SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Hasse-Minkowski Theorem
A local-to-global principle

av

Simon Edvard Natanael Vestberg

2025 - No M3

Hasse-Minkowski Theorem
A local-to-global principle

Simon Edvard Natanael Vestberg

Självständigt arbete i matematik 30 högskolepoäng, avancerad nivå

Handledare: Dan Petersen

2025

Hasse-Minkowski theorem
A Local to global principle

SIMON VESTBERG

Abstract. (English) In this paper, our aim is to describe a construction for the $p-$adic numbers and the Hasse-Minkowski theorem, where these are central. To this end, we will also touch upon some basic definitions and theorems regarding the following: quadratic forms, the Hilbert symbol and the Legendre symbol.

Key-words: $p-$adic numbers, Hasse-Minkowski theorem, local-to-global principle.

Abstract. (Swedish) Vårt mål i denna avhandling är att beskriva uppbyggnaden av de $p-$adiska talen och Hasse-Minkowskis sats, där de förstnämnda är en central del. För att uppnå detta berör vi även definitioner och satser angående: kvadratiska former, Hilbert symbolen och Legendre symbolen.

Nyckel-ord: $p-$adiska tal, Hasse-Minkowski, lokal-till-global principen.

Date: May 2025.

Contents

1. Introduction

The purpose of this thesis is not only to explore the construction of the $p-$adic numbers and to give a basic understanding of what these numbers are, but also to explain the Hasse-Minkowski theorem. Due to this, even though we will give a full construction of the $p-$adic numbers, the main goal is to state, prove and understand the so-called "local to global principle", which is also known as the Hasse-Minkowski theorem, which basically says that for certain equations (quadratic forms) to have rational solutions, it is necessary and sufficient that the equations have solutions in the completions of the rational numbers (id est in the reals and the $p-$adic numbers).

To be able to understand the proof of the Hasse-Minkowski theorem, we will not only need the understanding of the $p-$adic numbers but also understand exactly what these "certain equations" are to know when the theorem is applicable. To this end, we will also go through the basic definitions and theorems regarding these equations; quadratic forms.

The $p-$adic numbers were first introduced by a German mathematician named Kurt Hensel, who tried to bring concepts from mathematic series into number theory. Number theory is an ancient part of mathematics focused on numbers, and in particular integers, and the relations and properties between them. This in itself can be interesting enough; however, we can also view numbers as solutions to equations, and in this sense we can say that number theory aims to solve equations. This can be done in many different ways, one such way is studying the equation arithmetically modular some prime, but a more sophisticated way can be to search for solutions using the $p-$adic numbers. Indeed, as we mentioned earlier the Hasse-Minkowski theorem revolves around this. It is clear that if there is a rational solution to a certain equation it will yield solutions in the reals and the $p-$adic numbers (since the rational numbers embed into both the reals and the $p-$adic numbers). The Hasse-Minkowski theorem handles the converse; when is it possible for an equation with coefficients in the reals and the $p-$adic numbers to have a rational solution?

This way, you can use techniques that require complete fields (as both the reals and the $p-$adic numbers are complete fields) to search for a solution, for example Newton's method or Hensel's lemma (which is basically the $p-$adic analogue to Newton's method and can be applied to find roots of polynomials).

So, this might give you a clue to why the $p-$adic numbers are interesting, but the next natural question then is; what are they? This question will hopefully be answered in depth in the first part of this paper when we introduce the p-adic numbers and describe a construction of them. For now you will have to settle for this short informal explanation.

The $p-$adic numbers are the result of defining a different metric (id est a different distance function) on the rational numbers, in the same (or at least similar) way as the "normal" metric, the Euclidean absolute value, can be used to define the real numbers. Although the real numbers may at first seem more intuitive, and indeed they are, since we are taught these at an early age, the $p-$adic numbers are a great addition to a mathematicians toolbox. Further, even if they at first were introduced for a number theoretical purpose they also have applications in analysis, algebra and more (though this is nothing that will be discussed here in any more detail).

The major part of reference material consists of A Course in Arithmetic by Serre [Ser73] and $p-$adic numbers, $p-$adic analysis and Zeta-functions by Koblitz [Kob77].

2. The p-adic numbers

2.1. **Prerequisites.** Here we will go through some prerequisites needed for the construction of the $p-$adic numbers.

Definition 2.1.1. *Let X be a (non-empty) set, a function d from X to $\mathbb{R}_{\geq 0}$ is called a* `distance` *(or a* `metric`*), if the following criteria is met:*

$$(i)\ d(x,y) = 0 \text{ if and only if } x = y,$$
$$(ii)\ d(x,y) = d(y,x) \qquad \text{for all } x, y \in X,$$
$$(iii)\ d(x,y) \leq d(x,z) + d(z,y) \quad \text{for all } x, y, z \in X.$$

Moreover, we call the pair (X,d), the set together with the metric, a metric space.

Definition 2.1.2. *Let (X,d) be a metric space. We say that a sequence $\{x_1, x_2, ...\}$ with $x_i \in X$ for all i, is a* `Cauchy sequence` *(alternatively "sequence is Cauchy") if there exist, for each given $\epsilon > 0$, a $N \in \mathbb{Z}_{>0}$ such that $d(x_m, x_n) < \epsilon$ for all $m, n > N$.*

With this definition in mind we have this following definition regarding equivalence of two metrics.

Definition 2.1.3. *Two metrics d and d' are called* `equivalent` *if a sequence is Cauchy with regards to d then it is Cauchy with regards to d' and vice versa.*

In this paper we will mostly have X equal to $\mathbb{Q}$, the rational numbers, or $\mathbb{Q}_p$, the p-adic numbers (we will shortly define exactly what these are). Both of these are examples of fields, id est a set F together with two operations, call them addition and multiplication, such that F (respectively $F - \{0\}$) is a commutative group under addition (respectively multiplication) and the law of distributivity holds.

Definition 2.1.4. *Metrics d that come from the* `norm` *of a field F are maps, denoted as $||\ ||$, from F to $\mathbb{R}_{\geq 0}$ such that the following criteria hold:*

$$(i)\ ||x||=0 \text{ if and only if } x=0,$$
$$(ii)\ ||x\cdot y|| = ||x||\cdot||y||,$$
$$(iii)\ ||x+y|| \leq ||x||+||y||.$$

Note that when we say that a metric d comes from, alternatively induced by, a norm $||\ ||$ we simply mean that the metric is defined as $d(x,y) = ||x-y||$. If two norms induce equivalent metrics we say that the norms are equivalent. One example of an induced metric is the metric $d(x,y) = |x-y|$, where $|\ |$ is the standard absolute value norm in $\mathbb{Q}$. The metric $d(x,y)$ is then the "usual" distance between two numbers.

Definition 2.1.5. *Let p be a prime number and let $\text{ord}_p\ a$ be the greatest k such that $a \equiv 0\ (mod\ p^k)$, $a \in \mathbb{Z} - \{0\}$. (Recall that $a \equiv b\ (mod\ p)$ means that $p|(a-b)$), i.e. k is the highest power for p such that p^k divides a. Note that $\text{ord}_p\ (a_1 \cdot a_2) = \text{ord}_p\ a_1 + \text{ord}_p\ a_2$.*

For $a = 0$ we defined $\text{ord}_p\ a = \infty$, we also define for $x \in \mathbb{Q}$ (that is $x = \frac{a}{b}$) $\text{ord}_p\ x = \text{ord}_p\ a - \text{ord}_p\ b$. It is worth noting here that this definition does not depend on a or b, but only x, we could multiply a and b by some number c (so $x = \frac{ac}{bc}$) and still have $\text{ord}_p\ x = \text{ord}_p\ ac - \text{ord}_p\ bc = \text{ord}_p\ a + \text{ord}_p\ c - (\text{ord}_p\ b + \text{ord}_p\ c) = \text{ord}_p\ a - \text{ord}_p\ b$.

For example;

$$ord_7\ 98 = ord_7(49\cdot 2) = ord_7(7^2\cdot 2) = 2 \qquad ord_7\ 99 = 0.$$

Definition 2.1.6. *We define the map $|\ |_p$ from $\mathbb{Q}$ to $\mathbb{R}_{\geq 0}$ as:*

$$|x|_p = \begin{cases} 0, & \text{if } x = 0; \\ \frac{1}{p^{\text{ord}_p\ x}}, & \text{if } x \neq 0. \end{cases}$$

Proposition 2.1.7. *The map $|\ |_p$ defined in Definition 2.1.6 is a norm on $\mathbb{Q}$.*

Proof. We need to show that the criteria $(i) - (iii)$ for norms (Definition 2.1.4) hold for $|\ |_p$.

For (i) : if $x = 0$ we have, per definition, that $|x|_p = 0$. On the other hand, if $|x|_p = 0$ we have, since $\frac{1}{p^{\text{ord}_p\ x}} \neq 0$, that $x = 0$.

For (ii) : assume $x, y \neq 0$ (since if $x = 0$ or $y = 0$ the criterion clearly holds) then

$$|x\cdot y|_p = \frac{1}{p^{\text{ord}_p\ xy}} = \frac{1}{p^{\text{ord}_p\ x + \text{ord}_p\ y}} = \frac{1}{p^{\text{ord}_p\ x}} \cdot \frac{1}{p^{\text{ord}_p\ y}} = |x|_p \cdot |y|_p.$$

For (iii) : again assume that x, y and $x+y$ are all non-zero (if they are, (iii) holds trivially) and write $x = \frac{a}{b}$ and $y = \frac{c}{d}$ in their lowest terms. Then we have $x + y = \frac{ad+cb}{bd}$ and $\text{ord}_p\ (x+y) = \text{ord}_p\ (ad+cb) - \text{ord}_p\ bd$.

Furthermore, we have $\text{ord}_p\ (x+y) \geq \min\{\text{ord}_p\ ad, \text{ord}_p\ cb\} - \text{ord}_p\ bd$, since the greatest power of p that divides a sum of two numbers will be no less than the minimum of the greatest power of p dividing each of the numbers respectively. Now we can rewrite the right hand side of this, using properties from Definition 2.1.5 and of $\min\{a, b\}$ to get:

$$\begin{aligned}
&\min\{\text{ord}_p\ ad, \text{ord}_p\ cb\} - \text{ord}_p\ bd = \\
&\min\{\text{ord}_p\ a + \text{ord}_p\ d, \text{ord}_p\ c + \text{ord}_p\ b\} - \text{ord}_p\ b - \text{ord}_p\ d = \\
&\min\{\text{ord}_p\ a - \text{ord}_p\ b, \text{ord}_p\ c - \text{ord}_p\ d\} = \min\{\text{ord}_p\ \frac{a}{b}, \text{ord}_p\ \frac{c}{d}\} = \\
&\min\{\text{ord}_p\ x, \text{ord}_p\ y\}.
\end{aligned}$$

From this we can conclude that the third criterion holds because this shows that:

$$|x+y|_p = \frac{1}{p^{\text{ord}_p\ (x+y)}} \leq \max\{\frac{1}{p^{\text{ord}_p\ (x)}}, \frac{1}{p^{\text{ord}_p\ (y)}}\} = \max\{|x|_p, |y|_p\},$$

which clearly is $\leq |x|_p + |y|_p$ and we are done. □

One thing that is worth noting is that in this proof we actually proved a stronger inequality than what is needed for the map in Definition 2.1.6 to be called a norm ($|x+y|_p \leq \max\{|x|_p, |y|_p\}$ instead of the "normal" $||x+y|| \leq ||x|| + ||y||$). This leads us to an important distinction when it comes to norms, which in turn leads to the next definition.

Definition 2.1.8. *We say that a norm, respectively a metric, is* ***non-Archimedean*** *if the stronger inequality always holds, id est* $||x+y|| \leq \max\{||x||, ||y||\}$*, respectively* $d(x,y) \leq \max\{d(x,z), d(z,y)\}$*. Furthermore, we call norms (or metrics) that are not non-Archimedean simply* ***Archimedean****.*

We can see that if a metric is induced by a non-Archimedean norm the metric will also be non-Archimedean, since if this is the case we have $d(x,y) = ||x-y|| = ||(x-z)+(z-y)|| \leq \max\{||(x-z)||, ||(z-y)||\} = \max\{d(x,z), d(z,y)\}$.

