Diplomarbeit

Berechnung des Hilbert Symbols, quadratische Form-Äquivalenz und Faktorisierung ganzer Zahlen

(Computing the Hilbert symbol, quadratic form equivalence and integer factoring)

Angefertigt am Mathematischen Institut

Vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

10. September 2013

Von

Lars Ambrosius Wallenborn

geboren am 26.01.1985
in
Bonn

Contents

Kapitel 0

Deutsche Einleitung

0.1 Übersicht

In dieser Arbeit werde ich einen Algorithmus vorstellen, der das Problem der quadratischen rationalen Formäquivalenz mithilfe eines Orakels für das Faktorisieren ganzer Zahlen in Polynomialzeit entscheidet.
Einer Turingmaschine ein Orakel für ein Entscheidungsproblem zur Verfügung zu stellen, bedeutet, ihr zu ermöglichen das Problem in nur einem Schritt zu entscheiden. Ein Orakel für IntFact ermöglicht beispielsweise das effiziente Faktorisieren ganzer Zahlen. Für eine Komplexitätsklasse **C** schreiben wir beispielsweise $\mathbf{C}^{\text{IntFact}}$ für die Klasse an Problemen, die mithilfe eines Orakels für IntFact in **C** entschieden werden können.
Ich verwende die folgenden Kürzel für Komplexitätsklassen:

- **P** in polynomieller Zeit entscheidbares Problem.
- **NP** in nicht-deterministisch polynomieller Zeit entscheidbares Problem.
- **ZPP** in randomisiert polynomieller Zeit entscheidbares Problem.
- **EXP** in exponentieller Zeit entscheidbares Problem.

Für **C** notieren wir mit **coC** die Klasse der Entscheidungsprobleme deren Komplement in **C** liegt.
Die Menge aller Polynome vom Totalgrad d über dem Körper $\mathbb{F}$ notiere ich mit $\mathbb{F}[X_1, \ldots, X_n]^{=d}$. Zwei Formen $f, g \in \mathbb{F}[X_1, \ldots, X_n]^{=d}$ heißen äquivalent, wenn es eine invertierbare lineare Transformation τ in den Variablen gibt, sodass $f \circ \tau = g$. Wir schreiben dann auch $f \sim g$, siehe Abschnitt 2.3 für Details. Das zugehörige Entscheidungsproblem ist:

$$\text{QuadFormEquiv}_{\mathbb{F}} := \{ (f, g) \mid f, g \text{ quadratische Formen über } \mathbb{F} \text{ und } f \sim g \}.$$

Ich werde mich vor allem mit dem Spezialfall $\mathbb{F} = \mathbb{Q}$ beschäftigen. In [AS06b] wird gezeigt, dass $\text{QuadFormEquiv}_{\mathbb{Q}} \in \mathbf{EXP}$ gilt, also in exponentiell vielen Schritten in der Eingabegröße entschieden werden kann. Ein Hauptresultat der hier vorliegenden Arbeit ist eine Verbesserung dieser oberen Schranke zu:

$$\text{QuadFormEquiv}_{\mathbb{Q}} \in \mathbf{P}^{\text{IntFact}}.$$

Dies bedeutet, dass $\textsc{QuadFormEquiv}_{\mathbb{Q}}$ in polynomiell vielen Schritten möglich ist, wenn ein Orakel für IntFact zur Verfügung steht, IntFact also effizient entschieden werden kann. Da $\textsc{IntFact} \in \mathbf{NP} \cap \mathbf{coNP}$ liefert dies folgende verbesserte Schranke:

$$\textsc{QuadFormEquiv}_{\mathbb{Q}} \in \mathbf{NP} \cap \mathbf{coNP}.$$

Der dafür in dieser Arbeit geführte Beweis ist in sich geschlossen. Dabei werden nur grundlegende Eigenschaften des Hilbert-Symbols[1] und p-adischer Zahlen[2] verwendet. Außerdem ist meine Beweisstrategie sehr verschieden zu der in [Har08]. Dort wird zusätzlich gezeigt, dass das funktionale Problem, also auch das *Finden* der Äquivalenz τ, mithilfe eines Orakels für IntFact in Polynomialzeit möglich ist. Dieses Problem sei mit $\textsc{FuncQuadFormEquiv}_{\mathbb{Q}}$ notiert.
Eine weitere Aussage aus [Har08] ist, dass sich $\sqrt{-1} \pmod n$ effizient bestimmen lässt, wenn $\textsc{FuncQuadFormEquiv}_{\mathbb{Q}}$ effizient möglich ist. In dieser Arbeit werde ich zusätzlich zeigen, dass IntFact selbst in randomisierter Polynomialzeit möglich ist, wenn $\textsc{FuncQuadFormEquiv}_{\mathbb{Q}}$ effizient gelöst werden kann, was τ aus praktischer Sicht sehr interessant macht:

$$\textsc{IntFact} \in \mathbf{ZPP}^{\textsc{FuncQuadFormEquiv}_{\mathbb{Q}}}.$$

In Abschnitt 5.2 wird $\textsc{FuncQuadFormEquiv}_{\mathbb{F}}$ gelöst, indem es auf das Lösen einer diagonalen quadratischen Gleichung reduziert wird — ein Algorithmus dafür findet sich in [Sim05].
Am Ende werde ich einige verwandte bekannte Resultate über quadratische Formäquivalenz über verschiedenen Körpern aufzählen und beschreiben, warum das Problem kubischer Formäquivalenz wesentlich schwieriger zu sein scheint. Eventuell handelt es sich dabei sogar um den schwierigsten Grad. In diesem Kontext stelle ich des Weiteren einige interessante offene Fragen und Implikationen vor.

0.2 Komplexitätstheoretischer Kontext

Abbildung 1 stellt die Hauptresultate dieser Arbeit schematisch dar. Wenn dort ein Pfeil von einem Problem A auf ein Problem B zeigt, soll dies bedeuten, dass $B \in \mathbf{P}^A$, dass also B mithilfe eines Orakels für A effizient gelöst werden kann. Ein Zickzack-Pfeil gibt an, dass dafür außerdem der Zugriff auf eine Zufallsquelle nötig ist, dass also $B \in \mathbf{ZPP}^A$ gilt. Die Beweise für Aussagen schwarzer Pfeile sind in dieser Arbeit enthalten, ein Beweis für $\textsc{FuncQuadFormEquiv}_{\mathbb{Q}} \in \mathbf{P}^{\textsc{IntFact}}$ is referenziert.
Das Problem $\textsc{HilbertSymbol}_{\mathbb{Q}}$ ist zu entscheiden, ob die Gleichung

$$aX^2 + bY^2 = Z^2$$

mit $a, b \in \mathbb{Q}^*$ eine Lösung $(x, y, z) \in \mathbb{Q}^3 \setminus \{0\}$ hat.
Zur Definition der Probleme SqrtMod und QuadResidue in Abbildung 1 seien $x, n \in \mathbb{Z}$. Wir definieren nun eine Kurzschreibweise für „x ist ein quadratischer Rest modulo n“ oder, äquivalent, „x hat eine Quadratwurzel in $\mathbb{Z}/n\mathbb{Z}$“:

$$x \mathrm{R} n :\Leftrightarrow \exists\, s \in \mathbb{Z}/n\mathbb{Z} : s^2 \equiv x \pmod n.$$

[1] siehe Kapitel 3
[2] siehe zum Beispiel [Kob77]

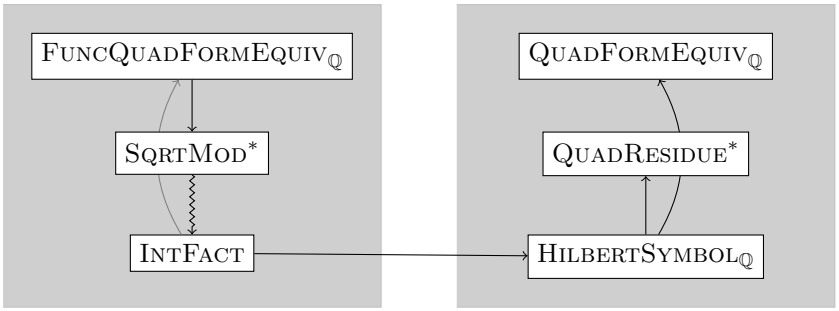


Abbildung 1: Schematische Darstellung der Resultate dieser Arbeit

Wenn $x\mathrm{R}n$ nicht gilt, schreiben wir $x\mathrm{N}n$. Hiermit definieren wir nun die genannten Probleme:

$$\textsc{QuadResidue} := \left\{ (x, n) \in \mathbb{Z}^2 \;\middle|\; x\mathrm{R}n \right\}$$

und das zugehörige funktionale Problem SqrtMod: Finde zu einem gegebenen x die Wurzel s mit $s^2 \equiv x \pmod{n}$ oder entscheide, dass es keine gibt.
QuadResidue* ist ein Spezialfall von QuadResidue: In diesem sei n von der Form pq mit für Primzahlen p und q für die gilt $p, q \equiv 1 \pmod{4}$. Weiterhin seien die Instanzen quadratische Reste bezüglich p und q oder bezüglich keinem der beiden — wir werden diese Relation durch $\{ p, q \} \vdash r$ notieren und sagen, dass sich $\{ p, q \}$ quadratisch einig über r ist. SqrtMod* ist die Einschränkung von SqrtMod auf invertierbare Instanzen.
Obwohl die Reduktion von IntFact zu FuncQuadFormEquiv$_{\mathbb{Q}}$ — und darum zu QuadFormEquiv$_{\mathbb{Q}}$ — kein neues Resultat ist, ist die Beweistechnik für QuadFormEquiv$_{\mathbb{Q}} \in \mathbf{P}^{\textsc{IntFact}}$, nämlich der Weg über HilbertSymbol$_{\mathbb{Q}}$, eine neue Idee. Das Vorkommen von HilbertSymbol$_{\mathbb{Q}}$ motiviert dann die Suche nach einer unteren Schranke. Ich werde zeigen, dass QuadResidue* eine solche ist.
Es sei zudem noch angemerkt, dass ich nicht nur zeigen werde, dass sich das Hilbert-Symbol mithilfe eines Orakels für IntFact in Polynomialzeit berechnen lässt, sondern auch, dass es sogar in quadratisch vielen Schritten möglich ist. Für einen Quantencomputer ist HilbertSymbol$_{\mathbb{Q}}$ mithilfe von Shors Faktorisierungsalgorithmus also sogar in $\mathcal{O}\left(n^6\right)$ Schritten möglich:

$$\textsc{HilbertSymbol}_{\mathbb{Q}} \in \mathbf{BQP}.$$

Die Komplexität von QuadResidue ist unbekannt. Die funktionale Version dieses Problems, SqrtMod, führt allerdings zu einem effizienten Algorithmus zum Faktorisieren ganzer Zahlen (siehe Abschnitt 5.3.2), was impliziert, dass es mindestens so schwierig ist wie IntFact. Gemeinhin ist dies als praktisch relativ schwierig oder zumindest von großem Interesse für die Kryptographie bekannt.
Ein anderer Grund für die Annahme, QuadResidue sei schwierig, ist, dass man für *generische Ring-Algorithmen* — dies sind Algorithmen, die einen Ring nur über eine „Schnittstelle“ für Addition, Multiplikation, Invertieren, Vergleichen

und zufällige Wahl eines Elements verwenden — mithilfe von IntFact quadratische Reste von Zahlen mit Jacobi-Symbol 1 unterscheiden kann. Ein Beweis dafür findet sich in [JS08]. Dort wird zudem gezeigt, dass es keinen effizienten Ring-Algorithmus für das Jacobi-Symbol geben kann. Obwohl die Komplexität von QuadResidue also unbekannt ist, muss ein effizienter Algorithmus irgendeine zusätzliche Struktur auf $\mathbb{Z}$ verwenden.

0.3 Klassifikation quadratischer Formen

Die zentrale Idee des Algorithmus ist eine Klassifikation rationaler quadratischer Formen, die sehr elegant in [Ser73] ausgearbeitet ist. Ich werde in dieser Arbeit auf die komplexitätstheoretischen Aspekte dieser Klassifikation eingehen. Es sei $\mathbb{F}$ ein beliebiger Körper, $f \in \mathbb{F}[X_1, \ldots, X_n]$ eine quadratische Form und $B = \{ b_1, \ldots, b_n \}$ eine Basis von $\mathbb{F}^n$. Definiere für $x, y \in \mathbb{F}^n$

$$x.y := \frac{f(x+y) - f(x) - f(y)}{2}$$

und ordne der Basis B durch $a_{ij} := b_i.b_j$ die symmetrische Matrix $A = (a_{ij})_{ij}$ zu. Dabei stellt sich heraus, dass sich eine Basis wählen lässt, bezüglich welcher A Diagonalgestalt hat.
Außerdem lässt sich zeigen, dass die assoziierten quadratischen Moduln genau dann isomorph sind, wenn die quadratischen Formen nach obiger Definition äquivalent sind. Die Isomorphie der quadratischen Moduln ist vollständig über gewisse Invarianten klassifiziert. Ein Beweis findet sich in Kapitel 4. Die Invarianten sind für deutschsprachige Leser in Abschnitt 0.6 definiert.
Eine rationale quadratische Form lässt sich natürlich auch als Form über den reellen Zahlen $\mathbb{R}$ oder über den p-adischen Zahlen $\mathbb{Q}_p$ für eine Primzahl p auffassen. Aus Notationsgründen definieren wir P als die Menge der Primzahlen und V als die Menge der Primzahlen zusammen mit dem Symbol ∞. Wir verwenden weiterhin die Konvention $\mathbb{Q}_\infty := \mathbb{R}$.
Wir sagen, eine Form „f repräsentiert 0 (über $\mathbb{F}$)“, wenn es $(x_1, \ldots, x_n) \in \mathbb{F}^n \setminus \{ \mathbf{0} \}$ gibt, sodass $f(x_1, \ldots, x_n) = 0$ gilt, wenn es also eine Nullstelle von f gibt, die nicht $\mathbf{0}$ ist. Das Hasse-Minkowski Theorem, auch bekannt als „Local-Global-Principle“, sagt nun aus, dass die Form f genau dann 0 über $\mathbb{Q}$ repräsentiert, wenn sie sowohl über $\mathbb{R}$ als auch über allen p-adischen Zahlen 0 repräsentiert. Es gilt ebenfalls die analoge Aussage für Äquivalenz: Für zwei quadratische Formen $f, g \in \mathbb{Q}[X_1, \ldots, X_n]$ gilt also:

$$f \sim g \text{ über } \mathbb{Q} \quad \Leftrightarrow \quad \forall\, v \in V\colon f \sim g \text{ über } \mathbb{Q}_v.$$

0.4 Hilbert-Symbol und Hilbert-Menge

Ein zentrales Resultat dieser Arbeit sind obere und untere Schranken für die Komplexität der Berechnung des Hilbert-Symbols über rationalen Zahlen.

