices to see that each of our coefficients lie in $\mathbb{Z}$.

For the first coefficient, we find:

$$q(v-w)+q(v)=q(w-v)+q(v)=q(w)-B_q(w-v,v)=q(w)+B_q(v-w,v)$$

so that $q(v-w)-B_q(v-w,v)=q(w)-q(v)=q(w)-\lambda\in\mathbb{Z}$.

For the second coefficient, we see:

$$B_q(nv,v-w)=nB_q(v,v)-B_q(nv,w)$$

and the former term is $n\cdot 2q(v)\in\mathbb{Z}$, and the latter term is integral because $nv,w\in\Lambda$.

□

References

[Cas] J.W.S. Cassels. *Rational Quadratic Forms*. Dover Books on Mathematics Series. Dover Publications, Incorporated, 2008.

[Ser] Jean-Pierre Serre. *A Course in Arithmetic*. Graduate Texts in Mathematics. Springer, 1973.