The difference between a non-Archimedean and an Archimedean norm (or at least *a* big difference), id est property (iii) of Definition 2.1.4, leads to the conclusion that non-Archimedean norms have a somewhat weird property. For an Archimedean norm property (iii) is the "normal" triangle-inequality, which means that, in let say $\mathbb{R}^2$ with the standard Euclidean metric $d(x,y) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2}$, the sum of two sides of a triangle is greater than the third.

While if we do this in a non-Archimedean norm on some field F. Suppose, for simplicity´s sake, $z = 0$, then the triangle-inequality (property (iii) of Definition 2.1.4) for non-Archimedean norms state: $||x-y|| \leq \max\{||x||, ||y||\}$. Now, let

us assume that the sides x and y in this "triangle" have non-identical "length", id est suppose $||x|| \neq ||y||$ and lets suppose $||x|| < ||y||$. Then we have

$$||x - y|| \leq ||y||.$$

However, $||y|| = ||x - (x - y)||$ and so

$$||y|| = ||x - (x - y)|| \leq \max\{||x||, ||x - y||\} = ||x - y||,$$

where the last equality holds since $||y|| \not\leq ||x||$.

Then we have $||x - y|| \leq ||y|| \leq ||x - y||$ so $||y|| = ||x - y||$, but this means that if the two "sides" x and y are not of the same length then the bigger of the two will be equal to the third side. In other words, in a non-Archimedean norm, every triangle is isosceles. We will therefore call the triangle-inequality for non-Archimedean norms the "isosceles triangle-inequality".

As we saw in the proof of the proposition above, we have that $|\ |_p$ is a non-Archimedean norm on $\mathbb{Q}$.

Theorem 2.1.9. ***(Ostrowski's theorem)*** *Every non-trivial norm $||\ ||$ on $\mathbb{Q}$ is either equivalent to $|\ |_p$ for some prime p or to $|\ |_\infty$.*

Before we prove this, we make some notational observations; when we talk about the trivial norm, we mean the norm such that $||0|| = 0$ and $||x|| = 1$ for all $x \neq 0$, and by the notation $|\ |_\infty$ we simply mean the normal absolute value norm. Note that this is strictly a notation and you should not infer any connection between $|\ |_\infty$ and $|\ |_p$.

Now we move on to the proof of the theorem.

Proof. We separate the proof into two cases.

Case (1); Assume there exists some $n \in \mathbb{N}_{>0}$ such that $||n|| > 1$, and let n_0 denote the smallest such n. Then $||n_0|| = n_0^\alpha$ for some $\alpha \in \mathbb{R}_{>0}$ (since $||n_0|| > 1$ we can find such an α). Then we can write any $n \in \mathbb{N}_{>0}$ to the base n_0, that is, we can write

$$n = a_0 + a_1 n_0 + a_2 n_0^2 + ... + a_k n_0^k \quad \text{with } 0 \leq a_i < n_0 \text{ and } a_k \neq 0.$$

Now, by property (iii) of Definition 2.1.4, we have that

$$\begin{aligned} ||n|| &\leq ||a_0|| + ||a_1 n_0|| + ||a_2 n_0^2|| + ... + ||a_k n_0^k|| \\ &= ||a_0|| + ||a_1|| n_0^\alpha + ||a_2|| n_0^{2\alpha} + ... + ||a_k|| n_0^{k\alpha}. \end{aligned}$$

We can, since all $a_i < n_0$ and the way we picked n_0 makes it so $||a_i|| \leq 1$, rewrite this further as:

$$\begin{aligned} ||n|| &\leq 1 + n_0^\alpha + n_0^{2\alpha} + ... + n_0^{k\alpha} \\ &= n_0^{k\alpha}(n_0^{-k\alpha} + n_0^{-(k-1)\alpha} + ... + n_0^{-\alpha} + 1) \end{aligned}$$

and since $n_0^k \leq n$ we have that

$$n_0^{k\alpha}(n_0^{-k\alpha} + n_0^{-(k-1)\alpha} + ... + n_0^{-\alpha} + 1) \leq n^\alpha \left(\sum_{i=0}^{\infty} (\frac{1}{n_0^\alpha})^i \right).$$

We can easily see, by for example the root test, that the sum in the parenthesis converges, since $n_0^\alpha = ||n_0|| > 1 \implies 1/n_0^\alpha < 1$. So we can view this sum as some finite constant, that we call C, then we have

$$||n|| \leq Cn^\alpha \qquad \forall n \in \mathbb{N}_{>0}.$$

Now if we fix an n and take some (large) m and replace n with n^m and take the m^{th}-root we get $||n|| \leq \sqrt[m]{C}n^\alpha$, then, if we let $m \to \infty$, we get (for fixed n) $||n|| \leq n^\alpha$.

To get the inequality the other direction, id est $||n|| \geq n^\alpha$, we first write n to the base n_0 again and note that $n_0^{k+1} > n \geq n_0^k$. Now write $||n_0^{k+1}||$ as $||n_0^{k+1} + n - n||$ which, by property (iii) of Definition 2.1.4, is $\leq ||n|| + ||n_0^{k+1} - n||$ and so, since $||n^{k+1}|| = ||n||^{k+1}$ by property (ii) (of the same definition) and $||n|| \leq n^\alpha$ by above, we have

$$\begin{aligned} ||n|| &\geq ||n_0^{k+1}|| - ||n_0^{k+1} - n|| \\ &\geq n_0^{(k+1)\alpha} - (n_0^{k+1} - n)^\alpha. \end{aligned}$$

Further, since $n \geq n_0^k$, we can write this as

$$\begin{aligned} ||n|| &\geq n_0^{(k+1)\alpha} - (n_0^{k+1} - n_0^k)^\alpha \\ &= n_0^{(k+1)\alpha} \left(1 - (1 - \frac{1}{n_0})^\alpha \right) \end{aligned}$$

Again, we can view the parenthesis as a constant, call it C', that depends on n_0 and α (but importantly not on n). Also recall that $n_0^{k+1} > n$, so what we have is

$$||n|| \geq C'n^\alpha$$

and now we can do as before, replacing n with n^m where m is large and n is fixed and take the m^{th}-root. Then similarly to before, letting $m \to \infty$, we get $||n|| \geq n^\alpha$ and we can conclude that $||n|| = n^\alpha$.

Now, by property (ii) of Definition 2.1.4, we have that $||x|| = |x|^\alpha$ for all $x \in \mathbb{Q}$, since

$$||x|| = ||\frac{n}{m}|| = \frac{||n||}{||m||} = \frac{n^\alpha}{m^\alpha} = |x|^\alpha.$$

We can see that this norm is equivalent to the (standard) absolute value norm $|\ |_\infty$. This is because, if the sequence x_i is Cauchy with respect to $|\ |_\infty$ then for any given $\epsilon > 0$ we can find $N \in \mathbb{N}$ such that $|x_n - x_m|^\alpha < \epsilon$ for all $n, m > N$, indeed, just choose N large enough so that $|x_n - x_m|_\infty < \epsilon^{\frac{1}{\alpha}}$ (which we know is possible since the sequence is Cauchy with respect to this norm). Hence $|\ |^\alpha$ is equivalent to $|\ |_\infty$ in this case.

Case (2): Now, instead, assume that $||n|| \leq 1$ for all $n \in \mathbb{N}_{>0}$. Since we assume $||\ ||$ to be non-trivial we know there exists some n such that $||n|| < 1$, let n_0 denote the smallest such n. Then n_0 must be a prime, otherwise we would have $n_0 = n_1 \cdot n_2$ with $n_1, n_2 < n_0$ but by our choice of n_0 we must have $||n_1|| = ||n_2|| = 1$ and so $||n_0|| = ||n_1 n_2|| = 1$ contradicting $||n_0|| < 1$. Now if q is another prime (id est $q \neq n_0$), we claim that $||q|| = 1$. To prove this claim, assume to the contrary that $||q|| < 1$, then there exists some (possible very large) N such that $||q^N|| = ||q||^N < \frac{1}{2}$. Similarly, since $||n_0|| < 1$, we can find some large N' such that $||n_0^{N'}|| < \frac{1}{2}$. Now, since q^N and $n_0^{N'}$ are relatively prime, we can find $a, b \in \mathbb{N}_{>0}$ such that $aq^N + bn_0^{N'} = 1$. We then have, by property (ii) and (iii) of Definition 2.1.4,

$$||1|| = ||aq^N + bn_0^{N'}|| \leq ||aq^N|| + ||bn_0^{N'}|| = ||a|| \cdot ||q^N|| + ||b|| \cdot ||n_0^{N'}||.$$

However, we have $||a|| \leq 1$ and $||b|| \leq 1$, which gives us

$$1 \leq ||a|| \cdot ||q^N|| + ||b|| \cdot ||n_0^{N'}|| \leq ||q^N|| + ||n_0^{N'}|| < \frac{1}{2} + \frac{1}{2} = 1,$$

which is a contradiction, whence we conclude that $||q|| = 1$.

Since this is true for any $q \neq n_0$ and since any positive integer n can be factorized into primes as $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ we have that the only $||p_i||$ that is not equal to 1 in $||n|| = ||p_1||^{a_1} ||p_2||^{a_2} \cdot \cdots \cdot ||p_k||^{a_k}$ will be $p_i = n_0$ (if such a p_i exists, otherwise we simply have $||n|| = 1$). Further, the a_i corresponding to this p_i will be equal $\operatorname{ord}_p n$. Hence, we have, if we denote $\rho = ||n_0|| < 1$

$$||n|| = \rho^{\operatorname{ord}_p n},$$

and similarly as in Case (1) we can see, by property (ii) of Definition 2.1.4, that this holds for all rational $x \neq 0$ (not only for n). Also similarly as before, we can see that this kind of norm is equivalent to $|\ |_p$, which finishes the proof. □

2.2. **Construction.** From here to the end of this section, we will take p to be a prime not equal to ∞.

We seek to describe the formal construction of the p-adic numbers, denoted as $\mathbb{Q}_p$, as equivalence classes of Cauchy sequences. However, after the proof of Theorem 2.2.3 it is advisable to forgo this convoluted way of thinking and instead think about these numbers more concretely (and we will discuss how to do this later in this section) as infinite sums.

Let S be the set of Cauchy sequences $\{x_i\}$ of rational numbers. These sequences have the property (since they are Cauchy) that for any $\epsilon > 0$ there exists a (strictly) positive integer N such that $|x_i - x_{i'}|_p < \epsilon$ for any $i, i' > N$. Furthermore, we say that two of these (Cauchy) sequences, $\{x_i\}$ and $\{y_i\}$, are *equivalent* if $|x_i - y_i|_p \to 0$ as $i \to \infty$. If $\{x_i\}$ and $\{y_i\}$ are equivalent we denote

this as $\{x_i\} \sim \{y_i\}$. We then define the *p-adic numbers*, denoted as $\mathbb{Q}_p$, as the set of equivalence classes on S.

We denote, for $x \in \mathbb{Q}$, the "constant" Cauchy sequence, id est the sequence $\{x_i\} = \{x, x, x, x, ...\}$, simply as $\{x\}$. Note that $\{x\} \sim \{x'\} \Leftrightarrow x = x'$. Further, we denote the zero sequence $\{0, 0, 0, ...\} = \{0\}$ as 0.

We define the norm of an equivalence class, $|x|_p$, as $\lim_{i\to\infty}|x_i|_p$, where $\{x_i\}$ is a representative of the equivalence class x. We know that this limit will exist because if $x = 0$ then, per definition, $|x|_p = 0$ and so $\lim_{i\to\infty}|x_i|_p = 0$. On the other hand, if $x \neq 0$ then we can find an ϵ such that for all N there exists an $i_N > N$ such that $|x_{i_N}|_p > \epsilon$.

If we then take N sufficiently large such that $|x_i - x'_i|_p < \epsilon$ whenever $i, i' > N$ we get that

$$|x_i - x_{i_N}|_p < \epsilon \quad \forall\ i > N.$$

Now, by the isosceles triangle-inequality and since $|x_{i_N}|_p > \epsilon$, we have that $|x_i|_p = |x_{i_N}|_p$ which implies that, for all $i > N$, $|x_i|_p$ is constant and equal to $|x_{i_N}|_p$. Moreover, this will be the value for $\lim_{i\to\infty}|x_i|_p$.