Definition 0.1. Es sei $\mathbb{F}$ ein Körper und $R \subseteq \mathbb{F}$ ein Unterring. Die folgende Abbildung soll die nicht triviale Lösbarkeit einer quadratischen diagonalen

Gleichung in drei Variablen mit Koeffizienten in $\mathbb{F}$ erfassen:

$$\begin{array}{rcl} \mathrm{h}_{R,\mathbb{F}}(\cdot,\cdot) : \mathbb{F}^* \times \mathbb{F}^* & \longrightarrow & \{-1,1\} \\ (a,b) & \longmapsto & \begin{cases} 1 & \text{wenn } \exists\,(x,y,z)\in R^3\setminus\{\mathbf{0}\}: \\ & \quad ax^2+by^2=z^2 \\ -1 & \text{sonst.} \end{cases} \end{array}$$

Für den Fall $R = \mathbb{F}$ schreiben wir auch $\mathrm{h}_{\mathbb{F}}(\cdot,\cdot) := \mathrm{h}_{\mathbb{F},\mathbb{F}}(\cdot,\cdot)$, für $\mathbb{F} = \mathbb{Q}_v, v \in V$, auch $(\cdot,\cdot)_v := \mathrm{h}_{\mathbb{Q}_v}(\cdot,\cdot)$. Die Abbildung $(\cdot,\cdot)_v$ heißt **Hilbert-Symbol**.

Das Hilbert-Symbol $(\cdot,\cdot)_v$ ist symmetrisch und bilinear. Diese und weitere nützliche Aussagen leiten sich aus dem folgenden, recht technischen, Theorem 0.3 her. Dort taucht unter anderem das Legendre-Symbol auf, das wie folgt definiert ist:

Definition 0.2 (Legendre-Symbol). Wir notieren das **Legendre-Symbol** für $p \in P \setminus \{2\}$ durch

$$\begin{array}{rcl} \left(\frac{\cdot}{p}\right) : \mathbb{F}_p & \longrightarrow & \{-1,0,1\} \\ x & \longmapsto & x^{\frac{p-1}{2}}. \end{array}$$

Eine der wichtigsten Eigenschaften des Legendre-Symbols ist: Für alle $x \in \mathbb{F}_p^*$ gilt $\left(\frac{x}{p}\right) = 1$ genau dann wenn $x \mathrm{R} p$.

Theorem 0.3. *Es sei $v \in V$ und $a, b \in \mathbb{Q}_v^*$. Das Hilbert-Symbol lässt sich dann wie folgt berechnen:*

Wenn $v = \infty$: *Es gilt $(a,b)_\infty = -1$ genau dann wenn $a, b < 0$.*

Wenn $v \in P$: *Wir schreiben $a = v^\alpha u$ und $b = v^\beta w$ wobei u und w v-adische Einheiten sind. Es gilt dann:*

Wenn $v = 2$: $(a,b)_v = (-1)^{\epsilon(u)\,\epsilon(w)+\alpha\,\omega(w)+\beta\,\omega(u)}$.

Wenn $v \neq 2$: $(a,b)_v = (-1)^{\alpha\beta\,\epsilon(v)} \left(\frac{u}{v}\right)^\beta \left(\frac{w}{v}\right)^\alpha$.

Hier ist $\left(\frac{u}{v}\right)$ das Legendre-Symbol modulo v (also $\left(\frac{\pi(u)}{v}\right)$, wobei $\pi\colon \mathbb{Z}_v^* \to \mathbb{F}_v^*$ die kanonische Projektion ist). Weiterhin ist $\epsilon(u) := \frac{u-1}{2}$ und $\omega(u) := \frac{u^2-1}{8}$.

Für das Hilbert-Symbol gilt zudem die folgende Aussage:

Theorem 0.4 (Hilbert). *Für $a, b \in \mathbb{Q}^*$ gilt $(a,b)_v = 1$ für fast alle $v \in V$ und*

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

Klassische Beweise für diesen Satz zeigen erst, dass die Teilmenge von V für die das Hilbert-Symbol gleich -1 ist, endlich ist. In dieser Arbeit werde ich zeigen, dass diese Menge mithilfe eines Orakels für IntFact sogar in $\mathcal{O}(n^2)$ Schritten aufgelistet werden kann (n ist die maximale Bitgröße der Zähler und Nenner von a und b).

Definition 0.5 (Hilbert-Menge). Definiere für $a, b \in \mathbb{Q}^*$:

$$\mathcal{H}_{a,b} := \{\, v \in V \mid (a,b)_v = -1 \,\}.$$

Mithilfe von Theorem 0.3 lässt sich das folgende Lemma beweisen, welches, zusammen mit der Tatsache dass

$$\forall\, a, b, c \in \mathbb{Q}^* : \quad \mathcal{H}_{a,bc} = \mathcal{H}_{a,b} \triangle \mathcal{H}_{a,c}$$

gilt, Theorem 0.4 als Korollar liefert. $\triangle$ bezeichnet hier die symmetrische Differenz.

Lemma 0.6. *Für $a, b \in P \cup \{ -1 \}$ ist $\mathcal{H}_{a,b}$ wie folgt*

(i). $\mathcal{H}_{-1,-1} = \{ \infty, 2 \}$

(ii). *Für* $p \in P$: $\mathcal{H}_{p,p} = \mathcal{H}_{-1,p} = \begin{cases} \varnothing & \text{wenn } p = 2 \text{ oder } \epsilon(p) \text{ ungerade} \\ \{ 2, p \} & \text{sonst.} \end{cases}$.

(iii). *Für* $p \in P \setminus \{ 2 \} : \mathcal{H}_{p,2} = \begin{cases} \{ 2, p \} & \text{wenn } \omega(p) \text{ gerade} \\ \varnothing & \text{sonst.} \end{cases}$

(iv). *Für verschiedene $p, q \in P \setminus \{ 2 \}$ gilt $\mathcal{H}_{p,q} \subseteq \{ 2, p, q \}$ und $|\mathcal{H}_{p,q}| = 2$.*

Unter Zuhilfenahme dieses Lemmas und der Linearität des Hilbert-Symbols liefert eine Faktorisierung der Zähler und Nenner von $a = \pm\frac{a_1}{a_2}, b = \pm\frac{b_1}{b_2} \in \mathbb{Q}^*$ dann die Hilbert-Menge $\mathcal{H}_{a,b}$ in polynomieller Zeit. Genauer benötigt man $\mathcal{O}\left(n^2\right)$ Schritte wobei $n = \log(\max\{ a_1, a_2, b_1, b_2 \})$.
Das Hasse-Minkowski-Theorem liefert, dass genau dann $aX^2 + bY^2 - Z^2$ eine nichttrivale Nullstelle in $\mathbb{Q}^3$ hat, wenn sie für alle $v \in V$ eine nicht triviale Nullstelle in $\mathbb{Q}_v^3$ hat. Es gilt also, dass für $a, b \in \mathbb{Q}^*$

$$aX^2 + bY^2 = Z^2$$

genau dann eine nicht triviale Lösung hat, wenn $\mathcal{H}_{a,b} = \varnothing$. Dies lässt sich aufgrund der gerade geführten Argumentation in polynomieller Zeit überprüfen, gesetzt den Fall, dass man in der Lage ist, ganze Zahlen effizient zu faktorisieren.

0.5 Untere Schranke für das Hilbert-Symbol

QuadResidue scheint, zumindest ohne ein Orakel für IntFact, ein sehr schwer zu lösendes Problem zu sein. Seine funktionale Version (also Wurzel ziehen modulo n) liefert beispielsweise einen randomisierten Algorithmus für IntFact, was ein Hinweis dafür ist, dass QuadResidue selbst auch relativ schwierig ist. Wir wollen QuadResidue darum auf den folgenden Spezialfall einschränken: Der Modulus n ist Produkt zweier Primzahlen $p, q \in P$, für die gilt $p, q \equiv 1 \pmod 4$. Außerdem soll gelten, dass $\{ p, q \}$ sich „quadratisch einig“ über r ist:

Definition 0.7 (Quadratische Einigkeit). Es sei $r \in \mathbb{Z}_{>0}$ eine natürliche Zahl und $Q \subseteq \mathbb{Z}_{>0}$ eine endliche Menge natürlicher Zahlen. Wir sagen, dass Q **sich quadratisch einig über** r **ist** und schreiben dann $Q \vdash r$, wenn jeder Faktor von r entweder ein quadratischer Rest modulo jedem oder keinem Element von Q ist. Etwas formaler ausgedrückt schreiben wir $Q = \{ n_1, \ldots, n_l \} \subseteq \mathbb{Z}_{>0}$ und die Primfaktorzerlegung von r:

$$r = \prod_{i=1}^{k} p_i \quad k \in \mathbb{N}, \forall\, i \in [1 : k] : p_i \in P.$$

Dann ist sich Q quadratisch einig über r genau dann, wenn für alle $i \in [1:k]$ gilt:

$$\forall\, j \in [1:l]\colon p_i \mathrm{R} n_j \quad \text{oder} \quad \forall\, j \in [1:l]\colon p_i \mathrm{N} n_j.$$

Abgesehen von einigen Spezialfällen werde ich in dieser Arbeit zeigen, dass r ein quadratischer Rest modulo n ist genau dann, wenn $\mathrm{h}_{\mathbb{Q}}(r,n) = 1$ gilt. Dies zeigt, dass $\textsc{QuadResidue}^* \leq_{\mathbf{p}} \textsc{HilbertSymbol}_{\mathbb{Q}}$. Diese Resultate sind schematisch in Abbildung 2 dargestellt.

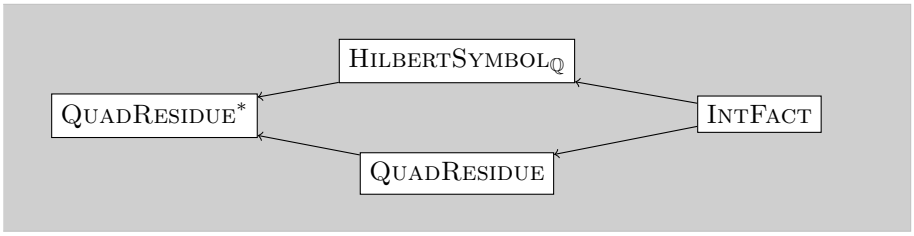


Abbildung 2: Schematische Darstellung der Komplexität des Hilbert-Symbols

0.6 Invarianten quadratischer Formen

Nun, da wir einen Eindruck von der Komplexität des Hilbert-Symbols haben, wollen wir das Problem quadratischer Formäquivalenz über rationalen Zahlen entscheiden. Dazu führen wir zunächst einige Invarianten ein. Es wird sich herausstellen, dass diese quadratische rationale Formen schon vollständig klassifizieren. Dies liefert zunächst die Entscheidbarkeit quadratischer rationaler Formäquivalenz. Da sich das Hilbert-Symbol aber mithilfe von IntFact berechnen lässt, liefern sie weiterhin einen Algorithmus zur Entscheidung quadratischer Formäquivalenz mithilfe eines Orakels für IntFact. Sei also f eine quadratische Form, $\{ b_1, \ldots, b_n \}$ eine Basis und $\cdot_{\bullet}\cdot$ das zugehörige Skalarprodukt.

Rang Der Rang einer quadratischen Form f ist definiert als der Rang der Matrix A wobei, wie oben erwähnt $A := (b_i \bullet b_j)_{i,j}$ für eine Basis $\{ b_i \}_i$. Dieser entspricht quasi der minimal „nötigen“ Anzahl von Variablen.

Diskriminante $\operatorname{disc}(f) := \det(A)$ in $\mathbb{Q}\Big/[\mathbb{Q}^*]^2$ (hier ist $[\mathbb{Q}^*]^2$ die Menge der Quadrate in $\mathbb{Q}^*$). Die Determinante ist eindeutig bis auf Multiplikation mit einem Quadrat, da ein Basiswechsel der quadratischen Form mit einer Matrix X für die Determinante eine Multiplikation mit $\det(X)^2$ bedeutet.

Signatur Eine quadratische Form mit vollem Rang lässt sich über $\mathbb{R}$ diagonalisieren. Das Paar (r,s), wobei r die Anzahl der positiven Koeffizienten und s die Anzahl der negativen Koeffizienten ist, heißt Signatur von f.

Hasse-Minkowski-Invariante Für $v \in V$ definiere

$$\begin{array}{rcl} \varepsilon\colon (\mathbb{Q}_v^*)^n & \longrightarrow & \{ \pm 1 \} \\ (a_1, \ldots, a_n) & \longmapsto & \prod_{i<j} (a_i, a_j)_v \,. \end{array}$$

Wenn (V, Q) ein nicht degenerierter quadratischer Modul von Rang n mit orthogonaler Basis $B = \{ b_1, \ldots, b_n \}$ ist, definiere weiterhin

$$\varepsilon(B) := \varepsilon(b_1.b_1, \ldots, b_n.b_n).$$

Es wird sich herausstellen, dass ε unabhängig von der Basis B ist, es sich dabei also um eine Invariante handelt.

Diese Invarianten klassifizieren quadratische Formäquivalenz über rationalen Zahlen schon vollständig und machen sie damit insbesondere entscheidbar. In Hinsicht auf die Komplexität ist die Hasse-Minkowski-Invariante am schwierigsten zu berechnen. Erst ein Orakel für das Hilbert-Symbol macht auch sie polynomiell berechenbar.