Definition 2.2.1. *The* `multiplication` *of two equivalence classes (of Cauchy sequences) x and y, written as $x \cdot y$, we define by taking two representatives $\{x_i\} \in x$ and $\{y_i\} \in y$ and let the sequence $\{x_i y_i\}$ represent the equivalence class $x \cdot y$.*

Note that this definition of multiplication is independent of which representatives of x and y we choose. To see this, consider another representative $\{x'_i\} \in x$ and $\{y'_i\} \in y$ then we would have

$$\begin{aligned} |x'_i y'_i - x_i y_i|_p &= |x'_i(y'_i - y_i) + y_i(x'_i - x_i)|_p \\ &\leq \max\{|x'_i(y'_i - y_i)|_p,\ |y_i(x_i - x'_i)|_p\}. \end{aligned}$$

Now let $i \to \infty$, then the first argument in the max function will be equal to $|x'_i|_p \cdot \lim|y'_i - y_i|_p = 0$ since $\{y_i\} \sim \{y'_i\}$, similarly for the second argument in the max function we got $|y_i|_p \cdot \lim|x'_i - x_i|_p = 0$ since $\{x_i\} \sim \{x'_i\}$. Hence $|x'_i y'_i - x_i y_i|_p = 0$ as $i \to \infty$ and so, per definition, $\{x'_i y'_i\} \sim \{x_i y_i\}$.

Definition 2.2.2. *The* `sum` *of two equivalence classes of Cauchy sequences x and y we define in a similar manner, by taking a representative of each class $\{x_i\} \in x$ and $\{y_i\} \in y$ and define $x + y$ as term-wise addition, id est $\{x_i + y_i\}$.*

It can be shown, again similarly to above, that the choice of representative does not matter.

We also define the additive inverse $-x$ of x in the obvious way. Meaning, take a representative $\{x_i\} \in x$ and define $-x$ to be the equivalence class represented by $\{-1 \cdot x_i\}$ where -1 is the constant sequence $\{-1, -1, -1, ...\}$.

However, for multiplicative inverses we can not take the "obvious" sequence $\{\frac{1}{x_i}\}$, since there is a possibility that some terms in the Cauchy sequence $\{x_i\}$ are equal to zero. So we have to be a bit careful here, but we can see that

a zero term in a Cauchy sequence can be replaced with a non-zero term (say $x_i = 0$, then replace this term with $x_i' = p^i$). Then, as long as $\{x_i\} \not\sim 0$ (id est $|x_i|_p \not\to 0$ as $i \to \infty$), the sequence $\{\frac{1}{x_i}\}$ will be Cauchy and this sequence will be the multiplicative inverse of $\{x_i\}$.

We can now show that the set of equivalence classes of Cauchy sequences, $\mathbb{Q}_p$, together with the operations addition and multiplication as defined above, is a field. Example gratia; distributivity is easily shown as: take representatives $\{x_i\}$, $\{y_i\}$ and $\{z_i\}$ of the elements $x, y, z \in \mathbb{Q}_p$ (id est representatives of the equivalence classes of Cauchy sequences). Then the equivalence class $x(y+z)$ can be represented by the sequence $\{x_i(y_i + z_i)\} = \{x_i y_i + x_i z_i\}$, which also represents the equivalence class $xy + yz$ (so these equivalence classes of Cauchy sequences are the same) and so distributivity holds.

Theorem 2.2.3. *Let $x \in \mathbb{Q}_p$ be an equivalence class of Cauchy sequences such that $|x|_p \leq 1$. Then x has precisely one representative Cauchy sequence $\{x_i\}$ that satisfies the following:*

$$\begin{aligned} &(i)\ 0 \leq x_i < p^i \qquad \forall i \in \mathbb{N}_{>0}, \\ &(ii)\ x_i \equiv x_{i+1} \pmod{p^i} \qquad \forall i \in \mathbb{N}_{>0}. \end{aligned}$$

Before we go into the proof of this theorem we state a lemma, which will prove helpful in the proof of the theorem.

Lemma 2.2.4. *Suppose $x \in \mathbb{Q}$ with $|x|_p \leq 1$ then we can find an integer $\alpha \in \mathbb{Z}$ such that $|\alpha - x|_p \leq p^{-i}$ holds for any i.*

Furthermore, this α can be picked from the set $\{0, 1, 2, .., p^i - 1\}$.

Proof. Since $x \in \mathbb{Q}$ we can write $x = \frac{a}{b}$, let this be its simplest form, and since $|x|_p \leq 1$ we know, by definition, that p does not divide b. Thus p^i and b are relatively prime, which means that we can find integers n, m such that $np^i + mb = 1$. Now let $\alpha = am$ then

$$|\alpha - x|_p = |am - \frac{a}{b}|_p = |\frac{a}{b}|_p |mb - 1|_p$$

and now, since $|\frac{a}{b}|_p = |x|_p \leq 1$ and $mb - 1 = np^i$, we have

$$|\frac{a}{b}|_p |mb - 1|_p \leq |np^i|_p = \frac{|n|_p}{p^i},$$

which is $\leq p^i$ since $|n|_p \leq 1$ for all integers.

Lastly, note that we can (if needed) add any multiple of p^i to α to make $\alpha \in \{0, 1, 2, .., p^i - 1\}$ hold, and this will not change the fact that $|\alpha - x|_p \leq p^i$. □

Now to the proof of Theorem 2.2.3:

Proof. We will first show the uniqueness (that there is precisely one representative and no more). Let x_i' be another sequence that fulfills both criteria (i) and (ii). Then, if $x_{i_0} \neq x_{i_0}'$ we have that $x_{i_0} \not\equiv x_{i_0} \pmod{p_0^i}$, since by (i) we have

$0 \leq x_{i_0} < p_0^i$ and $0 \leq x'_{i_0} < p_0^i$. However, this means that for all $i \geq i_0$, that $x_i \not\equiv x'_i \pmod{p_0^i}$ since $x_i \equiv x_{i_0} \not\equiv x'_{i_0} \equiv x'_i \pmod{p_0^i}$. Which in turn means

$$|x_i - x'_i|_p > p^{-i_0} \quad \forall\, i \geq i_0$$

and so $\{x_i\} \not\sim \{x'_i\}$ (which is what we wanted to show).

Now we show the existence. Suppose $\{y_i\}$ is a Cauchy sequence; then what we want to do is to find a sequence $\{x_i\}$, for which (i) and (ii) hold, which is equivalent to $\{y_i\}$.

Let, for $k \in \mathbb{N}_{>0}$, n_k be a natural number such that $|y_i - y_{i'}|_p \leq p^{-k}$ for all $i, i' \geq n_k$. We can, without loss of generality, assume that the n_k's are increasing (strictly) with k and, in particular, that $n_k \geq k$. This means that $|y_i|_p \leq 1$ for $i \geq n_1$ since

$$|y_i|_p = |y_i - y_{i'} + y_{i'}|_p \leq \max\{|y_{i'}|_p, |y_i - y_{i'}|_p\} \leq \max\{|y_{i'}|_p,\ p^{-1}\}$$

for all $i' \geq n_1$ and $|y_{i'}| \to |x|_p \leq 1$ as i' tends to infinity. Then, by Lemma 2.2.4, we can find a sequence of integers α_k such that

$$|\alpha_k - y_{n_k}|_p \leq p^{-k}$$

and also, by the lemma, this sequence of integers will satisfy (i) in the theorem. We claim that $\{\alpha_k\}$ is the sequence we are looking for, to prove this we have left to show that $\alpha_{k+1} \equiv \alpha_k \pmod{p^k}$ and that $\{\alpha_k\} \sim \{y_i\}$.

We have (using the old trick that $y_{n_{k+1}} - y_{n_{k+1}} = 0$)

$$\begin{aligned}
|\alpha_{k+1} - \alpha_k|_p &= |\alpha_{k+1} - y_{n_{k+1}} + y_{n_{k+1}} - y_{n_k} - (\alpha_k + y_{n_k})|_p \\
&\leq \max\{|\alpha_{k+1} - y_{n_{k+1}}|_p, |y_{n_{k+1}} - y_{n_k}|_p, |\alpha_k + y_{n_k}|_p\} \\
&\leq \max\{\frac{1}{p^{k+1}}, \frac{1}{p^k}, \frac{1}{p^k}\} \\
&= \frac{1}{p^k},
\end{aligned}$$

whence we can conclude that the first of the two assertions we wanted to prove is correct. For the second one, $\{\alpha_k\} \sim \{y_i\}$, we use the exact same technique to see that, given any k, for $i > n_k$ we have

$$|\alpha_i - y_i|_p \leq \frac{1}{p^k}.$$

Thus we have $|\alpha_i - y_i| \to 0$ as $i \to \infty$ and we are done.

□

Now, with this theorem under our belts, we can "forget" that p-adic numbers are equivalence classes of Cauchy sequences and instead think of them as sums that stretch infinitely to the right.

However, before we explain this in more detail we have to ask; what about a p-adic number x for which $|x|_p \leq 1$ does not hold? In this case, we can multiply x with a power of p, say p^n, where this power equals $|x|_p$. Then we can find a p-adic number $x' = xp^n$ for which $|x'|_p \leq 1$ does hold, thus, by the Theorem

2.2.3, x' is represented by a sequence $\{x'_i\}$ (with properties as in the theorem) and $x = x'p^{-n}$ is represented by $\{x_i\}$ (where $x_i = x'_i p^{-n}$). It is then practical to view the x'_i's to the base p, that is as

$$x'_i = a_0 + a_1 p + a_2 p^2 + ... + a_{i-1} p^{i-1}$$

where a_j are integers in $\{1, 2, .., p-1\}$ for $j \in \{0, 1, .., i-1\}$. Note that the second condition in Theorem 2.2.3 means that

$$x'_{i+1} = a_0 + a_1 p + a_2 p^2 + ... + a_{i-1} p^{i-1} + a_i p^i,$$

where a_j for $j \in \{0, 1, .., i-1\}$ are the same as in the expansion of x'_i. So we can, intuitively, think of x' as a number, written in base p, that stretches infinitely to the right, where we add a new digit whenever we go from x_i to x_{i+1}.

We can now view our original x as the decimal number, written to the base p, where the decimals (id est the digits to the right of the decimal sign) are finite and are represented by the negative powers of p (id est written to the left in the sum), but still has infinitely many digits for positive powers. That is, x can be written as:

$$x = \frac{a_0}{p^n} + \frac{a_1}{p^n - 1} + ... + \frac{a_{n-1}}{p} + a_n + a_{n+1} p + a_{n+2} p^2 + ...,$$

we call this the "p-adic expansion" of x.

Worth to note here, is that the uniqueness of the sequence in the theorem only applies because $|\ |_p$ is non-Archimedean, since in the Archimedean case we can represent terminating decimals with repeating 9's, id est $1 = 0,9999...$, while if two p-adic expansions converge to the same number $x \in \mathbb{Q}_p$ they are the same. That is, all the digits a_j are the same.

3. Hensel's lemma and Chevalley-Warning theorem

Theorem 3.1. ***(Hensel's Lemma)***

Let $f \in \mathbb{Z}_p[x]$ be a polynomial of degree n with $p-$adic integer coefficients (id est $f(x) = a_0 + a_1 x + ... + a_n x^n$) and let f' be its derivative. Furthermore, let $b \in \mathbb{Z}_p$ be a $p-$adic integer such that $f(b) \equiv 0 \pmod p$ and $f'(b) \not\equiv 0 \pmod p$.

Then there exists a unique $b' \in \mathbb{Z}_p$ such that

$$f(b') = 0 \qquad \text{and} \quad b' \equiv b \pmod p$$

Proof. We claim that it is possible to find a sequence of integers $b_1, b_2, b_3... \in \mathbb{Z}$ for which the following is true for all $n \geq 1$:

$$(i)\ \ f(b_n) \equiv 0 \pmod{p^{n+1}}$$
$$(ii)\ \ b_n \equiv b_{n-1} \pmod{p^n}$$
$$(iii)\ \ 0 \leq b_n < p^{n+1}.$$

Hensel's lemma follows immediately from this claim (as we will see when we have proven the claim), so let us prove this claim.

We do this using induction on n: if $n = 1$ let $\bar{b}_0 \in \{0, 1, 2, ..., p-1\}$ be the unique integer such that $\bar{b}_0 \equiv b \pmod{p}$ then $b_1 = \bar{b}_0 + \alpha_1 p$, for some integer $0 \leq \alpha_1 \leq p-1$ (els (ii) and (iii) would not hold). Thus,

$$\begin{aligned} f(b_1) = f(\bar{b}_0 + \alpha_1 p) &= a_0 + a_1(\bar{b}_0 + \alpha_1 p) + ... + a_n(\bar{b}_0 + \alpha_1 p)^n \\ &= \sum_{i=0}^{n} a_i(\bar{b}_0 + \alpha_1 p)^i \\ &= \sum_{i=0}^{n} a_1 \bar{b}_0^i + i a_1 \bar{b}_0^{i-1} \alpha_1 p + ... \end{aligned}$$

Note that we may ignore all terms that contain powers of p greater than or equal to 2, since we are looking at f modulo p^2. Now, this last sum we can separate as

$$\sum_{i=0}^{n} a_1 \bar{b}_0^i + \sum_{i=0}^{n} i a_1 \bar{b}_0^{i-1} \alpha_1 p \pmod{p^2}$$

and this is equal to $f(\bar{b}_0) + f'(\bar{b}_0)\alpha_1 p$.

By assumption (in the theorem itself) we have that $f(b) \equiv 0 \pmod{p}$ and so $f(\bar{b}_0) \equiv \beta p \pmod{p^2}$ for some integer $0 \leq \beta \leq p-1$ which means that, for $f(a_1)$ to be congruent to 0 modulo p^2, we must have $\beta p + f'(\bar{b}_0)\alpha_1 p$ be congruent to 0 modulo p^2, which is the same as saying $\beta + f'(\bar{b}_0)\alpha_1 \equiv 0 \pmod{p}$. However, since, by the second assumption of the theorem, $f'(b) \not\equiv 0 \pmod{p}$, we can solve this for the unknown α_1 (since this means we can divide by $f'(b)$) and by using Lemma 2.2.4 we can pick $\alpha_1 \in \{0, 1, .., p-1\}$ so that $\alpha_1 \equiv -\frac{\beta}{f'(\bar{b}_0)} \pmod{p}$. Note that this also means that α_1 is uniquely determined.

Now we continue with the induction step, so assume that $b_1, b_2, ..., b_{n-1}$ is "found" and that we want to find b_n. As in the case for $n = 1$ we have, by (ii) and (iii), that $b_n = b_{n-1} + \alpha_n p^n$ for an integer $0 \leq \alpha_n \leq p-1$. Also similarly to before, we look at the expansion of $f(b_{n-1} + \alpha_n p^n)$ (ignoring terms divisible by p^{n+1}), then we have

$$f(b_n) = f(b_{n-1} + \alpha_n p^n) \equiv f(b_{n-1}) + f'(b_{b-1})\alpha_n p^n \pmod{p^{n+1}}.$$

Again we work as we did in the base case and rewrite the equality we want to prove, $f(b_n) \equiv 0 \pmod{p^{n+1}}$, as

$$\beta' p^n + f'(b_{n-1})\alpha_n p^n \equiv 0 \pmod{p^{n+1}} \iff \beta' + f'(b_{n-1})\alpha_n \equiv 0 \pmod{p}.$$

Note that we can do this, since by the induction hypothesis $f(b_{n-1}) \equiv 0 \pmod{p^n}$, which means that $f(b_{n-1}) \equiv \beta' p^n \pmod{p^{n+1}}$ for some $\beta' \in \{0, 1, 2, ..., p-1\}$.

Now we again use the assumption that $f'(b) \not\equiv 0 \pmod{p}$ together with the fact that $b_{n-1} \equiv b \pmod{p}$ to see that $f'(b_{n-1}) \equiv f'(b) \not\equiv 0 \pmod{p}$ and thus we can find α_n in the same way as we did in the base case (by solving

$\beta' + f'(b_{n-1})\alpha_n \equiv 0 \pmod{p}$ for α_n). This concludes the induction step and so also the proof of the claim.

Now, as we alluded to before, we easily prove the theorem (by using the claim). Simply let $b' = \bar{b}_0 + b_1 p + b_2 p^2 + ...$ then the $p-$adic number $f(b')$ must equal 0 since $f(b') \equiv f(b_n) \equiv 0 \pmod{p^{n+1}}$ for all n. On the other hand, if we have a b' of this form, then we have a sequence of b_n that fulfills the criteria in the claim and since this sequence is unique it follows that b' is unique.

This concludes the proof of Hensel's lemma.

□

In the following theorem and lemma, let p be a prime number and q be a power of p. Also, let F be a field with q elements.

Theorem 3.2. ***(Chevalleys-Warning)*** *Let $\{f_i\}_{i=1}^k \subseteq F[x_1, ...x_n]$ be polynomials such that $\sum_i \deg f_i < n$ (in other words, the total degree of all f_i's should be less than n).*

Then, the number of common solutions $(a_1, ..., a_n) \in F^n$ is congruent to 0 modulo p (id est the number of common solutions are divisible by the characteristic of the field F).

Note that the conclusion of the theorem could also be stated as "the cardinality of the vanishing set for the polynomials $\{f_i\}_{i=1}^k$ is congruent to 0 modulo p."

Lemma 3.3. *Let l be an non-negative integer.*

Then, the sum

$$\sum_{x \in F} x^l = \begin{cases} -1 & \text{if } l \geq 1 \text{ and divisible by } q-1, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. First note that we use the convention that $0^0 = 1$, so $x^l = 1$ if $l = 0$.

Now, if $l = 0$, we have that all terms in the sum is equal to 1. Thus, $\sum x^l = q \cdot 1 = 0$ since F is of characteristic p.

Secondly, if $l \geq 1$ and divisible by $q-1$. We have, by Fermat, $x^l = 1$ for $x \neq 0$ (and $0^l = 0$). So, in this case, we have that $\sum x^l = (q-1) \cdot 1 = q \cdot 1 - 1 = -1$.

Lastly, in the case where $l \geq 1$ is not divisible by $q-1$. We have, by basic group theory, that F^* is cyclic of order $q-1$ and by this we know that there exists some $y \in F^*$ for which $y^l \neq 1$ and that $\sum x^l = \sum y^l x^l \iff (1-y^l)\sum x^l = 0$. This implies that the sum must equal 0.

This finishes the proof of the lemma. □

Proof. (Of Theorem 3.2).

Let $x \in F^n$ and define the product $P = \prod_{i=1}^k (1 - f_i^{q-1}(x))$. Furthermore, let $U \subseteq F^n$ denote the set of common zeros of the f_i's.

Then we have, if $x \in U$, that $P = 1$ since, in this case, $f_i(x) = 0$ for all i. On the other hand, if $x \notin U$, at least one $f_i(x) \neq 0$ and so, since F is of characteristic p, $f_i(x)^{q-1} = 1$ thus $P = 0$ in this case.

Note that this makes P a so-called characteristic function of U (since P: $F^n \mapsto \{0,1\}$).

Now, define $S(g) = \sum_{x \in F^*} g(x)$, this means that $S(P)$ will equal the number of $x \in F^*$ such that $f_i(x) = 0$ for all i, or in other words:

$$Card(U) \equiv S(P) \pmod{p}.$$

If we can now show that $S(P) = 0$ we are done (since, as we note below the statement of the theorem, this is equivalent to the conclusion of the theorem).

The assumption that $\sum_{i=1}^{k} \deg f_i < n$ implies that the degree of P is less than $n(q-1)$. This in turn means that P is a linear combination of monomials $x^l = x_1^{l_1} \cdots x_n^{l_n}$ with combined degree less than $n(q-1)$, id est $\sum_{j=1}^{n} l_j < n(q-1)$. However, by Lemma 3.3, since at least one of the l_j is $< q-1$ (since $\sum_{j=1}^{n} l_j < n(q-1)$) we then know that $S(P) = S(x^l) = 0$ (since in this case we have a $l \geq 1$ not divisible by $q-1$) and we are done.

□

4. Quadratic forms

In this section, we will discuss quadratic forms. We will begin by stating multiple definitions, and then we will state some theorems which we will use later in the paper (in the proof of the Hasse-Minkowski Theorem).

Definition 4.1. *Let V be a module over a commutative ring R. We say that $Q : V \to R$ is a **quadratic form** on V if the following criteria hold:*

1. *$Q(rx) = r^2 Q$ for all $r \in R$ and $x \in V$*
2. *The function $(x, y) \mapsto Q(x+y) - (Q(x) + Q(y))$ is a bilinear form (id est a function that is linear in each argument separately).*

*Further, we call the pair (V, Q) a **quadratic module**.*

Note that in this paper we only consider the case where the ring A is a field (namely Q or Q_p, which we will denote as k) of characteristic other than 2. This makes it so that we can define the scalar product associated with Q as the symmetric bilinear form:

$$(x, y) \mapsto x.y = \frac{1}{2}\Big(Q(x+y) - (Q(x) + Q(y))\Big).$$

It is worthy to note that $x.x = Q(x)$ which determines a bijection between quadratic forms and symmetric bilinear forms (this is only the case since we are not considering characteristic 2).

We will later in this section show that each quadratic form Q is equivalent to a quadratic form looking like $a_1 x_1^2 + a_2 x_2^2 + ... + a_n x_n^2$.

For now, consider a basis $\{e_1, .., e_m\}$ for V and the matrix $A = (a_{i,j})$ with $(a_{i,j}) = e_i.e_j$ (this will be a symmetric), which is the matrix for Q with respect to this basis. Then for an element $x \in V$, we can write $x = \sum_{i=1}^{n} x_i e_i$, we have

$$Q(x) = \sum_{i,j} a_{i,j} x_i x_j,$$

which shows that Q is, in the variables $x_1, ..., x_n$, a "standard" quadratic form.

Definition 4.2. *Two quadratic forms Q and Q' are called **equivalent** if their modules (V, Q) respectively (V', Q') are isomorphic.*

We denote two equivalent forms as $Q \sim Q'$.

If Q and Q' are two equivalent quadratic forms then their corresponding matrices A and A' is related as $Y.A.Y^t = A'$ for some invertible matrix Y.

Definition 4.3. *Let (V, Q) be a quadratic module and $x, y \in V$ be two elements from V. We say that x and y are **orthogonal** if $x.y = 0$*

Further, we say that two vector subspaces $V_1, V_2 \subset V$ is orthogonal if for any $x \in V_1$ and $y \in V_2$ we have that $x.y = 0$.

Definition 4.4. *Let (V, Q) be a quadratic module of rank n. Then we denote by $d(Q)$ the **discriminant** of the quadratic form Q. If we have an orthogonal basis $e = \{e_1, e_2, ...e_m\}$ for V and we put $a_i = e_i \cdot e_i$ then*

$$d(Q) = a_1 \cdots a_n.$$

Furthermore, we define $\epsilon(Q) = \prod_{i<j}(a_i, a_j)$, where (a_i, a_j) is the Hilbert symbol (which, in particular, means that, if $a_i, a_j \in k^*$, then $(a_i, a_j) = \pm 1$). We have that $\epsilon(Q) = \pm 1$.

Definition 4.5. *Let (V, Q) be a quadratic module and d be the discriminant of the quadratic form Q. We say that Q is `non-degenerate` if $d \neq 0$.*

Definition 4.6. *Let (V, Q) be a quadratic module and $x \in (V, Q)$ be an element of this module. We say that x is `isotropic` if $Q(x) = 0$.*

Furthermore, if $U \subseteq V$ is such that, for all $x \in U$, $Q(x) = 0$ we say that U is `isotropic`.

Theorem 4.7. *For each quadratic module (V, Q) there exists an orthogonal basis, id est if $\{e_1, ..., e_m\}$ is a basis for (V, Q) the basis elements are pairwise orthogonal.*

Before we start with the proof of this theorem, it is worth noting that saying that the basis elements are pairwise orthogonal is the same as the matrix A of Q with respect to this basis being a diagonal matrix.