Chapter 1

Introduction

1.1 Outline

In this thesis, I will present an elementary algorithm which is able to decide quadratic rational form equivalence in polynomial time using an oracle for IntFact. In this context, a form is a multivariate homogeneous polynomial of degree 2 and two forms f and g over a field $\mathbb{F}$ will be called equivalent if there exists an invertible linear transformation τ in the variables such that $f \circ \tau = g$. One then writes $f \sim g$ (see Section 2.3 for details). The corresponding decision problem is

$$\text{QuadFormEquiv}_{\mathbb{F}} := \{\, (f, g) \mid f, g \text{ are quadratic forms over } \mathbb{F}, f \sim g \,\}.$$

We are especially interested in the case $\mathbb{F} = \mathbb{Q}$. In [AS06b] it is shown that $\text{QuadFormEquiv}_{\mathbb{Q}} \in \mathbf{EXP}$. One main result of this work is to improve this to

$$\text{QuadFormEquiv}_{\mathbb{Q}} \in \mathbf{P}^{\text{IntFact}}.$$

Since $\text{IntFact} \in \mathbf{NP} \cap \mathbf{coNP}$, this immediately implies

$$\text{QuadFormEquiv}_{\mathbb{Q}} \in \mathbf{NP} \cap \mathbf{coNP}.$$

The self-contained proof of this statement will only use basic properties of the Hilbert-symbol (see Chapter 3) and well known facts about p-adic numbers (see [Kob77]). In addition, my strategy is different from the methods used in [Har08], where it is shown that even *finding* the equivalence can be done in polynomial time using an oracle for IntFact.

There, it is also proven that if finding the equivalence can be done without the oracle for IntFact, one can find $\sqrt{-1} \pmod{n}$ in polynomial time. Adding to this, I will prove that finding the equivalence also implies that IntFact itself can be done in randomized polynomial time, which makes it of large practical interest:

$$\text{IntFact} \in \mathbf{ZPP}^{\text{FuncQuadFormEquiv}_{\mathbb{Q}}}.$$

Furthermore in Section 5.2 I will give the details about the process of finding rational quadratic form equivalence. For this, I will make use of the algorithm presented in [Sim05] to solve diagonal rational equations efficiently.

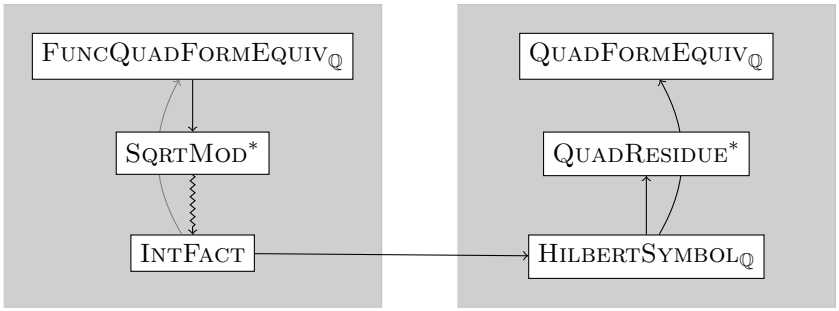


Figure 1.1: Schematic representation of the results of this work

Finally, I will list known results for quadratic form equivalence over different fields and explain why the problem of cubic form equivalence is much harder to solve – and may even be the hardest case. In this context, I will also discuss interesting open questions and their implications.

Figure 1.1 illustrates the results of this work. When an arrow points from a problem A to a problem B, it means that $B \in \mathbf{P}^A$ (or $B \in \mathbf{ZPP}^A$ when indicated by a zigzag arrow). The proofs for black arrows are included in this work — all others are referenced.

With QuadResidue* I refer to the special version of QuadResidue where the modulus n is of the form pq with prime numbers p and q for which it holds that $p, q \equiv 1 \pmod 4$. The instances are quadratic residues modulo both p and q or neither of them — we will denote this relation as $\{ p, q \} \vdash r$ and say that $\{ p, q \}$ quadratically agrees on r.

Even though the reduction from IntFact to FuncQuadFormEquiv$_{\mathbb{Q}}$ — and therefore to QuadFormEquiv$_{\mathbb{Q}}$ — is not a new result, the technique to prove QuadFormEquiv$_{\mathbb{Q}} \in \mathbf{P}^{\text{IntFact}}$ by taking the path over HilbertSymbol$_{\mathbb{Q}}$ is a new idea. The occurence of HilbertSymbol$_{\mathbb{Q}}$ then motivated to prove the lower bound QuadResidue* for the Hilbert symbol.

It is also noteworthy that I will not only show that HilbertSymbol$_{\mathbb{Q}}$ can be done in polynomial time given an oracle for IntFact, but also that it only takes a quadratic number of steps. In terms of quantum computation, it is well known that Shor's algorithm factors an integer of bitsize n in $\mathcal{O}\left(n^3\right)$ steps. So HilbertSymbol$_{\mathbb{Q}} \in \mathbf{BQP}$ or, more precisely, one needs $\mathcal{O}\left(n^6\right)$ steps to calculate the Hilbert symbol for an input of bitsize n on a quantum computer.

The complexity of QuadResidue is unknown. Its functional version though, which is SqrtMod, leads to an efficient algorithm to factor integers (see Section 5.3.2), which means that it is as hard as IntFact, which is believed to be quite hard or, at least, cryptographically interesting.

Another reason for the hardness of QuadResidue is that, in terms of *generic ring algorithms* — these are algorithms that use integers only over some sort of "interface" with calls for addition, multiplication, inversion, comparison and choosing random elements — IntFact reduces to distinguish quadratic residues from numbers with Jacobi symbol 1. This is shown in [JS08]. They also show that the Jacobi symbol does not have an efficient generic ring algorithm. So even though the complexity of QuadResidue is not known, an efficient algorithm

needs to make use of some structure of $\mathbb{Z}$ that a general ring does not possess.

1.2 Structure of this work

Here you can find the collected abstract of the differnt chapters

Preliminaries (Chapter 2)
In this chapter, we will give all basic definitions and notations that are used throughout this work. Additionally some basic results about finite fields, the Legendre symbol and p-adic numbers are established. If you are familiar with this topics, you can safely skip this chapter and only refer to it for reference.

The Hilbert Symbol (Chapter 3)
In this chapter, we will introduce the *Hilbert symbol* $(\cdot,\cdot)_v$ and prove some well-known properties that will afterwards be used to lay the foundation for the proof of the upper bound

$$\textsc{HilbertSymbol}_{\mathbb{Q}} \in \mathbf{P}^{\textsc{IntFact}}.$$

On the way, we define the *Hilbert set* for $a, b \in \mathbb{Q}^*$ by

$$\mathcal{H}_{a,b} := \{\, v \in V \mid (a,b)_v = -1 \,\}.$$

We will show that an oracle for IntFact leads to an algorithm that enumerates the Hilbert set in polynomial time.

Classification of Quadratic Forms (Chapter 4)
The goal of this chapter is to understand all preliminaries required for designing an algorithm which outputs a quadratic form equivalence. This case is significantly easier than equivalences for higher degree forms, because quadratic modules, which are closely connected to quadratic forms, are well-studied and have a lot of structure. We will define a quadratic module associated to a quadratic form and the isomorphy classes of these modules will correspond to the equivalence classes of the associated forms. Ultimately we will prove Witt's theorem yielding two useful corollaries:

(i). Every quadratic form is equivalent to a form $\sum_{i=1}^{n} a_i X_i^2$.

(ii). We have a cancellation rule for quadratic forms (the $\ominus$ will be defined on the way):

$$f \ominus h \sim g \ominus h \quad \Longrightarrow \quad f \sim g$$

This chapter is based on [Ser73].

Results (Chapter 5)
In this chapter, we will present the details about finding rational quadratic form equivalence and the proof that IntFact can be done in randomized polynomial time if $\textsc{FuncQuadFormEquiv}_{\mathbb{Q}} \in \mathbf{FP}$. This lower bound

$$\textsc{IntFact} \leq_{\mathbf{R}} \textsc{FuncQuadFormEquiv}_{\mathbb{Q}}$$

seriously improves the result by [Har08] where it is only shown that one can calculate $\sqrt{-1} \pmod{n}$ using an oracle for IntFact and that

$$\text{FuncQuadFormEquiv}_{\mathbb{Q}} \leq \text{IntFact}.$$

Outlook and Open Questions (Chapter 6)
In this chapter, we will give an outlook over possible development in the future and state some open questions that might be of interest.

1.3 Used literature and resources

The basis of this work is [Ser73]. Additionally, articles by Nitin Saxena et. al. provided valuable context — you can find references to all of them in the bibliography.

1.4 Acknowledgment

I would like to thank my supervisor, Prof. Dr. Nitin Saxena for his support, trust and every minute of his time. Every single meeting was illuminating and great fun. I also would like to thank Prof. Dr. Jens Franke for agreeing to be the second corrector of this diploma thesis.

Thanks go out to Jesko Hüttenhain and Tobias Gödderz for the good times studying mathematics together. Also thank you for all the support and feedback you gave me concerning this thesis — and everyhing else.

Special thanks go out to Anna-Eliane Müller and Oliver Nahm who managed to read the whole thesis despite the fact they never studied mathematics. I want to thank Luisa Schwartz for proofreading the whole thesis and even more for the constant support during this project.

Finally I whould like to thank my family, the people close to me and especially all of the above for inspiring me to become what I am.

Chapter 2

Preliminaries

Abstract

In this chapter, we will give all basic definitions and notations that are used throughout this work. Additionally some basic results about finite fields, the Legendre symbol and p-adic numbers are established. If you are familiar with this topics, you can safely skip this chapter and only refer to it for reference.

2.1 Basic definitions and notation

Notation 2.1 (General notations). Usually we will use the following symbols in this work: $\mathbb{F}$ is a field of arbitrary characteristic, $\mathbb{F}_q$ the field with q elements of characteristic p, where q is a power of a prime p. Furthermore, $\mathbb{F}[X_1, \ldots, X_n]$ is the ring of polynomials in n variables over the field $\mathbb{F}$. Also, we will use the following sets of numbers:

- $\mathbb{N}$ is the set of natural numbers. We use the convention that $0 \notin \mathbb{N}$ and denote $\mathbb{N}_0 := \mathbb{N} \cup \{ 0 \}$.
- $\mathbb{Z}$ is the set of integers.
- P is the set of primes and $V := P \cup \{ \infty \}$ the set of primes together with the symbol ∞.
- $\mathbb{Q}$ is the set of rational numbers, which is the quotient field of $\mathbb{Z}$.
- For $p \in P$ we denote by $\mathbb{Q}_p$ the field of p-adic numbers, which is the completion of $\mathbb{Q}$ with respect to the p-adic norm and Cauchy-sequences. For more details see for example [Kob77].
- $\mathbb{R}$ is the set of real numbers, which is the completion of $\mathbb{Q}$ with respect to the absolute value and Cauchy-sequences. Whenever it is convenient, we will also denote $\mathbb{R}$ by $\mathbb{Q}_\infty$.
- Every ring in this work will contain a 1 (i.e. a neutral element for multiplication).

- For $n \in \mathbb{N}_0$ we have the set of multiindices $\mathbb{N}_0^n$. For $d \in \mathbb{N}_0$ we define the set of multiindices of norm d:

$$[\mathbb{N}]^{=d} := \{\, \alpha \in \mathbb{N}_0^n \mid |\alpha|_1 = d \,\}$$

  and respectively

$$[\mathbb{N}]^{\leq d} := \{\, \alpha \in \mathbb{N}_0^n \mid |\alpha|_1 \leq d \,\},$$

  where $|(\alpha_1, \ldots, \alpha_n)|_1 = |\alpha_1| + \ldots + |\alpha_n|$. Finally define $[\mathbb{N}] := \bigcup_{d \in \mathbb{N}_0} [\mathbb{N}]^{=d}$.

- For $i, j \in \mathbb{Z}$, we define the set of integral numbers from i to j by

$$[i : j] := [i, j] \cap \mathbb{Z}.$$

  Sometimes, for shortness of notation, we will use

$$[j]_0 := [0 : j] \quad \text{and} \quad [j] := [1 : j].$$

- As a convention we set $\inf \varnothing := \infty$ and $\sup \varnothing := -\infty$.

- For a ring R, we denote by R^* the invertible elements of R.

- For $n \in \mathbb{N}$ we define a subgroup of the multiplicative group of $\mathbb{F}$ by

$$[\mathbb{F}^*]^n := \{\, x^n \mid x \in \mathbb{F}^* \,\}.$$

- For two functions $f, g \colon \mathbb{R} \to \mathbb{R}$ we write $f = \mathcal{O}(g)$ if and only if

$$\exists\, c, x_0 \in \mathbb{R}_{>0} \colon \forall\ x \in \mathbb{R}_{>x_0} \colon\ |f(x)| \leq c \cdot |g(x)|\,.$$

Definition 2.2 (Total Degree of a Polynomial). Let $f \in \mathbb{F}[X_1, \ldots, X_n]$ be a polynomial and write $f(X_1, \ldots, X_n) = \sum_{\alpha \in [\mathbb{N}]} a_\alpha x^\alpha$ with only finitely many $a_\alpha \neq 0$, then the **(total) degree of** f is given by

$$\deg(f) = \sup \{\, |\alpha| \mid a_\alpha \neq 0 \,\}.$$

Remark 2.3. Note that nonzero constants (elements of $\mathbb{F}^*$) have degree 0 and, since $\sup \{\, \varnothing \,\} := -\infty$, we have $\deg(0) = -\infty$.

Fact 2.4. *For polynomials $f, g \in \mathbb{F}[X_1, \ldots, X_n]$ and $h \in \mathbb{F}[x]$ it holds that*

$$\begin{aligned} \deg(f + g) &\leq \max(\deg(f), \deg(g)) \\ \deg(f \cdot g) &= \deg(f) + \deg(g) \\ \deg(h \circ g) &= \deg(h) \cdot \deg(g)\,. \end{aligned}$$

Remark 2.5. An easy corollary from the last equality in 2.4 is that when you replace a variable of a multivariate polynomial by an invertible linear combination of other variables, the degree does not change.

Definition 2.6. Let R be a ring, I an index set and $F = \{f_i\}_{i \in I}$ be a set of polynomials where $f_i \in R[X_1, \ldots, X_n]$ for any $i \in I$. Then we denote the set of common zeros by $V(F)$ i.e.

$$V(F) := \{\, x \in R^n \mid \forall\, f \in F \colon f(x) = 0 \,\}.$$

Remark 2.7. If the ring R in Definition 2.6 actually is a field, Hilbert's basis theorem yields that for an arbitrary set $F \subseteq \mathbb{K}[X_1, \ldots, X_n]$ there exists a finite subset $E \subseteq F$ such that $V(F) = V(E)$.

Definition 2.8. For $a, b \in \mathbb{Z}$ we say that a **divides** b if and only if

$$\exists\, k \in \mathbb{Z} \colon ak = b.$$

We then also write $a \mid b$. If $a \notin \{1, b\}$, we say that it is a **nontrivial divisor of** b. For $c \in \mathbb{Z}$ we write $a \equiv b \pmod{c}$ or $a \equiv_c b$ if $c \mid a - b$.