Proof. To prove this theorem, we will use induction on the number of basis elements (id est the dimension of V). For the base case, $m = 0$, it is trivially true. Now, if $m > 0$ pick an element $e_1 \in V$ such that $e_1.e_1 \neq 0$, if such an element exists (if it does not exist, we are already done since if $e.e = 0$ for

all elements $e \in V$ then all elements are isotropic and all bases of V would be orthogonal). Then, let H be the orthogonal complement of e_1 (clearly we then have $e_1 \notin H$) and we have $V = ke_1 \bigoplus H$, where k is whatever field V is a module over. However, this means that, since the dimension of H is less than m for which we can conclude, by induction, that H must have an orthogonal basis $\{e_2, ..., e_m\}$ and so $\{e_1, e_2, ..., e_m\}$ will be a orthogonal basis of V and we are done. □

Note that this theorem can be interpreted as: "any quadratic form is equivalent to a sum of squares". More concisely, if Q is a quadratic form then $Q \sim f = a_1x_2^2 + a_2x_2^2 + ... + a_mx_m^2$ with $a_1, ..., a_m \in k$ (from here on out, when we write f is a quadratic form, we mean a "standard" quadratic form, id est a sum of squares).

Definition 4.8. *Let f be a quadratic form (then, by Theorem 4.7, $f \sim a_1x_2^2 + a_2x_2^2 + ... + a_mx_m^2$). If the number of $a_j \neq 0$ equals i we say that f is of* ***rank*** *i.*

Now that we know that any quadratic form Q is equivalent to a quadratic form f we have some important and useful theorems to go through and consider.

Definition 4.9. *Let f be a quadratic form of n variables. We say that f* ***represent*** *an element $a \in k$ if there exists an $x = (x_1, ..., x_n) \in k^n$, not equal to 0, such that $f(x) = a$.*

Note that for a quadratic form f to represent 0, in light of this definition, this means that there must exist an isotropic element, different from 0, in the quadratic module.

Theorem 4.10. *Let f be a non-degenerate quadratic form such that f represent 0. Then f represent every element of k.*

A full proof for this can be found in exampli gratia [Ser73] on page 32, we however are more interested in the corollary.

Corollary 4.11. *Let f be a non-degenerate quadratic form of $n-1$ variables and let $a \in k^*$. Then, these statements are equivalent:*

(i) f represent a
(ii) f is equivalent to $f' + ay^2$, where f' is a quadratic form of $n-2$ variables.
(iii) Let $f'' = f - az^2$, then f'' represent 0.

Proof. The implication $(ii) \implies (i)$ is obvious, since, if $f \sim f' + ay^2$ we can simply take the element $(x_1, ...x_{n-2}, y) = (0, ..., 0, 1) \in k^{n-1}$ to make f represent a. On the other hand, if f represent a, we have $f(x) = x.x = a$ for some element $x \neq 0$ in the corresponding quadratic module V. Now, let H denote the orthogonal complement of x then, with the same argument as in Theorem 4.7 above, we have $V = kx \bigoplus H$. Further, if we let f' denote the quadratic form corresponding to a basis of H then we have, as desired, $f \sim f' + ay^2$ (and f' is a quadratic form of $n-2$ variables).

Now the implication $(ii) \implies (iii)$ is given, since $f \sim f' + ay^2$ represent a, $f'' = f - az^2 \sim f' + ay^2 - az^2$ will represent 0 (just take the element $(0, ..., 0, 1, 1) \in k^n$).

Lastly, we show the implication $(iii) \implies (i)$, assume $f'' = f - az^2$ represent 0, id est there exists a non-trivial element $(x_1, ..., x_{n-1}, z) \in k^n$. Then either $z = 0$ whence we conclude that f represent 0 and so, by Theorem 4.10, f represent all elements of k, and in particular f represent a. In the other case, when $z \neq 0$, we have an element $x = (\frac{x_1}{z}, ..., \frac{x_{n-1}}{z}) \in k^{n-1}$ such that $f(x) = a$ (id est f represent a). Either-way this show the desired implication and the proof is done.

□

Corollary 4.12. *Let f_1 and f_2 be two non-degenerate quadratic form of any rank greater than zero, and let $f = f_1 - f_2$. The following are equivalent:*

(i) *f represent 0.*
(ii) *There exists some element $a \in k^*$ such that both $f_!$ and f_2 represent a.*
(iii) *There exists some element $a \in k^*$ such that both $f_1 - ay^2$ and $f_2 - ay^2$ represent 0.*

Proof. From Corollary 4.11 it follows that $(ii) \iff (iii)$. Furthermore, $(ii) \implies (i)$ is immediate (since if f_1 and f_2 represent a we clearly have that $f = f_1 - f_2$ represent 0). For the implication $(i) \implies (ii)$: if f represent 0 then, per definition, there exists a non-trivial x such that $f(x) = 0$, write this x on form (x', x'') where $f_1(x') = f_2(x'')$. Now, either the element $a = f_1(x') = f_2(x'')$ does not equal 0 and we see that (ii) holds or $a = 0$ and in this case at least one of the forms, say f_1, represent 0. This means, by Theorem 4.10, f_1 represent all elements of k and, especially, all non-zero values that f_2 takes, so (ii) holds in this case as well. This ends the proof. □

Theorem 4.13. *Let f be a quadratic form of rank n. Then f represent 0 if and only if the following hold:*

(i) *$n = 2$ and $d(f) = -1$ (as an element of k^*/k^{*2}).*
(ii) *$n = 3$ and $\epsilon(f) = (-1, -d(f))$.*
(iii) *$n = 4$ and one of: $d(f) \neq 1$ or $d(f) = 1$ and $\epsilon(f) = (-1, -1)$.*
(iv) *$n \geq 5$.*

Note that when we say "$d(f) = -1$ as an element of k^*/k^{2*}" we mean that $d(f)$ equals the product of -1 by a square. The same also applies for $d(f)$ in the other cases of the theorem, and also in the upcoming corollary (which we will state before going into the proof of this theorem).

Corollary 4.14. *Let f be a quadratic form of rank n and let $a \in k^*/k^{*2}$. Then f represent a if and only if the following hold:*

(i) *$n = 1$ and $a = d(f)$ (as elements of k^*/k^{*2}, id est a equals a product of $d(f)$ by a square).*
(ii) *$n = 2$ and $\epsilon(f) = (a, -d(f))$.*

(iii) $n = 3$ *and one of:* $a \neq -d(f)$ *or* $a = -d(f)$ *and* $\epsilon(f) = (-1, -d(f))$.
(iv) $n \geq 4$.

We will only present the proof of case $n = 2$, since this is the only part used in this paper. However, the case $n = 3$ is quite easily proven using only the definition of Hilbert symbols (Definition 6.1) and some of its properties, while the last two cases require a bit more work. For a fully written proof we refer you to, for example, [Ser73] pages 37-38.

Proof. (Of Theorem 4.13 in the case $n = 2$). In this case, we have a quadratic form $f \sim a_1x_1^2 + a_2x_2^2$ and we can see that it is, for f to represent 0, necessary and sufficient that $-\frac{a_1}{a_2}$ is a square. Because, assume f represent 0 then

$$a_1x_1^2 + a_2x_2^2 = 0 \iff a_2x_2^2 = -a_1x_1^2 \iff x_2^2 = -\frac{a_1}{a_2}x_1^2 \iff \frac{x_2^2}{x_1^2} = -\frac{a_1}{a_2},$$

which implies that $-\frac{a_1}{a_2}$ is a square.

Note that the last equality is legal, since per definition of represent (Definition 4.9) we know that there exists a non-trivial $x \in k^2$ such that $f(x) = 0$ and if $x_1 = 0$ then $f(x_1, x_2) = a_2x_x^2 = 0$ which can only be true if $x_2 = 0$ (since by Definition 4.8 $a_2 \neq 0$).

Conversely, assume $-\frac{a_1}{a_2}$ is a square (and that f represent some element b) then

$$a_1x_1^2 + a_2x_2^2 = b \iff x_2^2 = \frac{b}{a_2} - \frac{a_1}{a_2}x_1^2.$$

Then we see that taking $x = (1, \sqrt{-\frac{a_1}{a_2}})$ gives us that f represent 0, as desired.

However, $-\frac{a_1}{a_2} = -a_1a_2 = -d(f)$ in k^*/k^{*2} which means that $-d(f)$ is a square, which in turn means $d(f) = -1$ (in k^*/k^{*2}).

This ends the proof for the case when f is of rank 2. □

5. Legendre symbol

In this section we are going to touch on the Legendre symbol and some of its basic properties. Recall that we say that a is a quadratic residue modulo p if there exists some integer x such that

$$x^2 \equiv a \pmod{p}.$$

Definition 5.1. *Let $p \neq 2$ be a prime number and x an integer. We define the* ***Legendre symbol*** *of x, denoted by $(\frac{x}{p})$, as*

$$\left(\frac{x}{p}\right) = \begin{cases} 1 & \text{if } a \not\equiv 0 \pmod{p} \text{ and } a \text{ is a quadratic residue modulo } p, \\ -1 & \text{if } a \text{ is not a quadratic residue modulo } p, \\ 0 & \text{if } a \equiv 0 \pmod{p}. \end{cases}$$

Alternatively, we could define the Legendre symbol of x, via an explicit formula, as the integer $x^{\frac{p-1}{2}} \pmod p \equiv \pm 1$.

Furthermore, if $x' \in \mathbb{F}_p$ is the image of $x \in \mathbb{Z}$ in the finite field $\mathbb{F}_p$, we simply write $(\frac{x}{p}) = (\frac{x'}{p})$.

First thing that is worthy to notice is that

$$\left(\frac{x}{p}\right)\left(\frac{y}{p}\right) = x^{\frac{p-1}{2}} y^{\frac{p-1}{2}} \pmod p = xy^{\frac{p-1}{2}} \pmod p = \left(\frac{xy}{p}\right).$$

In other words, it is multiplicative in the top argument, the Legendre symbol is in fact a character (however this is not something we will dive deeper into in this paper).

Now for some computations of the Legendre symbol, the following theorem will deal with the case when x equals $1, -1$ or 2. Before that though we want to define two functions that will clear up the notations in the following paragraphs a bit.

Definition 5.2. *Let a be an odd integer. We define the functions*

$$\varepsilon(a) \equiv \frac{a-1}{2} \pmod 2 = \begin{cases} 1 & a \equiv -1 \pmod 4 \\ 0 & a \equiv 1 \pmod 4 \end{cases}$$

and

$$\omega(a) \equiv \frac{a^2-1}{2} \pmod 2 = \begin{cases} 1 & a \equiv \pm 5 \pmod 8 \\ 0 & a \equiv \pm 1 \pmod 8. \end{cases}$$

Theorem 5.3. *Let p be a prime number, and q a power of p.*

We have that if $p = 2$ then every element of $\mathbb{F}_q$ is a square.

If $p \neq 2$ then the elements of $\mathbb{F}_q^*$ that are squares forms a subgroup H such that $|\mathbb{F}_q^* : H| = 2$, id est H has index 2.

Proof. For the first point consider the function $f : \mathbb{F}_q \to \mathbb{F}_q$ such that $f(a) = a^2$, recall that here we have q as a power of $p = 2$ meaning that $\mathbb{F}_q$ is of characteristic 2 and so, in this field, we have that $(a-b)^2 = a^2 - b^2$. Thus,

$$f(a) = f(b) \iff a^2 = b^2 \iff a^2 - b^2 = 0 \iff (a-b)^2 = 0 \iff a - b = 0.$$

Hence, f is injective and so (since $\mathbb{F}_q$ is a finite field) f is also surjective, id est f is an automorphism, and we conclude that in this case all elements of $\mathbb{F}_q$ are indeed squares.

For the second point, let $y \in \overline{\mathbb{F}_q}$ (id est let y be an element in the algebraic closure of $\mathbb{F}_q$) such that $y^2 = x$ where $x \in \mathbb{F}_q^*$.

Since $x \in \mathbb{F}_q^*$ we know that $x^{q-1} = 1$ and so $y^{q-1} = x^{\frac{q-1}{2}} = \pm 1$. Thus, x is a square in $\mathbb{F}_q$ if and only if $y^{q-1} = 1$ (id est $y \in \mathbb{F}_q^*$).

Moreover, this also means that the kernel of the function $x \mapsto x^{\frac{q-1}{2}}$ is precisely equal to $H = \mathbb{F}_q^{*2}$ and, since $\mathbb{F}_q^*$ is cyclic of order $q-1$, we have that $\mathbb{F}_q^{*2}$ has index 2 as stated. This completes the proof.

□

Theorem 5.4. *We have that for the Legendre symbol that the following is true:*

(i)

$$\left(\frac{1}{p}\right) = 1.$$

(ii)

$$\left(\frac{-1}{p}\right) = (-1)^{\varepsilon(p)}.$$

(iii)

$$\left(\frac{2}{p}\right) = (-1)^{\omega(p)}.$$

Proof. Since 1 is always a quadratic residue modulo p the (i) is clear. For (ii), if we use the alternative definition stated in definition 5.1, it is straightforward to see that $(\frac{-1}{p}) = (-1)^{\frac{p-1}{2}}$, but this is precisely as desired since $(-1)^k$ only depends on whether k is even or odd (id est we can view $(-1)^{\frac{p-1}{2}}$ as $(-1)^{\frac{p-1}{2} \pmod 2} = (-1)^{\varepsilon(p)}$).

For (iii) there is a bit more work needed. Let ζ denote the primitive 8^{th} root of unity in an algebraic closure $\overline{\mathbb{F}}_p$ of the finite field $\mathbb{F}_p$. Then we have that $\zeta^4 = -1$ which implies that $\zeta^2 + \zeta^{-2} = 0$, and so, for an element $y = \zeta + \zeta^{-1}$ we have $y^2 = \zeta^2 + \zeta^{-2} + 2 \cdot (\zeta\zeta^{-1}) = 2$ (id est y is the square root of the element 2 in $\mathbb{F}_p$). Which, by argument seen in the proof of Theorem 5.3, means that $(\frac{x}{p}) = y^{p-1}$.

Now, since we are working modulo p, we have

$$y^p = \zeta^p + \zeta^{-p}$$

and so, if $p \equiv \pm 1 \pmod 8$, we get $y^p = \zeta^p + \zeta^{-p} = y$ and

$$\left(\frac{2}{p}\right) = y^{p-1} = 1.$$

On the other hand, if $p \equiv \pm 5 \pmod 8$, we get (since $\zeta^4 = -1$) $y^p = \zeta^p + \zeta^{-p} = \zeta^5 + \zeta^{-5} = \zeta^4(\zeta^1 + \zeta^{-1}) = -(\zeta^1 + \zeta^{-1}) = -y$. Which implies, in the case where $p \equiv \pm 5 \pmod 8$, that $y^{p-1} = -1$.

We can thusly conclude that (iii) holds (since this was how we defined the function $\omega(a)$). This finishes the proof of the theorem. □

Lastly, we have the quadratic reciprocity law that state, if p and q are distinct odd primes, that $(\frac{p}{q})(\frac{q}{p}) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$. This can, alternatively, be written (using the $\varepsilon(n)$ from Definition 5.2 and the fact that the Legendre symbol is multiplicative in the top argument, since p and q are distinct we have $(\frac{p^2}{q}) = 1$ we can move one Legendre symbol to the right-hand side) as $(\frac{p}{q}) = (\frac{q}{p})(-1)^{\frac{p-1}{2}\frac{q-1}{2}}$ and this is how Gauss stated the quadratic reciprocity law.

Theorem 5.5. ***(Gauss)*** *Let p and q be two distinct odd primes then*

$$\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)(-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

To prove this theorem we will make use of two lemmas that we will state and prove shortly. First we make a note that we can use the Gauss sum

$$y = \sum_{x \in \mathbb{F}_q} \left(\frac{x}{q}\right) w^x.$$

We are able to do this, and it will be well-defined, since if we let $x \in \mathbb{F}_q$ and $w \in \overline{\mathbb{F}_p}$ such that w is an q^{th} root of unity (this means that w^x is well-defined since $w^q = 1$) the sum will indeed be well-defined.

Note that we will abuse the notation a bit and let q also denote the image of q in $\mathbb{F}_p$.

Lemma 5.6. $y^2 = (-1)^{\varepsilon(q)} q$

Proof. We prove this by manipulating the Gauss sum stated above. Recall that the Legendre symbol is multiplicative in the top argument we have:

$$y^2 = \sum_{x_1 \in \mathbb{F}_q} \left(\frac{x_1}{q}\right) w^{x_1} \cdot \sum_{x_2 \in \mathbb{F}_q} \left(\frac{x_2}{q}\right) w^{x_2} = \sum_{x_1, x_2 \in \mathbb{F}_q} \left(\frac{x_1 x_2}{q}\right) w^{x_1 + x_2}.$$

This can, with a simple variable change, be rewritten as

$$\sum_{x_1, x_2 \in \mathbb{F}_q} \left(\frac{x_1 x_2}{q}\right) w^{x_1+x_2}. = \sum_{z_1 \in \mathbb{F}_q} w^{z_1} \left[\sum_{z_2 \in \mathbb{F}_q} \left(\frac{z_2(z_1 - z_2)}{q}\right)\right].$$

Then, if $z_2 \neq 0$, we have (by the multiplicative nature of the Legendre symbol and the fact that $z_2(z_1 - z_2) = -1(1 - z_1 z_2^{-1}) z_2^2$) that

$$\left(\frac{z_2(z_1 - z_2)}{q}\right) = \left(\frac{(z_1 z_2^{-1} - 1) z_2^2}{q}\right) = \left(\frac{-1}{q}\right)\left(\frac{(1 - z_1 z_2^{-1})}{q}\right)\left(\frac{z_2^2}{q}\right).$$

This can be simplified; from Theorem 5.4 we have $(\frac{-1}{q}) = (-1)^{\varepsilon(q)}$ and from the definition we can see that $(\frac{z_2^2}{q}) = 1$ and so $(\frac{-1}{q})(\frac{(1-z_1 z_2^{-1})}{q})(\frac{z_2^2}{q}) = (-1)^{\varepsilon(q)}(\frac{(1-z_1 z_2^{-1})}{q})$.

Now, if we denote

$$C_{z_1} = \sum_{z_2 \in \mathbb{F}_q^*} \left(\frac{(1 - z_1 z_2^{-1})}{q}\right)$$

we have

$$y^2 = (-1)^{\varepsilon(q)} \sum_{z_1 \in \mathbb{F}_q} C_{z_1} w^{z_1}.$$

If $z_1 = 0$ then $C_0 = \sum_{z_2 \in \mathbb{F}_q^*} (\frac{1}{q}) = q - 1$ where the last equality comes from that there are $q - 1$ elements in $\mathbb{F}_q^*$ and, by Theorem 5.4, $(\frac{1}{q}) = 1$.

On the other hand, if $z_1 \neq 0$ we have that $z' = 1 - z_1 z_2^{-1}$ runs over $\mathbb{F}_q - \{1\}$ and so

$$C_{z_1} = \sum_{z' \in \mathbb{F}_q} \left(\frac{z'}{q}\right) - \left(\frac{1}{q}\right) = -\left(\frac{1}{q}\right) = -1.$$

Where the second to last equality comes from that we know, by Theorem 5.3, that the squares in $\mathbb{F}_q^*$ form a subgroup of index 2, id est the number of squares in $\mathbb{F}_q^*$ equal the number of non-squares and the last equality is, again, Theorem 5.4.

Putting all this together, we have that

$$(-1)^{\varepsilon(q)} y^2 = \sum_{z_1 \in \mathbb{F}_q} C_{z_1} w^{z_1} = q - 1 - \sum_{z_1 \in \mathbb{F}_q^*} w^{z_1} = q$$

and we are done. □

Lemma 5.7. $y^{p-1} = (\frac{p}{q})$

Proof. The proof of this lemma is quite straightforward; we again make use of the Gauss sum but also the fact that $\overline{\mathbb{F}_p}$ is of characteristic p.

We have that

$$y^p = \sum_{x \in \mathbb{F}_q} \left(\frac{x}{q}\right) w^{xp} = \sum_{z \in \mathbb{F}_q} \left(\frac{zp^{-1}}{q}\right) w^z = \left(\frac{p^{-1}}{q}\right) \sum_{x \in \mathbb{F}_q} \left(\frac{z}{q}\right) w^z = \left(\frac{p^{-1}}{q}\right) y.$$

Now, let b be the multiplicative inverse of a then $1 \equiv ab \pmod p$ and so $a \equiv a^2 b \pmod p$ which means that

$$1 \cdot \left(\frac{b}{p}\right) = \left(\frac{a^2}{p}\right)\left(\frac{b}{p}\right) = \left(\frac{a^2 b}{p}\right) = \left(\frac{a}{p}\right).$$

This means that

$$y^p = \left(\frac{p^{-1}}{q}\right) y = \left(\frac{p}{q}\right) y$$

and we can, dividing both side with y, see that

$$y^{p-1} = \left(\frac{p}{q}\right)$$

as we wanted to prove. □

The proof of Theorem 5.5 is now instantaneous by the above lemmas and by Theorem 5.4.

Proof. (of Theorem 5.5)

By Lemma 5.6 and 5.7 we have:

$$\left(\frac{(-1)^{\varepsilon(q)} q}{p}\right) = y^{p-1} = \left(\frac{p}{q}\right)$$

and by Theorem 5.4:

$$\left(\frac{(-1)^{\varepsilon(q)}}{p}\right) = (-1)^{\varepsilon(q)\varepsilon(p)}$$

(since if $\varepsilon(q) = 0$ then $(-1)^{\varepsilon(q)\varepsilon(p)} = 1$ no matter what $\varepsilon(p)$ is, as it should be. On the other hand, if $\varepsilon(q) = 1$ the above-mentioned theorem gives the formula).

Then, since the Legendre symbol is multiplicative in the top argument and by Definition 5.2, we get the reciprocity law as stated:

$$\left(\frac{q}{p}\right) = (-1)^{\frac{q-1}{2}\frac{p-1}{2}}\left(\frac{p}{q}\right).$$

□

6. Hilbert symbol

Throughout this section we will let k denote either $\mathbb{R}$ or $\mathbb{Q}_p$ (id est either the field of real numbers or the field of $p-$adic numbers for some p prime).

Definition 6.1. *Let $a, b \in k^*$. We call the number $(a, b) = \pm 1$ the **Hilbert symbol** of a and b and define it as*

- $(a, b) = 1$ *if $z^2 - ay^2 - bx^2 = 0$ has a solution other than $(0, 0, 0) \in k^3$*
- $(a, b) = -1$ *if such a solution does not exist.*

Note that we can multiply a and b with squares without changing the value of the Hilbert symbol (a, b). Hence, (a, b) is a map from $k^*/k^{*2} \times k^*/k^{*2}$ to $\{1, -1\}$.

We will denote, if clarification is needed, the Hilbert symbol as $(a, b)_p$, respectively $(a, b)_\infty$, with $a, b \in \mathbb{Q}_p$, respectively $a, b \in \mathbb{R}$ (id est if the form $z^2 - ay^2 - bx^2$ represent 0 or not, in $\mathbb{Q}_p$ respectively $\mathbb{R}$).

Proposition 6.2. *Let $a, b, c, d \in k^*$ and assume $a \neq 1$ in the Hilbert symbols containing $1 - a$. The following are properties of the Hilbert symbol:*

(i) $(a, b) = (b, a)$ *and* $(a, c^2) = 1$
(ii) $(a, -a) = 1$ *and* $(a, 1 - a) = 1$
(iii) $(a, b) = 1 \implies (ad, b) = (d, b)$
(iv) $(a, b) = (a, -ab) = (a, (1 - a)b)$

Proof. For (i) : if $(x', y', z') \in k^3$ is a solution to quadratic form $z^2 - ay^2 - bx^2 = 0$ then (y', x', z') is gonna be a solution to $z^2 - by^2 - ax^2 = 0$, thus $(a, b) = (b, a)$. Further, if b equals a square c^2 then clearly $z^2 - ay^2 - bx^2 = z^2 - ay^2 - c^2x^2 = 0$ has a solution in k^3, hence $(a, c^2) = 1$.