Algorithm 1 Extended Euclidean Algorithm

Input: $a, b \in \mathbb{N}$
Output: $(\gcd(a,b), s, t)$ with $\gcd(a,b) = sa + tb$

```
set x := 0, s := 1, y := 1, t := 0
while b ≠ 0 do
    set q := a div b
    set r := a mod b
    set (a, b) := (b, r)
    set (x, s) := (s − q · x, x)
    set (y, t) := (t − q · y, y)
end while
return (a, s, t)
```

Lemma 2.9 (Bézout's Lemma). $\forall\, a, b \in \mathbb{Z}\ \exists\, s, t \in \mathbb{Z}$:

$$\gcd(a,b) = sa + tb$$

Proof. Algorithm 1 yields such numbers. □

Theorem 2.10 (Chinese Remainder Theorem). *If $m, n \in \mathbb{Z}$ are coprime then*

$$\forall\, a, b \in \mathbb{Z}\ \exists\, x \in \mathbb{Z} \colon \begin{array}{ll} x \equiv a & \pmod{n} \text{ and} \\ x \equiv b & \pmod{m} \end{array}$$

or equivalently:

$$\gcd(n,m) = 1 \quad \Rightarrow \quad \mathbb{Z}/nm\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}\,.$$

Theorem 2.11 (Dirichlet Theorem). *For all $a, m \in \mathbb{N}$ with $\gcd(a,m) = 1$ there exist infinitely many primes $p \in P$ such that $p \equiv a \pmod{m}$.*

Proof. One can find a proof in [Ser73, Chapter IV] □

2.2 Complexity Theory

In this section, we will give an overview of the computational model used in this work and basic conventions, notations and complexity classes.

Whenever we speak of a Turing machine or an algorithm, we mean a Turing machine over a finite alphabet with three tapes:

- input tape: on this tape the input is given.
- output tape: when the machine M stops on an input after a finite number of steps, the output x has been written on this tape. We denote it by $M(x)$ and call it the **output of** M **on** x. The number of steps a Turing machine needs to stop is called the **time** it needs.
- work tape: this is the tape the machine uses during computation.

Of course this machine can perform exactly the same operations as any other "standard" Turing machine (e.g. with only one tape).

Now mathematical objects, as numbers, graphs and polynomials, can be encoded as strings of bits. Whenever we give such objects to algorithms, we mean that we write their encoding on the input tape. We assume a canonical encoding if it is obvious — e.g. one can encode natural numbers by their binary representation — and specify it, if it is not. "Writing an object on the input tape" or "giving it to an algorithm" means running the Turing machine on the encoding of the object.

A **decision problem** is a problem that has a boolean answer. A turing machine **decides** such a problem L, if it outputs 1 for every $x \in L$ and 0 for every $x \notin L$. If such a Turing machine exists, the problem is called **decidable**, and **undecidable** otherwise. We will use the following **complexity classes** of decidable problems in this work

- **P**: problems that can be decided in a polynomial number of steps in the size of the input.
- **EXP**: problems that can be decided in an exponential number of steps in the size of the input.
- **NP**: for "yes"-instances of problems in this class there exist certificates which can be checked in polynomial time.
- **coNP**: for "no"-instances of problems in this class there exist certificates which can be checked in polynomial time.

Allowing the turing machine to access a source of randomness, yields another complexity class:

- **ZPP**: problems that can be decided in a polynomial number of steps in the size of the input given a source of randomness.

Since we often want to talk about solutions to problems that are more complex than just one bit, we define **functional problems** in addition to decision problems.

Definition 2.12 (Functional problems). Let I, C be two sets, called **instances** and **certificates**. A binary relation $P \subseteq I \times C$ is called **functional problem**. The set **FP** consists of functional problems for which there exists a Turing machine M that runs for every $x \in I$ in polynomial time in the size of x with either $(x, M(x)) \in P$ or $M(x)$ is empty, meaning there is no $c \in C$ such that $(x, c) \in P$.

Definition 2.13 (Oracle Turing machines and complexity classes). Let L be a language or a functional problem. An **oracle Turing machine** M^L **with oracle** L is a turing machine with an extra tape — the **oracle tape** — and three extra states QUERY, ANSWERYES and ANSWERNO. M^L behaves like an ordinary Turing machine, except when M enters the QUERY state, with a word x on the query tape, the content of the oracle tape is replaced (in one step) as follows:

- If L is a decision problem, the content of the oracle tape is deleted and the machine continues execution in the ANSWERYES state if $x \in L$ and in the ANSWERNO state otherwise.
- If L is a functional problem, the content of the oracle tape is replaced by a certificate c such that $(x, c) \in L$ or deleted entirely if no such certificate exists. In the first case, the next state is ANSWERYES and ANSWERNO in the latter.

The complexity class of decision problems solvable by an algorithm in class $\mathbf{C}$ with an oracle for a L is denoted by $\mathbf{C}^L$.

Notation 2.14. When we are talking about the polynomial reducibility of two problems A and B, the following terms are equivalent:

- "$A \leq_{\mathbf{p}} B$" or even "$A \leq B$"
- "There exists a (polynomial) reduction from A to B."
- "A reduces (polynomially) to B."
- "A is a problem that is (polynomially) solvable by using an oracle for B."

If we additionally allow randomness within the reduction, we denote this by

- "$A \leq_{\mathbf{p},\mathbf{R}} B$" or even "$A \leq_{\mathbf{R}} B$"

Notation 2.15 (Notation in algorithms).

- Whenever we write **assert** in an algorithm, the truth value of the statement following is tested. If the statement is true, the algorithm continues normally and if it is false, the algorithm terminates and returns **false**.
- Let X be a finite set. When we write $x \in_R X$ in an algorithm, x is chosen uniformly at random from the set X.

2.3 Polynomial equivalence

Definition 2.16 (Polynomial Equivalence). Let $\mathbb{L}/\mathbb{F}$ be a field extension. Two polynomials $f, g \in \mathbb{F}[X_1, \ldots, X_n]$ are said to be **equivalent over** $\mathbb{L}$ if there exists an invertible linear transformation τ with coefficients in $\mathbb{L}$ sending each X_i to a linear combination of the $X_1, \ldots, X_n$:

$$f(\tau(X_1, \ldots, X_n)) = g(X_1, \ldots, X_n).$$

We then write $f \sim_{\mathbb{L}} g$. If $\mathbb{L} = \mathbb{F}$, we simply write $f \sim g$ and say that f and g are **equivalent** or, if we want to stress the invertible linear transformation, **equivalent via** τ.

Remark 2.17. An invertible linear transformation τ on the variables of a polynomial $f \in \mathbb{F}[X_1, \ldots, X_n]$ can also be expressed by a matrix $A \in \mathrm{Gl}_n(\mathbb{L})$ acting on $\mathbb{F}[X_1, \ldots, X_n]$. So abusing notation a little bit, we can say that

$$f \sim_{\mathbb{L}} g \quad \Longleftrightarrow \quad \exists\, A \in \mathrm{Gl}_n(\mathbb{L}) : f \circ A = g.$$

In this case, we say that f **is equivalent to** g **via** A.

Fact 2.18. *The equivalence of polynomials is indeed an equivalence relation.*

Proof. For all $f, g, h \in \mathbb{F}[X_1, \ldots, X_n]$:

Step 1 (Reflexivity). $f \sim_{\mathbb{L}} f$ via the identity matrix in $\mathrm{Gl}_n(\mathbb{L})$.

Step 2 (Symmetry). If $f \sim_{\mathbb{L}} g$ via $A \in \mathrm{Gl}_n(\mathbb{L})$ then $g \sim_{\mathbb{L}} f$ via A^{-1}.

Step 3 (Transitivity). If $f \sim_{\mathbb{L}} g$ via $A \in \mathrm{Gl}_n(\mathbb{L})$ and $g \sim_{\mathbb{L}} h$ via $B \in \mathrm{Gl}_n(\mathbb{L})$, then $f \sim_{\mathbb{L}} h$ via $B \cdot A \in \mathrm{Gl}_n(\mathbb{L})$. □

Example 2.19. Let $f = X^2 + Y^2$ and $g = 2X^2 + 2Y^2$ be polynomials over $\mathbb{Q}$. For the invertible linear transformation τ defined by

$$X \mapsto X + Y$$
$$Y \mapsto X - Y$$

we have that $f \circ \tau = g$, so $f \sim g$ over rationals. We could also say, that

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in \mathrm{Gl}_n(\mathbb{Q})$$

is invertible and calculate

$$f\left(\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right) = f\begin{pmatrix} x+y \\ x-y \end{pmatrix} = (x+y)^2 + (x-y)^2 = 2x^2 + 2y^2 = g(x, y)$$

Example 2.20. Consider $f, g \in \mathbb{Q}[X]$ with $f = X^2$ and $g = 2X^2$. Then f and g are not equivalent over $\mathbb{Q}$ but they are equivalent over $\mathbb{R}$ via $\tau : X \mapsto \sqrt{2}X$.

Fact 2.21. *Equivalent polynomials have the same degree.*

Proof. Let f and g be equivalent polynomials via $A \in \mathrm{Gl}_n(\mathbb{L})$. So A replaces every variable by a linear combination of X_i which does not change the degree (see Remark 2.5). □

Definition 2.22. For $d \in \mathbb{N}$ a polynomial $f \in \mathbb{F}[X_1, \ldots, X_n]$ of the form

$$f(X_1, \ldots, X_n) = \sum_{\alpha \in [\mathbb{N}]^{=d}} a_\alpha x^\alpha$$

is called **homogeneous polynomial of degree** d. Homogeneous polynomials are also called **forms**. Furthermore define:

- $\mathbb{F}[X_1, \ldots, X_n]^{=d}$ the forms of degree d.
- $\mathbb{F}[X_1, \ldots, X_n]^{\leq d}$ the forms of degree at most d.

In this section, we will state what we exactly mean by "giving a polynomial to an algorithm" and state the central decision problems of this work. We will also define other problems that will turn out to be connected to the problem of form equivalence.

Definition 2.23 (Polynomials as input). We assume every polynomial $f \in \mathbb{F}[X_1, \ldots, X_n]$ with total degree d to be given in **expanded** form:

$$f(X_1, \ldots, X_n) = \sum_{0 \leq i_1 + \ldots + i_n \leq d} a_{i_1,\ldots,i_n} X_1^{i_1} \cdot \cdots \cdot X_n^{i_n}$$

with $a_{i_1,\ldots,i_n} \in \mathbb{F}$. This can also be written in a more elegant way:

$$f(X_1, \ldots, X_n) = \sum_{\alpha \in [\mathbb{N}]^{\leq d}} a_\alpha X^\alpha$$

where $X = (X_1, \ldots, X_n)$ and $a_\alpha \in \mathbb{F}$.

Definition 2.24. We define the following decision problems:

$$\textsc{PolyEquiv}_{d,\mathbb{F}} := \left\{ (f,g) \in \mathbb{F}[X_1, \ldots, X_n]^2 \;\middle|\; \begin{array}{c} n \in \mathbb{N}, f \sim g, \\ \deg(f) = d = \deg(g) \end{array} \right\},$$
$$\textsc{FormEquiv}_{d,\mathbb{F}} := \left\{ (f,g) \in \left(\mathbb{F}[X_1, \ldots, X_n]^{=d}\right)^2 \;\middle|\; n \in \mathbb{N}, f \sim g \right\}.$$

For shortage of notation, we also define:

$$\begin{aligned} \textsc{QuadPolyEquiv}_{\mathbb{F}} &:= \textsc{PolyEquiv}_{2,\mathbb{F}} \\ \textsc{QuadFormEquiv}_{\mathbb{F}} &:= \textsc{FormEquiv}_{2,\mathbb{F}} \\ \textsc{CubicPolyEquiv}_{\mathbb{F}} &:= \textsc{PolyEquiv}_{3,\mathbb{F}} \\ \textsc{CubicFormEquiv}_{\mathbb{F}} &:= \textsc{FormEquiv}_{3,\mathbb{F}} \end{aligned}$$

Every such problem also has a canonical functional version: Define the sets of instances and certificates by

$$I_d := \bigcup_{n \in \mathbb{N}} \left(\mathbb{F}[X_1, \ldots, X_n]^{=d}\right)^2 \qquad C := \bigcup_{n \in \mathbb{N}} \mathrm{Gl}_n(\mathbb{F})$$

and the functional problem of form equivalence:

$$\textsc{FuncFormEquiv}_{d,\mathbb{F}} := \{ ((f,g), A) \in I_d \times C \mid f \circ A = g \}.$$

Again, we define

$$\begin{aligned} \textsc{FuncQuadFormEquiv}_{\mathbb{F}} &:= \textsc{FuncFormEquiv}_{2,\mathbb{F}} \\ \textsc{FuncCubicFormEquiv}_{\mathbb{F}} &:= \textsc{FuncFormEquiv}_{3,\mathbb{F}}. \end{aligned}$$

Remark 2.25. Note that the size of an instance is not the degree of the form (which is constant for any problem $\textsc{FormEquiv}_{d,\mathbb{F}}$ or $\textsc{FuncFormEquiv}_{d,\mathbb{F}}$) but the number of bits we need, to encode it. This depends on the size of the coefficients and the number of variables n.

2.4 Connected problems

It turns out that the complexity of the problem CubicFormEquiv is connected to the complexity of the problem to decide if two given commutative algebras are isomorphic or not. So we first formalize, how algebras can be given to an algorithm as input and then define the problem of commutative algebra isomorphism and its local version. Every algebra that we will give to an algorithm must be finite dimensional.

Definition 2.26 ($\mathbb{F}$-algebras as input). Let $\mathbb{F}$ be a field and A be a finitely dimensional commutative $\mathbb{F}$-algebra with $\mathbb{F}$-basis $b_1, \ldots, b_n \in A$. We now want to capture the multiplicative structure of the algebra and therefore write every product of base elements as a linear combination of all base elements:

$$\forall\, i,j,k \in [1:n] \colon \exists\, a_{ijk} \in \mathbb{F} \colon b_i b_j = \sum_{k=1}^{n} a_{ijk} b_k.$$

We call the a_{ijk} **structure coefficients**.

Fact 2.27. *Let A be an $\mathbb{F}$-algebra with additive basis $\{ b_i \}_{i\in[1:n]}$ and structure coefficients $\{ a_{ijk} \}_{i,j,k\in[1:n]}$ then:*

$$A \cong \mathbb{F}[X_1, \ldots, X_n] \Big/ \left(X_i X_j - \sum_{k=1}^{n} a_{ijk} X_k \right)_{i,j\in[1:n]}.$$

To specify an isomorphism $\psi : A \to B$ it is sufficient to write for every i the element $\psi(b_i)$ as linear combination of $b_1, \ldots, b_n$ in B.

Definition 2.28 (Ring decomposition). A ring R is **decomposable** if there exist subrings $R_1, R_2 \subseteq R$ such that:

(i). $R_1 \cdot R_2 = R_2 \cdot R_1 = \{0\}$

(ii). $R_1 \cap R_2 = \{0\}$

(iii). $R = R_1 + R_2$

We then denote such a ring by $R = R_1 \times R_2$. The subrings R_1, R_2 are called **components** of R. We say that R **decomposes non-trivially** if and only if $R_1, R_2 \neq \{0, 1\}$. If a ring does not decompose non-trivially, it is called **local**. Since every algebra is also a ring, we can also call an algebra local.

Remark 2.29. Often a ring/algebra is called local if it has a unique maximal ideal. Since our algebras are finite dimensional, this is equivalent to the the above definition of local by [AM69, Exercise 8.3].

Definition 2.30.

$$\begin{aligned} \textsc{CommAlgIso}_{\mathbb{F}} := \{(A,B) \mid\ & A, B \text{ commutative } \mathbb{F}\text{-algebras with} \\ & \text{basis } b_1, \ldots, b_n \text{ and } A \cong B\} \\ \textsc{LocalCommAlgIso}_{\mathbb{F}} := \{(A,B) \mid\ & A, B \text{ local commutative } \mathbb{F}\text{-algebras with} \\ & \text{basis } b_1, \ldots, b_n \text{ and } A \cong B\} \end{aligned}$$

Fact 2.31. *The following statements hold trivially:*

- $\textsc{FormEquiv}_{d,\mathbb{F}} \leq \textsc{PolyEquiv}_{d,\mathbb{F}}$
- $\textsc{LocalCommAlgIso}_{\mathbb{F}} \leq \textsc{CommAlgIso}_{\mathbb{F}}$
- $\textsc{CubicFormEquiv}_{\mathbb{F}} \leq \textsc{FormEquiv}_{d,\mathbb{F}}$

You can find a discussion about the connection between the form equivalence problems and the algebra isomorphism problems in Section 6.2, especially interesting is Theorem 6.4.

For the sake of completeness, we will formalize the decision problem of IntFact and its functional version here:

Definition 2.32.

$$\textsc{IntFact} := \left\{ (n,k) \in \mathbb{N}^2 \;\middle|\; k < n \text{ and } \exists\, d \in [2:k] \colon d \mid n \right\}.$$

The functional version can be defined as follows: Let the instances be $I := \mathbb{N}$ and the certificates be

$$C := \bigtimes_{k \in \mathbb{N}} (P \times \mathbb{N})^k.$$

With these now define

$$\textsc{FuncIntFact} := \left\{ (z, ((p_1, r_1), \ldots, (p_k, r_k))) \in I \times C \;\middle|\; z = \prod_{i=1}^{k} p_i^{r_i} \right\}.$$

Given an oracle for the decision problem, one can easily find the lowest prime factor of n by binary search:

Algorithm 2 LowestPrimeFactor-With-IntFact-Oracle

Input: $n \in \mathbb{N}$
Output: $\min\{\, p \in P \mid p \mid n \,\}$.

```
1:  l := 2
2:  h := n − 1
3:  while h > l do
4:      k := ⌊(l+h)/2⌋
5:      if (n, k) ∈ IntFact then
6:          h := k
7:      else
8:          l := k + 1
9:      end if
10: end while
11: if l + 1 | n then
12:     return l + 1
13: else if l | n then
14:     return l
15: else
16:     return n
17: end if
```

Theorem 2.33. *Algorithm 2 is correct and runs in polynomial time in the bit-size of n.*

After we found the lowest prime factor p of an input n, we can find the whole factorization by applying the algorithm to $\frac{n}{p}$ repeatedly (since there are only $\mathcal{O}(\log(n))$ prime factors, this method leads to the factorization in polynomial time).

This can be summarized as follows:

Theorem 2.34.

$$\textsc{FuncIntFact} \in \mathbf{FP}^{\textsc{IntFact}}.$$

In Section 5.3 we will also show that finding a quadratic form equivalence enables us to take square roots modulo n which implies that IntFact can be done in randomized polynomial time — which is shown in Section 5.3.2. The complexity of the decision version of the problem of finding a square root modulo n is unknown, see for example Problem 11 in [AM]. We will use this decision version to lower bound the decision problem of HilbertSymbol, defined in Definition 3.5. Later we shall see, that this problem will enable us to *decide* quadratic rational form equivalence.

Definition 2.35 (quadratic residue / square roots modulo n). Let $x, n \in \mathbb{Z}$. We now define shorthand notation to write that "x is a quadratic residue modulo n" or, equivalently, that "x has a square root in $\mathbb{Z}/n\mathbb{Z}$. We write

$$x \mathrm{R} n :\Leftrightarrow \exists\, s \in \mathbb{Z}/n\mathbb{Z} : s^2 \equiv x \pmod{n}$$

and $x\mathrm{N}n$ otherwise. With this, we define

$$\textsc{QuadResidue} := \left\{ (x,n) \in \mathbb{Z}^2 \;\middle|\; x \mathrm{R} n \right\}.$$

Let $I := \mathbb{Z}^2$ and

$$I^* := \left\{ (x,n) \in \mathbb{Z}^2 \;\middle|\; x \in (\mathbb{Z}/n\mathbb{Z})^* \right\}$$

be the instances and $C := \mathbb{Z}$ be the certificates. Now define

$$\textsc{SqrtMod} := \left\{ ((x,n),s) \in I \times C \;\middle|\; x \equiv s^2 \pmod{n} \right\},$$
$$\textsc{SqrtMod}^* := \left\{ ((x,n),s) \in I^* \times C \;\middle|\; x \equiv s^2 \pmod{n} \right\}.$$

In contrast, the problem of taking square roots of *integral* numbers is much easier to decide and solve. The reason for this is that the natural ordering of $\mathbb{Z}$ can be used for a binary search:

Algorithm 3 Integral-Sqrt

Input: $z \in \mathbb{Z}$.
Output: $\sqrt{z}$ if $\sqrt{z} \in \mathbb{Z}$, **false** else.

1: **assert** $z \geq 0$
2: **set** $x := z$ **DIV** 2, $S := \{ x \}$
3: **while** $x^2 \neq z$ **do**
4: **set** $x := (x + (z$ **DIV** $x))$ **DIV** 2
5: **assert** $x \notin S$.
6: **set** $S := S \cup \{ x \}$
7: **end while**
8: **return** x

2.5 Finite fields

Notation 2.36. For $n \in \mathbb{Z}$ and $x \in \mathbb{F}$ we write:

$$n.x := \begin{cases} 0 & \text{if } n = 0, \\ \underbrace{1 + \ldots + 1}_{n\text{-times}} & \text{if } n > 0, \\ -(\underbrace{1 + \ldots + 1}_{n\text{-times}}) & \text{if } n < 0. \end{cases}$$

Definition 2.37. A commutative ring R with $1 \neq 0 \in R$ is called **integral domain** if it has no zero divisors i.e. $\nexists\, x, y \in R \setminus \{0\} \colon x.y = 0$.

Remark 2.38. The image of the map $\varphi \colon \mathbb{Z} \to \mathbb{F}, n \mapsto n.1$ is an integral domain. Hence $\varphi(\mathbb{Z})$ is isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/p\mathbb{Z}$ for a prime p. Its field of fractions is therefore isomorphic to $\mathbb{Q}$ or $\mathbb{Z}/p\mathbb{Z} =: \mathbb{F}_p$. In the first case we say that $\mathbb{F}$ **is of characteristic zero** and in the second case that $\mathbb{F}$ **is of characteristic** p denoted by $\mathrm{char}(\mathbb{F})$.

Lemma 2.39. *It holds that* $\sqrt{1} = \pm 1$ *in* $\mathbb{F}_q$.

Proof. Since $\mathbb{F}_q$ contains no zero-divisors and because

$$(X - 1)(X + 1) = X^2 - 1$$

we have that $X^2 - 1 = 0$ if and only $X - 1 = 0$ or $X + 1 = 0$ which is equivalent to $X = \pm 1$. □

Lemma 2.40. *For a finite field* $\mathbb{F}$ *it holds that* $p := \mathrm{char}\,(\mathbb{F})$ *is prime and* $|\mathbb{F}| = p^{[\mathbb{F}:\mathbb{F}_p]}$.

Proof. Since $\mathbb{F}$ is finite it does not contain $\mathbb{Q}$ and therefore has a prime characteristic. Let $n = \dim_{\mathbb{F}_p}(\mathbb{F}) =: [\mathbb{F}, \mathbb{F}_p]$ be the dimension of $\mathbb{F}$ over $\mathbb{F}_p$, which is finite since $\mathbb{F}$ is finite. Let $\{b_1, \ldots, b_n\} \subseteq \mathbb{F}$ be a basis. Now we can write

$$\mathbb{F} = \mathbb{F}_p\{b_1, \ldots, b_n\} := \left\{ \sum_{i=1}^{n} \lambda_i b_i \;\middle|\; \lambda_i \in \mathbb{F}_p \right\}$$

and see that $|\mathbb{F}| = p^n$. □

Lemma 2.41. *If* $\mathrm{char}(\mathbb{F}) = p$ *and* $x \in \mathbb{F}$ *it holds that*

$$\binom{p}{k}.x = \begin{cases} x & \text{if } k \in \{0, p\} \\ 0 & \text{else} \end{cases}$$

Proof.

Case 1 ($k < 0$ or $k > p$). The statement holds by definition of the binomial coefficient.

Case 2 ($k \in \{0, p\}$). We have

$$\binom{p}{k}.x = \frac{p!}{p! \cdot 1}.x = x.$$

Case 3 ($0 < k < p$)**.** Since p is prime, it divides $p!$ but not $k!$ and therefore it only divides the numerator, not the denominator. This enables us to calculate:

$$\binom{p}{k}.x = \frac{(p-1)!}{(p-k)!k!} \cdot \underbrace{p.x}_{=0} = 0. \qquad \square$$

Definition/Proposition 2.42. *The map*

$$\begin{array}{rccc} \sigma\colon & \mathbb{F} & \longrightarrow & \mathbb{F} \\ & x & \longmapsto & x^{\mathrm{char}(\mathbb{F})} \end{array}$$

is called ***Frobenius*** *and is an isomorphism.*

Proof.

Step 1 (Additivity). Let $x, y \in \mathbb{F}$, we calculate

$$\sigma(x+y) = (x+y)^p = \sum_{k=0}^{p} \binom{p}{k} x^k y^{p-k} \overset{2.41}{=} x^p + y^p = \sigma(x) + \sigma(y)$$

Step 2 (Multiplicity). Since fields are commutative, it clearly holds that

$$\sigma(xy) = (xy)^p = x^p y^p = \sigma(x)\sigma(y)$$

Step 3 (Injectivity). Let $x \in \mathbb{F}$ with $\sigma(x) = x^p = 0$. Since $\mathbb{F}$ is a fields it has no zero-divisors and $x = 0$.

Step 4 (Surjectivity). We know that σ is injective and as a field-homomorphism it therefore is surjective. $\square$

Lemma 2.43. *For every generator $y \in \mathbb{F}_q^*$ the map*

$$\begin{array}{rccc} & \mathbb{F}_q^* & \longrightarrow & \mathbb{F}_q^* \\ & x & \longmapsto & xy \end{array}$$

is an isomorphism.

Proof. Since y is a generator, we have that for every $z \in \mathbb{F}_q^*$ there exists $k_z \in \mathbb{N}$ such that $y = z^{k_z}$. We now define the map

$$\begin{array}{rccc} & \mathbb{F}_q^* & \longrightarrow & \mathbb{F}_q^* \\ & z & \longmapsto & y^{k_z - 1} \end{array}$$

which is obviously an inverse to the morphism in the statement. $\square$

Corollary 2.44. *Let $y \in \mathbb{F}_q^*$ be a generator and $k \in \mathbb{Z}$:*

$$\sum_{x \in \mathbb{F}_q^*} x^k = \sum_{x \in \mathbb{F}_q^*} (xy)^k.$$

Proof. Simply reindex the sum by the isomorphism from Lemma 2.43. $\square$

Fact 2.45. *Let q be a power of 2, then every element of $\mathbb{F}_q$ is a square.*

Proof. By Definition/Proposition 2.42 the Frobenius $x \mapsto x^2$ is an automorphism of $\mathbb{F}_q$. □

Theorem 2.46. *Let $p \in P \setminus \{2\}$ and q be a power of p, then the sequence*

$$\{1\} \longrightarrow \left[\mathbb{F}_q^*\right]^2 \xhookrightarrow{\iota} \mathbb{F}_q^* \xrightarrow{\varphi} \{\pm 1\} \longrightarrow \{1\}$$

with

$$\begin{array}{rccc} \varphi\colon & \mathbb{F}_q^* & \longrightarrow & \{\pm 1\} \\ & x & \longmapsto & x^{(q-1)/2} \end{array}$$

and the canonical inclusion ι is exact.

Proof.

Step 1 (The map φ). φ is well-defined because of the fact that $\mathbb{F}_q^*$ is cyclic of order $q-1$ and Lemma 2.39. Since $\varphi(1) = 1$, we need to show that $\exists\, y \in \mathbb{F}_q^*\colon \varphi(y) = -1$ to conclude that φ is surjective. So suppose that φ is not surjective i.e. $\forall\, x \in \mathbb{F}_q^*\colon \varphi(x) = 1$. But since $-1 \neq 1$ in $\mathbb{F}_q^*$, it holds that not every element in $\mathbb{F}_q^*$ is a square.