For (ii) : if $b = -a$ then the quadratic form $z^2 - ay^2 - bx^2 = z^2 - ay^2 + ax^2$ will have a zero at $(0, 1, 1)$ (id est $(a, -a) = 1$), and if $b = 1 - a$ then $z^2 - ay^2 - bx^2 = z^2 - ay^2 - (1 - a)x^2 = z^2 - ay^2 - x^2 + ax^2$ will have a zero at $(1, 1, 1)$, (which implies that $(a, 1 - a) = 1$).

For (iii) : if $(a,b) = 1$ then the quadratic form $z^2 - ay^2 - bx^2 = 0$ has a solution (not equal to $(0,0,0)$). Now, either b is a square of some element b' whence we can see that the form has a zero at $(1,0,b')$ and $k(\sqrt{b}) = k$ and the norm elements of $k(\sqrt{b})^*$ equals k^*, or b is not a square whence we know that the solution $(x',y',z') \neq (0,0,0)$ must be such that $y' \neq 0$ (else b would be a square). This means that a is the norm of $\frac{z}{y} + \beta\frac{x}{y}$, where β denotes a square root of b. In either case we have that a belongs to the group of norms of $k(\sqrt{b})^*$ (denoted $Nk(\sqrt{b})^*$) and we have that $d \in Nk(\sqrt{b})^* \Leftrightarrow ad \in Nk(\sqrt{b})^*$, this proves (iii).

For (iv) : this follows from $(i) - (iii)$.

□

It is worth to note that (iii), in the proposition above, is a special case of

$$(ad,b) = (a,b)(d,b),$$

which demonstrates the bilinearity of the Hilbert symbol. The actual proof that the Hilbert symbol is bilinear we get from the following theorem, since the formulae for the Hilbert symbols given in the theorem are bilinear.

Theorem 6.3. *For $k = \mathbb{R}$, if $a > 0$ or $b > 0$ we have that $(a,b) = 1$ and if both $a, b < 0$ we have $(a,b) = -1$.*

For $k = \mathbb{Q}_p$, let a, b be written in the form $p^\alpha u$ respectively $p^\beta v$, with u, v $p-$adic units, then we have

$$(a,b) = (-1)^{\alpha\beta\varepsilon(p)}\left(\frac{u}{p}\right)^\beta\left(\frac{v}{p}\right)^\alpha, \qquad p \neq 2,$$

$$(a,b) = (-1)^{\varepsilon(u)\varepsilon(v)+\alpha\omega(v)+\beta\omega(u)}, \qquad p = 2.$$

Theorem 6.4. ***(Hilbert)*** *Let $a, b \in \mathbb{Q}_p$. Then $(a,b)_v = 1$ for all $v \in V$ (possibly excluding some finite amount of elements) and*

$$\prod_{v\in V}(a,b)_v = 1.$$

Proof. First note that by the observation above, that the Hilbert symbol is bilinear, it is enough to prove the theorem in the cases where a, b equals -1 or some prime q. We make use of Theorem 6.3 in all of the following cases to calculate the Hilbert symbol.

Case $a = -1,\ b = -1$: we have $(a,b)_\infty = (a,b)_2 = -1$ and $(a,b)_v = 1$ for all $v \in V - \{2,\infty\}$, thus the product equals 1 as desired.

Case $a = -1,\ b = q$: here we have two sub-cases ($q = 2$ and $q \neq 2$). If $q = 2$ we have $(a,b)_v = (-1,2)_v = 1$ for all $v \in V$. If $q \neq 2$, we have $(a,b)_v = 1$ for all $v \in V - \{2,q\}$ and $(a,b)_2 = (a,b)_v = (-1)^{\varepsilon(q)}$. Thus, in both cases, the product equals 1.

Case $a = q,\ b = q'$: where q' is another prime (possibly distinct from q). If $q = q'$ we have for all $v \in V$, by Proposition 6.2 case (iv) together with (i), that $(a,b)_v = (q,q)_v = (-1,q)_v$ and we can refer back to the previous case.

If $q \neq q'$ and $q' = 2$ we have $(a, b) = (q, 2) = 1$ for all $v \in V - \{2, q\}$, for $v = 2, q$ we have $(q, 2)_2 = (-1)^{\omega(q)}$ and

$$(q, 2)_q = \left(\frac{2}{q}\right) = (-1)^{\omega(q)},$$

where this last equality comes from Theorem 5.4.

If $q \neq q'$ and $q, q' \neq 2$, we have $(a, b)_v = (q, q')_v = 1$ for all $v \in V - \{2, q, q'\}$, for $v = 2, q, q'$ we have $(a, b)_2 = (q, q')_2 = (-1)^{\varepsilon(q)\varepsilon(q')}$ and

$$(a, b)_q = (q, q')_q = \left(\frac{q'}{q}\right), \quad (a, b)_{q'} = (q, q')_{q'} = \left(\frac{q}{q'}\right).$$

However, by Theorem 5.5 (quadratic reciprocity), we have

$$\left(\frac{q}{q'}\right)\left(\frac{q'}{q}\right) = (-1)^{\varepsilon(q)\varepsilon(q')}$$

from which we can conclude that the product must be equal to 1, which finishes the proof.

□

This next theorem proves the existence of rational numbers given a Hilbert symbol, id est given elements $a_i \in \mathbb{Q}^*$ and a collection $(e_{i,v})_{i \in I, v \in V}$ (where I is a set of indices) of numbers ± 1 there exists a $x \in \mathbb{Q}$ such that $(a_i, x)_v = e_{i,v}$ $\forall i \in I, v \in V$.

Before we actually state and prove the theorem, we will first present some lemmas that will be needed for the proof. The first two of these lemmas will be stated without proof (proofs for these can be found exempli gratia in [Ser73] page 24 and 74-75 respectively).

Lemma 6.5. *(**Chinese remainder theorem**) Let $a_1, .., a_k$ and $l_1, ...l_k$ be integers such that $\gcd(l_i, l_j) = 1$ for all $i \neq j$, id est all the l_i are relatively prime, and $0 \leq a_i < l_i$ for all i. Then there exists an integer a such that*

$$a \equiv a_i \pmod{l_i} \quad \text{for all } i.$$

Lemma 6.6. *(**Dirichlet theorem**) Let $a, n \in \mathbb{Z}_{\geq 1}$ be relatively prime integers. Then there exists infinitely many primes p for which $p \equiv a \pmod{n}$.*

Lemma 6.7. *(**Approximation lemma**) Let S be a finite subset of V and let $\prod_{v \in S} \mathbb{Q}_v$ be equipped with the standard product topology. Then the image of $\mathbb{Q}$ is dense in $\prod_{v \in S} \mathbb{Q}_v$.*

Proof. Suppose $S = \{\infty, p_1, ..., p_n\}$, with p_i's being distinct primes, we then want to show that $\mathbb{Q}$ is dense in $\mathbb{R} \times \mathbb{Q}_{p_1} \times ... \times \mathbb{Q}_{p_n}$. We do this by showing that any point, $(x_\infty, x_1, x_2, ..., x_n)$, in this product is a closure point of $\mathbb{Q}$ (id est every open neighbourhood of this point contains at least one point of $\mathbb{Q}$).

So, let $(x_\infty, x_1, x_2, ..., x_n) \in \mathbb{R} \times \mathbb{Q}_{p_1} \times ... \times \mathbb{Q}_{p_n}$ be a point of the product. We can assume $x_i \in \mathbb{Z}_{p_i}$ for $1 \leq i \leq n$, if not we may simply multiply with

some integer to make it so. Then, given any $\epsilon > 0$ and any integer $N > 0$, we want to prove that there exists an $x \in \mathbb{Q}$ such that

$$|x - x_\infty| < \epsilon \quad \text{and} \quad \text{ord}_{p_i}(x - x_i) \geq N \quad \text{for all } i \in \{1, ..., n\}.$$

By the Chinese remainder theorem (applied to $l_i = p_i^N$) there exists some $y \in \mathbb{Z}$ such that $\text{ord}_{p_1}(y - x_i) \geq N$ for all i.

Now, pick some prime number $q \geq 2$ (actually any number that is relatively prime with all p_i's would work). Then, since $q^m \to \infty$ as $m \to \infty$, we have that the rational numbers of the form $\frac{a}{q^m}$, with $a \in \mathbb{Z}$ and $m \geq 0$, are dense in $\mathbb{R}$.

Now, pick a number b on the form above (id est $b = \frac{a}{q^m}$) such that

$$|y - x_\infty + bp_1^N \cdots p_n^N| \leq \epsilon.$$

Then $x = y + bp_1^N \cdots p_n^N$ (which will be a rational number) will be as desired and we are done. □

Now to the theorem I alluded to above which will make use of Dirichlet's theorem (Lemma 6.6) and the approximation lemma (Lemma 6.7).

Theorem 6.8. *Let $(a_i)_{i \in I}$ and $(\epsilon_{i,v})_{i \in I, v \in V}$ be collections of numbers from $\mathbb{Q}^*$ respective numbers equal to $\{\pm 1\}$ (here I is a finite index set). We have that there exists an $x \in \mathbb{Q}^*$ such that $(a_i, x)_v = \epsilon_{i,v}$ for all $i \in I$ and for all $v \in V$ if and only if the following conditions hold:*

(i) All, but a finite amount, $\epsilon_{i,v}$ equals 1.
(ii) $\prod_{v \in V} \epsilon_{i,v} = 1$ for all $i \in I$.
(iii) There exists, for all $v \in V$, $x_v \in \mathbb{Q}_v^*$ such that, for all $i \in I$, $(a_i, x_v)_v = \epsilon_{i,v}$.

Proof. The "if" part of (i) and (ii) follows from Theorem 6.4. Furthermore, for (iii) we can simply take $x_v = x$.

The other direction is more work; let $(\epsilon_{i,v})_{i \in I, v \in V}$ be a collection of numbers equal to ± 1 that satisfies the three conditions $(i), (ii)$ and (iii) in the theorem. We may also, since we are free to multiply a_i by the square of some integer, assume that the a_i's are integers. Now, let $S, T \subset V$ be two subsets of V where $v \in S$ if $v = \infty, 2$ or is a prime factor of any a_i and $v \in T$ if there exists $i \in I$ such that $\epsilon_{i,v} = -1$. We can first note that these sets will be finite (since the collections (a_i) and $(\epsilon_{i,v})$ are finite) and we may also note that we can, and will indeed, separate the proof into two cases: if $S \cap T = \emptyset$ or if the intersection is not empty (which we will call the general case).

Case $S \cap T = \emptyset$: Put

$$t = \prod_{l \in T - \{\infty\}} l \quad \text{and} \quad s = 8 \cdot \prod_{l \in S - \{2, \infty\}} l.$$

We have, since $S \cap T = \emptyset$, that these two integers, s and t, are relatively prime and both are ≥ 1 whence, by Lemma 6.6, we can find a prime number p such that $p \equiv t \pmod{s}$ and $p \notin S \cup T$. We then want to show that $x = tp$ will be as desired (id est, for all $i \in I$ and $v \in V$, $(a_i, x)_v = \epsilon_{i,v}$).

If $v \in S$ then, again since $S \cap T = \emptyset$, we have $\epsilon_{i,v} = 1$ and we are left to check if $(a_i, x) = 1$. We can first note that if $v = \infty$ then, since both $t > 0$ and $p > 0$ we have $x > 0$, and so $(a_i, x)_\infty = 1$ by Theorem 6.3. On the other hand, if $v \neq \infty$, we have that v is a prime number l and (since x and t are $l-$adic units, because $l \neq p$ by our choice of p) $x = tp \equiv t^2 \pmod{s}$ thus $x \equiv t^2 \pmod 8$ if $l = 2$ and $x \equiv t^2 \pmod l$ if $l \neq 2$, either way x is a square in $\mathbb{Q}_q^*$ and we can conclude that $(a_i, x)_v = 1$. Since, if we consider the equation $z^2 - cy^2 - bx^2 = 0$, and if b is square of some element b' then the equation will have a solution at $(1, 0, b')$ and so $(c, b) = 1$.

If $v \notin S$, then v is a prime number $l \neq 2$ (since 2 is in S) and we also know that l is not a prime factor of a_i (since all of these are also in S), this means that a_i is a $l-$adic unit. Then by Theorem 6.3 we have that, for all $b \in \mathbb{Q}_l^*$,

$$(a_i, b)_l = \left(\frac{a_i}{l}\right)^{\operatorname{ord}_l b}.$$

.