Step 2 (Exactness at $\mathbb{F}_q^*$:). Let Ω be an algebraic closure of $\mathbb{F}_q$. For $x \in \mathbb{F}_q^*$, let $y \in \Omega$ be such that $y^2 = x$ (such a y exists, since y is a root of $X^2 - x \in \Omega[X]$). We now claim that

$$y^{q-1} = 1 \Leftrightarrow y \in \mathbb{F}_q^*.$$

Because then

$$y \in \mathbb{F}_q^* \Leftrightarrow y^{q-1} = 1 \Leftrightarrow x^{(q-1)/2} = 1 \Leftrightarrow \varphi(x) = 1$$

which means that $x \in \ker(\varphi) \Leftrightarrow x = y^2$ for some $y \in \mathbb{F}_q^*$. To prove the claim, note that y is a fixpoint of the Frobenius and since $\mathbb{F}_q[y]$ is a finite extension of $\mathbb{F}_q$ the group $\mathrm{Aut}_{\mathbb{F}_q}(\mathbb{F}_q[y])$ is cyclic and generated by the Frobenius. So there exists only one $\mathbb{F}_q$-automorphism of $\mathbb{F}_q[y]$ meaning that $\mathbb{F}_q = \mathbb{F}_q[y]$ and $y \in \mathbb{F}_q$. □

Remark 2.47. If $q = p$, the exactness at $\mathbb{F}_p^*$ — i.e. $\mathrm{im}(\iota) = \ker(\varphi)$ — means that the numbers x where $x^{(p-1)/2} = 1$ are exactly the squares in $\mathbb{F}_p^*$.

Remark 2.48. Theorem 2.46 can be reformulated to "for $p \neq 2$ we have that $\left[\mathbb{F}_q^*\right]^2$ is a subgroup of index 2 in $\mathbb{F}_q^*$" since the short exact sequence induces an isomorphism

$$\mathbb{F}_q^* \Big/ \left[\mathbb{F}_q^*\right]^2 \cong \{\pm 1\}$$

which means that half of the elements of $\mathbb{F}_q^*$ are squares.

2.6 Legendre symbol

Definition 2.49 (Legendre symbol). For $p \in P \setminus \{2\}$ we denote the **Legendre symbol** by

$$\begin{array}{rcl} \left(\frac{\cdot}{p}\right) : \mathbb{F}_p & \longrightarrow & \{-1,0,1\} \\ x & \longmapsto & x^{\frac{p-1}{2}} \end{array}$$

Lemma 2.50 (Properties of the Legendre symbol). *Let p be an odd prime*

(i). For $x \in \mathbb{F}_p^*$, $\left(\frac{x}{p}\right) = \pm 1$.

(ii). For $x, y \in \mathbb{F}_p$: $\left(\frac{x}{p}\right)\left(\frac{y}{p}\right) = \left(\frac{xy}{p}\right)$.

(iii). For $x \in \mathbb{F}_p^*$: $\left(\frac{x}{p}\right) = 1 \Leftrightarrow x \in \left[\mathbb{F}_p^*\right]^2$

(iv). If y is a square root $x \in \mathbb{F}_p^*$ in an algebraic closure of $\mathbb{F}_p$, then $\left(\frac{x}{p}\right) = y^{p-1}$.

Remark 2.51. Since the Legendre symbol is linear in the upper argument and evaluates to 1 for squares, it can also be interpreted as a map

$$\mathbb{F}_p \Big/ \left[\mathbb{F}_p^*\right]^2 \to \{-1,0,1\}.$$

Proof.

(i). The map φ defined in Theorem 2.46 agrees with the Legende symbol on $\mathbb{F}_q^*$.

(ii). For $x = 0$ or $y = 0$ the statement is trivial and so it is for $x, y \in \mathbb{F}_p^*$:

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

(iii). Clear by Remark 2.47.

(iv). Since $y^2 = x$, one gets:

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

Definition 2.52 (Some Homomorphisms to $\mathbb{Z}/2\mathbb{Z}$). Define the maps

$$\begin{array}{rcl} \epsilon\colon (\mathbb{Z}/4\mathbb{Z})^* & \longrightarrow & \mathbb{Z}/2\mathbb{Z} \\ n & \longmapsto & \frac{n-1}{2} \end{array} \quad \text{and} \quad \begin{array}{rcl} \omega\colon (\mathbb{Z}/8\mathbb{Z})^* & \longrightarrow & \mathbb{Z}/2\mathbb{Z} \\ n & \longmapsto & \frac{n^2-1}{8}. \end{array}$$

Fact 2.53 (Values of ϵ and ω for odd integers). *It holds that*

$$\epsilon(1) = 0,\ \epsilon(-1) = -1,\ \omega(\pm 1) = \omega(\pm 3) = 0,\ \omega(\pm 5) = 1.$$

This can be reformulated as

$$\epsilon(n) = \begin{cases} 0 & \text{if } n \equiv 1 \pmod 4 \\ 1 & \text{if } n \equiv -1 \pmod 4 \end{cases} \qquad \omega(n) = \begin{cases} 0 & \text{if } n \equiv \pm 1 \pmod 8 \\ 1 & \text{if } n \equiv \pm 5 \pmod 8 \end{cases}$$

Theorem 2.54. *The following formulas hold:*

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

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

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

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

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

(iii). Let Ω be an algebraic closure of $\mathbb{F}_p$ and $\alpha \in \Omega$ be an 8th root of unity. Then the element $y = \alpha + \alpha^{-1}$ verifies $y^2 = 2$, because $\alpha^4 = -1$ implies $\alpha^2 + \alpha^{-2} = 0$. We have

$$y^p = \alpha^p + \alpha^{-p}.$$

If $p \mod \pm 1 \pmod 8$, this implies $y^p = y$, thus $\left(\frac{2}{p}\right) = y^{p-1} = 1$. If $p = \pm 5 \pmod 8$, one finds $y^p = \alpha^5 + \alpha^{-5} = -\left(\alpha + \alpha^{-1}\right) = -y$ — this again follows from $\alpha^4 = -1$. We deduce from this that $y^{p-1} = -1$, which implies the statement. □

Remark 2.55. Theorem 2.54 translates to

- 1 is a square modulo p if and only if $p \equiv 1 \pmod 4$.
- 2 is a square modulo p if and only if $p \equiv \pm 1 \pmod 8$.

Theorem 2.56. *The Legendre symbol can be calculated in linear time in the bitsize of the input.*

Proof. Let $p \in P \setminus \{2\}$ and $x \in \mathbb{F}_p$. One simply calculates $\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}}$ by square-and-multiply. Additionally reduce modulo p in every step, such that the quantity does not become too large. □

Theorem 2.57 (Quadratic reciprocity law). *For distinct $p, l \in P \setminus \{2\}$:*

$$\left(\frac{p}{l}\right) = \left(\frac{l}{p}\right)(-1)^{\epsilon(p)\,\epsilon(l)}.$$

Proof. Let Ω be an algebraic closure of $\mathbb{F}_p$ and $\omega \in \Omega$ be a primitive l-th root of unity. Now define

$$y := \sum_{x \in \mathbb{F}_l} \left(\frac{x}{l}\right) \omega^x$$

which is well defined, since by definition we have of ω that $\omega^l = 1$ and therefore that the ω^x are well defined. We will now prove two statements:

(i). $y^2 = (-1)^{\epsilon(l)}\, l$.

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

which together imply the theorem:

$$\begin{aligned}(-1)^{\epsilon(l)\,\epsilon(p)}\left(\frac{l}{p}\right) &= \left((-1)^{\epsilon(l)}\right)^{\epsilon(p)}\left(\frac{l}{p}\right) \overset{2.54}{=} \left(\frac{(-1)^{\epsilon(l)}}{p}\right)\left(\frac{l}{p}\right) \\ &\overset{2.50(\text{ii})}{=} \left(\frac{(-1)^{\epsilon(l)}\, l}{p}\right) \overset{(\text{i})}{=} \left(\frac{y^2}{p}\right) = y^{p-1} \overset{(\text{ii})}{=} \left(\frac{p}{l}\right)\end{aligned}$$

(i). First note that

$$\begin{aligned}y^2 &= \left(\sum_{x\in\mathbb{F}_l}\left(\frac{x}{l}\right)\omega^x\right)\left(\sum_{x\in\mathbb{F}_l}\left(\frac{x}{l}\right)\omega^x\right) = \sum_{x,z\in\mathbb{F}_l}\left(\frac{xz}{l}\right)\omega^{x+z} \\ &= \sum_{u\in\mathbb{F}_l}\omega^u\left(\sum_{t\in\mathbb{F}_l}\left(\frac{t(u-t)}{l}\right)\right) \\ &= \sum_{u\in\mathbb{F}_l}\omega^u\left(\sum_{t\in\mathbb{F}_l^*}\left(\frac{t(u-t)}{l}\right)\right).\end{aligned}$$

If $t \neq 0$ we have

$$\left(\frac{t(u-t)}{l}\right) \overset{2.50(\text{ii})}{=} \left(\frac{-t^2}{l}\right)\left(\frac{1-ut^{-1}}{l}\right) \overset{2.54}{=} (-1)^{\epsilon(l)}\left(\frac{1-ut^{-1}}{l}\right).$$

This two equalities imply:

$$\begin{aligned}(-1)^{\epsilon(l)}\, y^2 &= \sum_{u\in\mathbb{F}_l}\omega^u\,(-1)^{\epsilon(l)}\left(\sum_{t\in\mathbb{F}_l^*}\left(\frac{t(u-t)}{l}\right)\right) \\ &= \sum_{u\in\mathbb{F}_l}\omega^u\,(-1)^{\epsilon(l)}\left(\sum_{t\in\mathbb{F}_l^*}(-1)^{\epsilon(l)}\left(\frac{1-ut^{-1}}{l}\right)\right) \\ &= \sum_{u\in\mathbb{F}_l}\omega^u\left(\sum_{t\in\mathbb{F}_l^*}\left(\frac{1-ut^{-1}}{l}\right)\right) = \sum_{u\in\mathbb{F}_l}\omega^u C_u\end{aligned}$$

where $C_u := \sum_{t\in\mathbb{F}_l^*}\left(\frac{1-ut^{-1}}{l}\right)$. Now note that

$$C_0 = \sum_{t\in\mathbb{F}_l^*}\left(\frac{1}{l}\right) = l-1$$

and for $u \neq 0$ perform an index shift $s := 1 - ut^{-1}$ to get

$$\begin{aligned}C_u &= \sum_{t\in\mathbb{F}_l^*}\left(\frac{1-ut^{-1}}{l}\right) = \sum_{s\in\mathbb{F}_l\setminus\{1\}}\left(\frac{s}{l}\right) \\ &= \sum_{s\in\mathbb{F}_l}\left(\frac{s}{l}\right) - \left(\frac{1}{l}\right) = -\left(\frac{1}{l}\right) \overset{2.54(\text{i})}{=} -1,\end{aligned}$$

where $\sum_{s\in\mathbb{F}_l}\left(\frac{s}{l}\right) = 0$ because in $\mathbb{F}_l$ there are as many squares as non squares. Hence we calculate

$$(-1)^{\epsilon(l)}\, y^2 = \sum_{u\in\mathbb{F}_l} C_u\omega^u = C_0 + \sum_{u\in\mathbb{F}_l^*} C_u\omega^u = l - 1 - \sum_{u\in\mathbb{F}_l^*}\omega^u = l,$$

finally implying the statement $y^2 = l\,(-1)^{\epsilon(l)}$.

(ii). Since Ω is of characteristic p, we have

$$\begin{aligned} y^p &= \sum_{x\in\mathbb{F}_l}\left(\frac{x}{p}\right)\omega^{xp} \overset{2.44}{=} \sum_{z\in\mathbb{F}_l}\left(\frac{zp^{-1}}{l}\right)\omega^z \\ &\overset{2.50(\text{ii})}{=} \left(\frac{p^{-1}}{l}\right)\sum_{z\in\mathbb{F}_l}\left(\frac{z}{l}\right)\omega^z = \left(\frac{p^{-1}}{l}\right) y = \left(\frac{p}{l}\right) y, \end{aligned}$$

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

One can generalize the Legendre symbol to the Jacobi symbol that is defined not only for odd primes p but for odd integers greater than 1:

Definition 2.58 (Jacobi symbol). For $n \in \mathbb{N}_{>1}$ with prime factorization $n = \prod_{i=1}^r p_i^{r_i}$ where $p_i \in P$ and pairwise different we define the **Jacobi symbol** as the map

$$\begin{aligned} (\cdot/n) : \mathbb{Z} &\longrightarrow \{1, 0, -1\} \\ a &\longmapsto \prod_{i=1}^r \left(\frac{a}{p_i}\right)^{r_i}. \end{aligned}$$

Remark 2.59. If n is prime in the above definition, the Jacobi symbol coincides with the Legendre symbol. This is the reason why they are often both denoted by $\left(\frac{a}{n}\right)$. We will stick to the notation $\left(\frac{x}{p}\right)$ for the Legendre symbol and (a/n) for the Jacobi symbol.

Remark 2.60. The definition that reduces the Legendre symbol to the Jacobi symbol stated above is not feasible to formulate an efficient algorithm for the Jacobi symbol since the factorization of n is needed. Nevertheless, the Jacobi symbol can be calculated efficiently: The fastes known algorithm — also in terms of implementation and practice — is given in [BZ10] and runs in $\mathcal{O}(M(n)\log(n))$ where $M(n)$ denotes the time needed to multiply two n-bit numbers. Since for $n \in P$, the Jacobi symbol coincides with the Legendre symbol, this statement renders Theorem 2.56 more precisely.

2.7 Group of Norms

In linear algebra, one defines the determinant of an endomorphism as follows: For an n-dimensional $\mathbb{F}$-vector space V we can choose a basis and write an endomorphism $\varphi\colon V \to V$ as a matrix $A = (a_{ij})_{1\le i,j\le n} \in \mathbb{F}^{n\times n}$ with respect to this basis. Now the determinant of φ is defined as the determinant of this matrix, which is independent of the chosen basis.

$$\det(\varphi) := \det(A) := \sum_{\pi\in\mathcal{S}_n} \operatorname{sgn}(\pi)\prod_{i=1}^n a_{i\pi(i)}$$

where $\mathcal{S}_n$ is the group of permutations on the numbers $[n]$.

Definition 2.61 (Norm with respect to an algebraic extension). Let $\mathbb{L}/\mathbb{F}$ be a finite algebraic field extension – which means that $\mathbb{L}$ is a finite dimensional $\mathbb{F}$-vector space – and for every $\alpha \in \mathbb{L}$ define a map

$$\begin{array}{rcl} m_\alpha \colon \mathbb{L} & \longrightarrow & \mathbb{L} \\ x & \longmapsto & \alpha \cdot x \end{array}$$

which is an $\mathbb{F}$-linear endomorphism of $\mathbb{L}$. The **norm with respect to the algebraic extension** $\mathbb{L}/\mathbb{F}$ is defined as

$$N_{\mathbb{L}/\mathbb{F}}(\alpha) := \det(m_\alpha).$$

Remark 2.62. The Norm with respect to a field extension is not to be confused with the norms that are used to complete $\mathbb{Q}$ which then yield p-adic numbers $\mathbb{Q}_p$ or the real numbers $\mathbb{R}$ (see for example [Kob77]).