If $l \notin T \cup \{p\}$ then $x = tp$ is a $l-$adic unit, which means that $\operatorname{ord}_l x = 0$, which in turn means that the formula above shows that $(a_i, x)_l = 1$. However, this is exactly what we want because, since $l \notin T$, we have, by construction, that $\epsilon_{i,l} = 1$.

If $l \in T$ (note that this means that $l \neq p$ by our choice of p) then, since $x = tp$ and t is the product of the elements of T, we have $\operatorname{ord}_l x = 1$. Furthermore, by condition (iii) there exists some $x_l \in \mathbb{Q}_l^*$ such that $(a_i, x_l)_l = \epsilon_{i,l}$ for all $i \in I$ and we know, since $l \in T$, that at least one of these have $\epsilon_{i,l} = -1$. This means that $\operatorname{ord}_l x_l \equiv 1 \pmod 2$ hence we have, for all $i \in I$,

$$(a_i, x)_l = \left(\frac{a_i}{l}\right) = (a_i, x_l)_l = \epsilon_{i,l}$$

as desired.

Only thing left is if $l = p$, this we can infer from the previous cases together with the product formula (Theorem 6.4) and conclude that

$$(a_i, x)_p = \prod_{v \neq p} (a_i, x)_v = \prod_{v \neq p} \epsilon_{i,v} = \epsilon_{i,p}$$

and we are done.

Left to prove is the so called "general case" where $S \cap T \neq \emptyset$. Since the squares of $\mathbb{Q}_v^*$ forms an open set we have, by Lemma 6.7, that there exists a $x' \in \mathbb{Q}^*$ for which $\frac{x'}{x_v}$ is a square in $\mathbb{Q}_v^*$ for all $v \in S$. This means, specifically, that for all $v \in S$ we have $(a_i, x')_v = (a_i, x_v)_v = \epsilon_{i,v}$. Now, if we define $\varepsilon_{i,v} = \epsilon_{i,v} \cdot (a_i, x')_v$, then first note that if $v \in S$ we have $\varepsilon_{i,v} = 1$. Furthermore, the collection $(\varepsilon_{i,v})_{i \in I, v \in V}$ satisfies the conditions (i), (ii) and (iii), this means that we can use the previous case (where S and T are disjoint). So, by Case $S \cap T = \emptyset$, there exists a $y \in \mathbb{Q}^*$ such that $(a_i, y)_v = \varepsilon_{i,v}$ for all $i \in I$ and for all $v \in V$.

Now we can find an x with the desired properties, namely let $x = yx'$ and we are done.

□

7. Hasse-Minkowski theorem

Theorem 7.1. ***(Hasse-Minkowski theorem)*** *Let f be a quadratic form, written as*

$$f = a_1x_1^2 + a_2x_2^2 + ... + a_nx_n^2, \qquad a_i \in \mathbb{Q}^*.$$

Then for f to represent 0 it is necessary and sufficient, that f_v represent 0, for all $v \in V$. Where V is equal to the set of all primes union infinity.

This is also, sometimes, called the local to global principle, because in other words this means that for f to have a "global" zero it must have "local" zeroes everywhere, and vice versa. Writing it like this the necessary part is clear, since if f "misses" a local zero then it cannot have a global zero, thus we only need to show the sufficiency.

Proof. First note that we can, by exchanging f by a_1f, assume that $a_1 = 1$.

We will consider different cases of n and prove the theorem separately for $n = 2, 3, 4$ and $n \geq 5$.

1) Case $n = 2$: Here we have $f = x_1^2 - a_2x_2^2$, since f_v represent 0, for all $v \in V$, we have, in particular, that f_∞ represent 0 which means that $a_2 > 0$. Further, consider the prime decomposition of a_2

$$a_2 = \prod_p a_2^{\operatorname{ord}_p a_2}$$

then, since f_p represent 0 and so a_2 is a square in $\mathbb{Q}_p$ (because $f_p = x_1^2 - a_2x_2^2 = 0 \implies a_2 = \frac{x_1^2}{x_2^2} = (\frac{x_1}{x_2})^2$), we have that $\operatorname{ord}_p a_2$ is even. This means that a_2 is a square in $\mathbb{Q}$ and f represent 0.

2) Case $n = 3$: Here we have $f = x_1^2 - a_2x_2^2 - a_3x_3^2$ (for easier notation lets write $a_2 = a$ and $a_3 = b$). We can assume, since we are free to multiply a and b with any square, that a, b are square free integers (id est $\operatorname{ord}_p a$ and $\operatorname{ord}_p b$ equals 0 or 1 for all p primes). We may also, without loss of generality, assume that $|a| \leq |b|$. We now prove that f represent 0 if f_v represent 0, for all $v \in V$, by the use of induction on the integer $k = |a|+|b|$.

If $k = 2$; in this case we have $f = x_1^2 \pm x_2^2 \pm x_3^2$, we can eliminate the case where $f = x_1 + x_2 + x_3$ since f_∞ represent 0 (which is impossible in this case), in the other cases f represent 0. For example, if $f = x_1^2 + x_2^2 - x_3^2$ we can find rational numbers x_1, x_2 and x_3 such that $f = 0$.

So lets assume $k \geq 3$ (in other words, since a and b are integers, $|b| \geq 2$) and write b as a decomposition into, since b square free, distinct primes

$$b = \pm p_1 p_2 \cdots p_i.$$

Further, assume that one of the primes in this decomposition equals p, then we want to show that a is a square modulo p.

First note that this is clear if $a \equiv 0 \pmod{p}$. On the other hand, if $a \not\equiv 0 \pmod{p}$, a is a $p-$adic unit; by assumption (that f_p represent 0), there exists $x = (x_1, x_2, x_3) \in (\mathbb{Q}_p)^3$ such that $x_1^2 - ax_2^2 - bx_3^2 = 0$. Moreover, we can suppose that x is primitive, since if not we can simply divide $x_1^2 - ax_2^2 - bx_3^2 = 0$ by p as many times as needed for one of x_1, x_2 or x_3 to not be divisible by p anymore (note that this does not change a or b in any way). Now, since $b \equiv 0 \pmod{p}$, we have $x_1^2 - ax_2^2 \equiv 0 \pmod{p}$ and it follows that, if $x_2 \equiv 0 \pmod{p}$ then $x_1 \equiv 0 \pmod{p}$ and $p^2 \mid bx_3^2$ (since $bx_3^2 = x_1^2 - ax_2^2$), which means that $x_3 \equiv 0 \pmod{p}$ (since $\mathrm{ord}_p(b) = 1$), contradicting the assumption that x is primitive. Hence $x_2 \not\equiv 0 \pmod{p}$, and thus a is a square modulo p and so it is a square modulo b, since, because $b = \pm p_1 p_2 \cdots p_i$, we have that $\mathbb{Z}/b\mathbb{Z} = \prod_{l=1}^{i} \mathbb{Z}/p_l\mathbb{Z}$.

We can therefore find integers m, b' such that $m^2 = a + bb'$, where $|m| \leq \frac{|b|}{2}$. Moreover, we can, by rewriting this equation as $bb' = m^2 - a$, see that bb' is a norm of the field extension $k(\sqrt{a})$ over k where $k = \mathbb{Q}$ or $k = \mathbb{Q}_v$. Which means that f represent 0 if and only if f' represent 0, where

$$f' = x_1^2 - ax_2^2 - b'x_3^2.$$

This is the case since $f = x_1^2 - ax_2^2 - bx_3^2$ represent 0 if and only if $(a, b) = 1$ (which is the case if b is a norm of $k(\sqrt{a})/k$). So, since bb' is a norm of $k(\sqrt{a})/k$, we have $(a, bb') = (a, b)(a, b') = 1$, id est f represent 0 if and only if f' represent 0.

Now write $b' = b''u$ where $b'', u \in \mathbb{Z}$ and b'' square free, then $|b''| < b$ since we have, because $|b| \geq 2$ and $|a| \leq |b|$,

$$|b'| = \left| \frac{m^2 - a}{b'} \right| \leq \frac{|b|}{4} + 1 < |b|.$$

Whence we can conclude that the induction hypothesis applies to the quadratic form

$$f'' = x_1^2 - ax_2^2 - b''x_3^2,$$

which is equivalent to f', and so it will represent 0 in $\mathbb{Q}$, which means (by previous remark) that f will represent 0 in $\mathbb{Q}$ and we are done with this case.

3) Case $n = 4$:

In this case we write $f = ax_1^2 + bx_2^2 - (cx_3^2 + dx_4^2)$. For $v \in V$ we have by Corollary 4.12 that, since f_v represent 0, there exists a $y_v \in \mathbb{Q}_v^*$ such that both $ax_1^2 + bx_2^2$ and $cx_3^2 + dx_4^2$ represent y_v and by Corollary 4.14 (note that this also applies to $\mathbb{Q}_\infty = \mathbb{R}$) this is the same as the following equations being satisfied

$$(y_v, -ab)_v = (a, b)_v \text{ and } (y_v, -cd)_v = (c, d)_v \quad \forall\, v \in V.$$

After applying Theorem 6.8 (which we can apply since $\prod_{v\in V}(a,b)_v = \prod_{v\in V}(c,d)_v = 1$), we can procure a $y \in \mathbb{Q}^*$ such that

$$(y,-ab)_v = (a,b)_v \text{ and } (y,-cd)_v = (c,d)_v \quad \forall\ v \in V.$$

This means that the form $ax_1^2 + bx_2^2 - yz^2$ represent 0 in $\mathbb{Q}_v$ for all $v \in V$ and so will (by case 2) above) represent 0 in $\mathbb{Q}$. Thus $ax_1^2 + bx_2^2$ represent y in $\mathbb{Q}$, similarly, with the same argument, we have that $cx_3^2 + dx_4^2$ represent y in $\mathbb{Q}$ and with this we can conclude that f represent 0.

4) Case $n \geq 5$:

In this case we write f on the form $f = h - g$ where $h = a_1x_1^2 + a_2x_2^2$ and $g = a_3x_3^2 + a_4x_4^2 + ... + a_nx_n^2$.

Let $S \subset V$ be the subset of V such that $2, \infty$ and all p such that $\text{ord}_p(a_i) \neq 0$ for, at least, one $i \geq 3$ (id est for one of the coefficients of g) is in S. For $v \in S$ we have that, since f_v represent 0 by assumption, there exists $a_v \in \mathbb{Q}_v$ such that both h and g represent a_v in $\mathbb{Q}_v$. This in turn means that there exist $x_i^v \in \mathbb{Q}_v$ such that

$$h(x_1^v, x_2^v) = a_v = g(x_3^v, x_4^v, ..., x_n^v).$$

Now, by the so called approximation theorem (Theorem 6.7), there exists $x_1, x_2 \in \mathbb{Q}$ for which, if $h(x_1, x_2) = a$, we have that $\frac{a}{a_v} \in \mathbb{Q}_v^{*2}$, $\forall\ v \in S$. Now if we consider the quadratic form

$$f_1 = az^2 - g$$

then, if $z = 1$, we have $az^2 = a \cdot 1 = a$ which means that $\frac{a}{a_v} \in \mathbb{Q}_v^{*2}$. Moreover, this also means that, if $v \in S$, g will represent both a_v and a in $\mathbb{Q}_v$ and so f_1 will represent 0 in $\mathbb{Q}_v$. On the other hand, if $v \notin S$; as a direct consequence of Theorem 3.2 we have that polynomials without constant terms (like the polynomials we have here) have a non-trivial solution and by Hensel's lemma (Theorem 3.1) this solution can be lifted to a true solution.

We see, in any case, that f_1 represent 0 in $\mathbb{Q}_v$ and by the induction hypothesis, since f_1 is of rank $n-1$, f_1 will represent 0 in $\mathbb{Q}$, which means that g represent a in $\mathbb{Q}$ and so, since h represent a, f will represent 0 in $\mathbb{Q}$ and we are done.

□

References

[Kob77] Neal Koblitz. *p-adic Numbers, p-adic Analysis, and Zeta-Functions*. Springer-Verlag, 1977.

[Ser73] Jean-Pierre Serre. *A course in arithmetic*. Springer-Verlag, 1973.