Fact 2.63. *A norm with respect to a field extension is multiplicative.*

Proof. This directly follows from the multiplicativity of the determinant. □

Fact 2.64. *The set $N_{\mathbb{L}/\mathbb{F}}(\mathbb{L}^*)$ is a multiplicative abelian subgroup of $\mathbb{L}$.*

Proof. For $a, b \in \mathbb{L}^*$, we have by Fact 2.63 that

$$N_{\mathbb{L}/\mathbb{F}}(a) N_{\mathbb{L}/\mathbb{F}}(b) = N_{\mathbb{L}/\mathbb{F}}(ab) \in N_{\mathbb{L}/\mathbb{F}}(L^*)$$

Furthemore we have that $1 = N_{\mathbb{L}/\mathbb{F}}(1) \in N_{\mathbb{L}/\mathbb{F}}(L^*)$ and that for any $a \in \mathbb{L}^*$ it holds that

$$N_{\mathbb{L}/\mathbb{F}}(a) N_{\mathbb{L}/\mathbb{F}}\left(a^{-1}\right) = N_{\mathbb{L}/\mathbb{F}}\left(aa^{-1}\right) = N_{\mathbb{L}/\mathbb{F}}(1) = 1,$$

which means that $N_{\mathbb{L}/\mathbb{F}}(a)^{-1} = N_{\mathbb{L}/\mathbb{F}}\left(a^{-1}\right) \in N_{\mathbb{L}/\mathbb{F}}(\mathbb{L})$. Commutativity is inherited from the muliplicative group of the field $\mathbb{L}$. □

2.8 Chevalley–Warning Theorem

Definition 2.65. We define for $n \in \mathbb{N}$ the **sum function**

$$\begin{array}{rcl} S \colon \left\{ \mathbb{F}_q^n \to \mathbb{F}_q \right\} & \longrightarrow & \mathbb{F}_q \\ f & \longmapsto & \sum_{x \in \mathbb{F}_q^n} f(x). \end{array}$$

Lemma 2.66. *For $k \in \mathbb{N}_0$ it holds that:*

$$S\left(X^k\right) = \begin{cases} -1 & \text{if } k \geq 1 \text{ and } q-1 \mid k \\ 0 & \text{else.} \end{cases}$$

Where we agree that $0^0 = 1$.

Proof. We distinguish different cases for $k \in \mathbb{N}_0$:

Case 1 ($k = 0$). In this case, all summands are equal to 1, hence $S(X^k) = q.1 = 0$ because $\mathbb{F}_q$ is of characteristic p and q is a multiple of p.

Case 2 ($k \geq 1$ and $q-1 \mid k$). We have for any $x \in \mathbb{F}_q$:

$$x^k = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x \neq 0 \end{cases}$$

hence $S\left(X^k\right) = (q-1).1 = q.1 - 1.1 = 0 - 1 = -1$.

Case 3 ($k \geq 1$ and $q-1 \nmid k$). Since $q-1$ is the order of the cyclic group $\mathbb{F}_q^*$, there exists $y \in \mathbb{F}_q^*$ such that $y^k \neq 1$. We calculate

$$S(X^k) = \sum_{x \in \mathbb{F}_q} x^k = \sum_{x \in \mathbb{F}_q^*} x^k \overset{2.44}{=} \sum_{x \in \mathbb{F}_q^*} y^k x^k = y^k S(X^k)$$

to get $(1-y^k)S(X^k) = 0$ which implies that $S(X^k) = 0$. □

Fact/Definition 2.67. Let $F \subseteq \mathbb{F}[X_1, \ldots, X_n]$ be a finite set of polynomials. Define the function

$$\chi_F := \prod_{f \in F} (1 - f^{q-1}).$$

In the case $\mathbb{F} = \mathbb{F}_q$, this is the characteristic function of the vanishing set of F:

$$\forall\, x \in \mathbb{F}_q^n \colon \chi_F(x) = \begin{cases} 1 & \text{if } x \in V(F), \\ 0 & \text{if } x \notin V(F). \end{cases}.$$

Proof. Let $x \in \mathbb{F}_q^n$ be arbitrary. If $x \in V(F)$, every f is zero at x and hence $\chi_F(x) = 1$. If $x \notin V(F)$ on the other hand, there exists an $f \in F$ s.t. $f(x) \neq 1$ and hence $(f(x))^{q-1} = 1$ implying $\chi_F(x) = 0$. □

Corollary 2.68. *For a finite set of polynomials $F \subseteq \mathbb{F}_q[X_1, \ldots, X_n]$ it holds that*

$$|V(F)| \equiv S(\chi_F) \pmod{p}.$$

Lemma 2.69. *Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ be a multiindex. It holds that*

$$\exists\, i \in [1:n] \colon S\left(X_i^{\alpha_i}\right) = 0 \quad \Longrightarrow \quad S\left(X^\alpha\right) = 0.$$

Proof.

$$\begin{aligned} S(X^\alpha) &= \sum_{(x_1,\ldots,x_n) \in \mathbb{F}_q^n} x_1^{\alpha_1} \cdots x_n^{\alpha_n} \\ &= \sum_{(x_1,\ldots,\hat{x}_i,\ldots,x_n) \in \mathbb{F}_q^{n-1}} \sum_{x_i \in \mathbb{F}_q} x_1^{\alpha_1} \cdots x_n^{\alpha_n} \\ &= \sum_{(x_1,\ldots,\hat{x}_i,\ldots,x_n) \in \mathbb{F}_q^{n-1}} x_1^{\alpha_1} \cdots \widehat{x_i^{\alpha_i}} \cdots x_n^{\alpha_n} \underbrace{\sum_{x_i \in \mathbb{F}_q} x_i^{\alpha_i}}_{\overset{2.66}{=} 0} = 0 \end{aligned}$$

□

Theorem 2.70 (Chevalley–Warning). *For any $F \subseteq \mathbb{F}_q[X_1, \ldots, X_n]$ we have*

$$\sum_{f \in F} \deg(f) < n \quad \Rightarrow \quad |V(F)| \equiv 0 \pmod{p}.$$

Proof. First notice that the condition $\sum_{f\in F} \deg(f) < n$ implies that F is finite, since $\mathbb{F}_q$ is finite and there are only finitely many degree bounded polynomials. By Corollary 2.68 it is left to show that

$$S(\chi_F) \equiv 0 \pmod{p}.$$

For the degree of χ_F we calculate:

$$\deg(\chi_F) \leq \sum_{f\in F} (q-1) \cdot \deg(f) = (q-1) \sum_{f \in F} \deg(f) < (q-1)n.$$

Thus χ_F is a linear combination of monomials:

$$\chi_F = \sum_{\alpha \in [\mathbb{N}]^{<n(q-1)}} a_\alpha X^\alpha \quad \text{for } a_\alpha \in \mathbb{F}.$$

By the pigeonhole principle, we therefore have that there is at least one $i \in [1:n]$ such that $\alpha_i < q-1$. Lemma 2.69 consequently implies that $S(X^\alpha) = 0$. Now calculate

$$\begin{aligned} S(\chi_F) &= \sum_{x \in \mathbb{F}^n} \sum_{\alpha \in [\mathbb{N}]^{<n(q-1)}} a_\alpha x^\alpha \\ &= \sum_{\alpha \in [\mathbb{N}]^{<n(q-1)}} a_\alpha \sum_{x \in \mathbb{F}^n} x^\alpha \\ &= \sum_{\alpha \in [\mathbb{N}]^{<n(q-1)}} a_\alpha \underbrace{S(X^\alpha)}_{\overset{2.69}{=} 0} = 0. \end{aligned}$$

□

Corollary 2.71. *Let $F \subseteq \mathbb{F}_q[X_1, \ldots, X_n]$ with $\sum_{f\in F} \deg(f) < n$ and no $f \in F$ have a constant term, then there exists a nontrivial common zero, i.e.*

$$\exists\, x \in V(F) \setminus \{\, 0 \,\}.$$

Proof. Since $1 = |\{\, 0 \,\}|$ is not divisible by p, we have $|V(F)| \neq |\{\, 0 \,\}|$. □

Corollary 2.72. *All quadratic forms over $\mathbb{F}_q$ in at least 3 variables have a nontrivial zero.*

Remark 2.73. In geometric language Corollary 2.72 means that every conic over a finite field has a rational point.

2.9 p-adic numbers

Let p be a prime number in this section. We will introduce the ring $\mathbb{Z}_p$ of so called p-adic integers and the field $\mathbb{Q}_p$ of so called p-adic numbers. They are widely used in number theory and can be thought of as the attempt to

calculate in $\mathbb{Z}/p\mathbb{Z}$ but without losing information. We will also make use of them to define the Hilbert symbol in Chapter 3. One can define p-adic numbers algebraically — as an inverse limit — or analytically — as the completion of $\mathbb{Q}$ with respect to a so-called non-Archimedean norm. In this chapter, we will give the algebraic approach, for a more detailed one that constructs them in analogy to the construction of $\mathbb{R}$ from $\mathbb{Q}$, see [Kob77].

Definition 2.74 (Inverse System). Let $(I, \leq)$ be a directed partially ordered set and $(A_i)_{i\in I}$ be a family of rings. Furthermore let $f_{ij} : A_j \to A_i$ for all $i \leq j$ be a family of homomorphisms with the following properties:

(i). $\forall\, i \in I : f_{ii} = \mathrm{id}_{A_i}$,

(ii). $\forall\, i \in I$ with $i \leq j \leq k : f_{ik} = f_{ij} \circ f_{jk}$.

The pair $((A_i)_{i\in I}, (f_{ij})_{i\leq j\in I})$ is called an **inverse system of rings and morphisms over** I (or shorter an **inverse system over** I) and the morphisms f_{ij} are called transition morphisms of the system.

Remark 2.75. An inverse systems/limit is often also called **projective system/limit**.

Definition 2.76 (Inverse Limit). Let $((A_i)_{i\in I}, (f_{ij})_{i\leq j\in I})$ be an inverse system. The following subring of the direct product is called **inverse limit**:

$$\varprojlim_{i\in I}(A_i) := \left\{ a \in \prod_{i\in I} A_i \;\middle|\; a_i = f_{ij}(a_j)\ \forall\, i \leq j \right\}$$

where a_i denotes the i-th component of $a \in \prod_{j\in I} A_j$.

Lemma 2.77. *Let (D_n, f_n) be a inverse system over $\mathbb{N}$. Then we have*

$$\forall\, n \in \mathbb{N}\colon D_n \neq \varnothing \text{ and } |D_n| < \infty \quad \Longrightarrow \quad \varprojlim_{i\in\mathbb{N}}(D_i) \neq \varnothing.$$

Proof. For ease of notation set

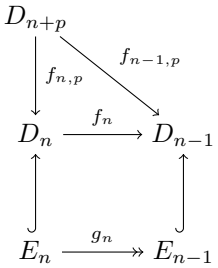
$$D := \varprojlim_{i\in\mathbb{N}}(D_i),$$

$$\forall\, n, p \in \mathbb{N}\colon \quad f_{n,p} := f_{n+1} \circ \cdots \circ f_{n+p} \quad \text{and} \quad D_{n,p} := f_{n,p}(D_{n+p})$$

First we observe the statement for the case that the $f_n : D_n \to D_{n-1}$ are surjective: Since D_1 is non-empty we can choose $d_1 \in D_1$ and since f_2 is surjective, we can find a pre-image in D_2. This process can be continued inductively to obtain an element in D implying that $D \neq \varnothing$.

We now reduce the lemma to this special case. For a fixed n, the $D_{n,p}$ form a decreasing family of finite non-empty subsets. Since $D_{n,p}$ as a subset of a finite set is finite, the sequence of natural numbers $|D_{n,p}|$ is decreasing. But the $D_{n,p}$ are non-empty and this sequence is therefore bounded by 1, hence convergent. So for any $n \in \mathbb{N}$, this means that there exists an $N_n \in \mathbb{N}$ such that for every $p \in \mathbb{N}_{\geq N_n}$ we have $D_{n,p} = D_{n,p+1}$. Set $p := \max\{ N_n, N_{n-1} \}$ and $E_n := D_{n,N_n}$ — which is non-empty — and note that $E_n \subseteq D_n$. Now define $g_n := f_n|_{E_n}$

as the restriction of f_n to this limit set E_n to get the following commutative diagram



The surjectivity of g_n follows by the commutativity:

$$E_{n-1} = D_{n-1,p} = f_{n-1,p}\left(D_{n+p}\right) = f_n\left(f_{n,p}\left(D_{n+p}\right)\right) = f_n\left(E_n\right) = g_n\left(E_n\right).$$

By the remark made at the beginning, the limit E of (E_n, g_n) is non-empty. Finally by the functoriality of the inverse limit we have that $E \subseteq D$. This ultimatly implies that D is non-empty. □

Fact/Definition 2.78 (p-adic integers). For $n \geq 1$ define $A_n := \mathbb{Z}/p^n\mathbb{Z}$ and denote the canonical projection by $\pi_n : A_n \to A_{n-1}$. The sequence

$$\ldots \longrightarrow\!\!\!\!\!\rightarrow A_n \xrightarrow{\pi_n} A_{n-1} \xrightarrow{\pi_{n-1}} \ldots \xrightarrow{\pi_3} A_2 \xrightarrow{\pi_2} A_1$$

forms a inverse system over $\mathbb{N}$ (silently add the identities and all compositions). Its' inverse limit is called the set of **p-adic integers** which is denoted by $\mathbb{Z}_p$. Every element $x \in \mathbb{Z}_p$ is a sequence

$$x = (x_n)_{n\in\mathbb{N}} = (x_1, x_2, \ldots) \quad \text{with} \quad \begin{array}{ll} x_n \in A_n = \mathbb{Z}/p^n\mathbb{Z} & \forall\, n \in \mathbb{N} \text{ and} \\ \pi_n\left(x_n\right) = x_{n-1} & \forall\, n \geq 2. \end{array}$$

Note that $\mathbb{Z}_p$ equiped with a coordinate-wise addition and multiplication becomes a ring, i.e. a subring of the product $\prod_{n\in\mathbb{N}} A_n$. Every integer $m \in \mathbb{N}$ defines such a sequence $(x_n)_{n\in\mathbb{N}}$ by setting for every $n \in \mathbb{N}$:

$$x_n \equiv m \pmod{p^n}.$$

This makes $\mathbb{Z}$ a subring of $\mathbb{Z}_p$. If we want to stress the difference, we will call $\mathbb{Z}$ *rational integers* in contrary to p-adic integers. The sum of two p-adic integers that are defined by rational integers is the p-adic integer that is defined by the sum of the two rational integers. Furthemore equip A_n with the discrete topology and note that $\mathbb{Z}_p \subseteq \prod_{n\in\mathbb{N}} A_n$ is compact.

Proof. First we note that the system above is a inverse system by construction: We silently added the identities so (i) in Definition 2.74 is fullfilled; similarly (ii) is true since we also added all compositions of π_n which trivially satisfy this property. The coordinates-wise addition and multiplication on $\mathbb{Z}_p$ makes it a ring because these operations commute with the modulo operation. Finally $\mathbb{Z}_p$ is closed in a product of compact spaces and is therefore compact. □

Example 2.79. The rational integer 35 defines the 2-adic integer

$$(1, 3, 3, 3, 3, 35, 35, \ldots)$$

and the 2-adic defined by 77 is

$$(1, 1, 5, 13, 13, 13, 77, 77, \ldots).$$

The sum of this two 2-adic integers is

$$\begin{array}{rcccccccccc} & (1, & 3, & 3, & 33, & 35, & 35, & 35, & 35, & \ldots) \\ + & (1, & 1, & 5, & 13, & 13, & 13, & 77, & 77, & \ldots) \\ = & (0, & 0, & 0, & 0, & 16, & 48, & 112, & 112, & \ldots) \end{array}$$

which is exactly the 2-adic defined by 112.

Example 2.80. The 3-adics defined by 35 and 77 are

$$(2, 8, 8, 35, 35, \ldots) \quad \text{and} \quad (2, 5, 23, 77, 77, \ldots).$$

Fact 2.81.

(i). $\pi_n\left(p^k\right) = p^k$ *for* $n, k \in \mathbb{N}$ *with* $n > 1$, $k \geq 0$.

(ii). $\ker(\pi_n) = p^{n-1}A_n$ *for* $n \in \mathbb{N}$ *with* $n > 1$.

Proof.

(i). This is clear.

(ii). One calculates

$$\begin{aligned} \ker(\pi_n) &:= \{\, x \in \mathbb{Z}/p^n\mathbb{Z} \mid \pi_n(x) = 0 \,\} \\ &= \left\{\, p^{n-1}y \;\middle|\; y \in \mathbb{Z}/p^n\mathbb{Z} \,\right\} = p^{n-1}A_n. \end{aligned}$$ □

Proposition 2.82. *For every* $n \in \mathbb{N}$, *the sequence of abelian groups*

$$0 \longrightarrow \mathbb{Z}_p \xrightarrow{p^n} \mathbb{Z}_p \xrightarrow{\epsilon_n} A_n \longrightarrow 0$$

where

$$\begin{array}{rcl} p^n : \mathbb{Z}_p & \longrightarrow & \mathbb{Z}_p \\ x & \longmapsto & p^n x \end{array} \quad \textit{and} \quad \begin{array}{rcl} \epsilon_n : \mathbb{Z}_p & \longrightarrow & A_n \\ (x_1, x_2, \ldots) & \longmapsto & x_n \end{array}$$

is exact.

Proof. The exactness is equivalent to the following four statements:

Step 1 ($\ker(p^n) = 0$). Let $x = (x_1, x_2, \ldots) \in \mathbb{Z}_p$ with $px = 0$ or, in other words, $\forall\, k \in \mathbb{N}$: $px_{k+1} = 0$. This means that x_{k+1} is of the form $p^k y_{k+1}$ with $y_{k+1} \in A_{k+1}$. We then calculate for $k \in \mathbb{N}$:

$$x_k = \pi_{k+1}(x_{k+1}) = \pi_{k+1}\left(p^{k-1}y_{k+1}\right) \overset{2.81(i)}{=} p^{k-1}\pi_{k+1}\left(y^{k+1}\right) = 0.$$

Consequently, $px = 0$ implies that $x = 0$ and therefore $p^n x = 0$ implies that $x = 0$.

Step 2 ($\operatorname{im}(p^n) \subseteq \ker(\epsilon_n)$)**.** Obviously $\epsilon_n(p^n\mathbb{Z}_p) = \{0\}$.

Step 3 ($\ker(\epsilon_n) \subseteq \operatorname{im}(p^n)$)**.** Let $x = (x_1, x_2, \ldots) \in \ker(\epsilon_n)$ i.e. $\epsilon_n(x) = 0$ which implies

$$\forall\, m \geq n\colon x_m \equiv 0 \pmod{p^n}.$$

We can now take for every $m \geq n$ the preimage $y_{m-n} \in A_{m-n}$ of x_m under the isomorphism $A_{m-n} \to p^n\mathbb{Z}/p^m\mathbb{Z} \subseteq A_m$. This means that $x_m = p^n y_{m-n}$. The y_{m-n} define a sequence $y := (y_1, y_2, \ldots)$ which is an element of $\mathbb{Z}_p$ and one calculates $p^n y = x$

Step 4 ($\ker(\epsilon_n) \subseteq \operatorname{im}(p^n)$)**.** ($\operatorname{im}(\epsilon_n) = A_n$): ϵ_n is defined as the projection on the n-th component, which is exactly A_n. □

Corollary 2.83. *For every $n \in \mathbb{N}$ it holds that*

$$\mathbb{Z}/p^n\mathbb{Z} \cong \mathbb{Z}_p/p^n\mathbb{Z}_p$$

Lemma 2.84. $\forall\, x \in \mathbb{Z}_p\colon x \in \mathbb{Z}_p^* \Leftrightarrow p \nmid x$.

Proof. We will show the statement for A_n which will imply the result for $\mathbb{Z}_p$. So let $\pi := \pi_n \circ \cdots \circ \pi_2 : A_n \to A_1 = \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ be the projection and $x \in A_n = \mathbb{Z}/p^n\mathbb{Z}$.

First assume that $x \in A_n^*$. This means that $\exists\, y \in A_n\colon xy = 1$ which still holds after applying π. So $\pi(x)$ and $\pi(y)$ are invertible in $\mathbb{F}_p$ and therefore x and y are not divisible by p.

Now assume that $p \nmid x$. This means $x \notin pA_n$ and therefore $x \neq 0$ in $\mathbb{F}_p$ or $\pi(x) \in \mathbb{F}_p^*$. This means that $\exists\, b \in A_n$ with $x = \pi(x) + bp$. Now choose a preimage y of x^{-1} under π and calculate:

$$\begin{array}{rrcl} & x & = & \pi(x) + bp \\ \Leftrightarrow & xy & = & 1 + bpy \\ \Leftrightarrow & xy & = & 1 - pz \end{array}$$

for $z := -yb$. One has

$$(1 - pz)(1 + pz + \ldots + p^{n-1}z^{n-1}) = 1$$

which proves that $y(1 + pz + \ldots + p^{n-1}z^{n-1})$ is the inverse of x in A_n. □

Definition 2.85. We call the invertible elements of $\mathbb{Z}_p$ **p-adic units**.

Proposition 2.86 (Decomposition in $\mathbb{Z}_p$). $\forall\, x \in \mathbb{Z}_p \setminus \{0\}$:

$$\exists!\, u \in \mathbb{Z}_p^*, \exists!\, n \geq 0 : \quad x = p^n u$$

Proof. If $x \in \mathbb{Z}_p$ is not zero, there exists a large n s.t. $\epsilon_n(x) \neq 0$. Now let n be minimal with this property and write $x = p^n u$ where u is not divisible by p. By Lemma 2.84 this implies $u \in \mathbb{Z}_p^*$. The uniqueness of the decomposition is obvious. □

Notation 2.87. Let $x \in \mathbb{Z}_p \setminus \{0\}$ and write it in the form $p^n u$ with $u \in \mathbb{Z}_p^*$. The integer n is called the **p-adic valuation** of x and we denote it by $\operatorname{ord}_p(x)$. We put $\operatorname{ord}_p(0) = \infty$ and can (symbolically) write

$$\operatorname{ord}_p(xy) = \operatorname{ord}_p(x) + \operatorname{ord}_p(y) \quad \text{and} \quad \operatorname{ord}_p(x+y) \geq \inf\left(\operatorname{ord}_p(x), \operatorname{ord}_p(y)\right).$$

Remark 2.88. One can use the p-adic valuation to define a norm that satsifies a stronger version of the triangle inequality: $|x+y| \leq \max\{|x|, |y|\}$. Such norms are called non-Archimedean. Ostrowski's theorem states that — up to equivalence of norms — the only norms on $\mathbb{Q}$ are the absolute value and such non-Archimedean norms.

Fact 2.89. *For the p-adic valuation it holds that for all $x \in \mathbb{Z}_p$:*

$$x \in p^n\mathbb{Z}_p \Leftrightarrow \mathrm{ord}_p(x) \geq n.$$

Proof. The statement is trivial for $x = 0$ so let $x \neq 0$ and decompose $x = p^k u$ as in Proposition 2.86 (i.e. $k = \mathrm{ord}_p(x)$ and $u \in \mathbb{Z}_p^*$). Now apparently $x \in p^n\mathbb{Z}_p$ is equivalent to $k \geq n$. □

Corollary 2.90. *$\mathbb{Z}_p$ is an integral domain.*

Proof. Suppose that $\mathbb{Z}_p$ is not an integral domain, then there exist $x, y \in \mathbb{Z}_p \setminus \{0\}$ with $xy = 0$. With the formula from Notation 2.87 we calculate

$$\infty = \mathrm{ord}_p(0) = \mathrm{ord}_p(xy) = \mathrm{ord}_p(x) + \mathrm{ord}_p(y)$$

but since $x, y \neq 0$ we know that $\mathrm{ord}_p(x), \mathrm{ord}_p(y) \neq \infty$ which is a contradiction.□

Proposition 2.91. *The topology on $\mathbb{Z}_p$ can be defined by the distance*

$$d(x, y) := p^{-\mathrm{ord}_p(x-y)}.$$

The ring $\mathbb{Z}_p$ then becomes a complete metric space in which $\mathbb{Z}$ is dense.

Proof. This follows from the equivalent definition of p-adic numbers as a completion of $\mathbb{Q}$ with respect to the p-adic norm. Details can for example be found in [Kob77]. □

Definition 2.92 (p-adic numbers). The field of fractions of $\mathbb{Z}_p$ is called the field of p-adic numbers and denoted by $\mathbb{Q}_p$.

Fact 2.93. *In analogy to Proposition 2.86 every element $x \in \mathbb{Q}_p^*$ can be written uniquely as $x = p^n u$ with $n \in \mathbb{Z}$ and $u \in \mathbb{Z}_p^*$. We also extend the p-adic valuation defined in Notation 2.87 to $\mathbb{Q}_p^*$. One has $\mathrm{ord}_p(x) \geq 0$ if and only if $x \in \mathbb{Z}_p$.*

Proposition 2.94. *The field $\mathbb{Z}_p$ together with the topology defined by*

$$\begin{array}{rccc} d\colon & \mathbb{Q}_p \times \mathbb{Q}_p & \longrightarrow & \mathbb{R} \\ & (x, y) & \longmapsto & p^{-\mathrm{ord}_p(x-y)} \end{array}$$

is locally compact and contains $\mathbb{Z}_p$ as an open subring. The field $\mathbb{Q}$ is dense in $\mathbb{Q}_p$.

Proof. Again, this can be seen easily if one defines p-adic numbers as it is done in [Kob77]. □

Notation 2.95. For a polynomial $f \in \mathbb{Z}_p[X_1, \ldots, X_m]$ and $n \in \mathbb{N}$, we denote by f_n the polynomial with coefficients in A_n (see Fact/Definition 2.78) obtained by reduction modulo p^n and Corollary 2.83.

Proposition 2.96. *For* $\left\{ f^{(i)} \right\}_{i\in I} \subseteq \mathbb{Z}_p[X_1,\ldots,X_m]$ *it holds that*

$$V\left(\left\{ f^{(i)} \right\}_{i\in I}\right) \neq \varnothing \quad \Longleftrightarrow \quad \forall\, n \in \mathbb{N}_{>1} \colon V\left(\left\{ f_n^{(i)} \right\}_{i\in I}\right) \neq \varnothing.$$

Proof. Define

$$D := V\left(\left\{ f^{(i)} \right\}\right) \quad \text{and} \quad \forall\, n \in \mathbb{N}_{>1} \colon D_n := V\left(\left\{ f_n^{(i)} \right\}_{i\in I}\right).$$

and note that D_n is finite since A_n is finite. By the functoriality of the inverse limit, we have $D = \varprojlim(D_n)$. The statement then follows from Lemma 2.77. □

Fact/Definition 2.97. $x = (x_1,\ldots,x_m) \in \mathbb{Z}_p^m$ is called primitive if one of the x_i is invertible, that is, if the x_i are not all divisible by p.

Proof. Obvious by Lemma 2.84. □

Proposition 2.98. *Let* $\left\{ f^{(i)} \right\}_{i\in I} \subseteq \mathbb{Z}_p[X_1,\ldots,X_m]$ *be a set of homogeneous polynomials. Then the following statements are equivalent:*

(i). The $f^{(i)}$ *have a nontrivial common zero in* $\mathbb{Q}_p^m$.

(ii). The $f^{(i)}$ *have a common primitive zero in* $\mathbb{Z}_p^m$.

(iii). $\forall\, n \in \mathbb{N}_{>1}$ *: the* $f_n^{(i)}$ *have a common primitive zero in* A_n^m.

Proof.

Step 1 ($(ii) \Rightarrow (i)$**).** Primitive zeros in $\mathbb{Z}_p^m$ are nontrivial in $\mathbb{Q}_p^m$.

Step 2 ($(i) \Rightarrow (ii)$**).** Let $x = (x_1,\ldots,x_m)$ be a nontrivial common zero of the $f^{(i)}$, put

$$h := \min\left(\operatorname{ord}_p(x_1),\ldots,\operatorname{ord}_p(x_m)\right) \text{ and } y = p^{-h}x.$$

Now $y \in \mathbb{Z}_p^m$ and is a primitive zero of all the $f^{(i)}$.

Step 3 ($(ii) \Leftrightarrow (iii)$**).** Follows from Proposition 2.96. □

2.10 Hensel's lemma

In this section, we will prove some lemmas that enable us to "lift" a solution of a polynomial equation modulo some prime p to a unique solution modulo some higher power of p. We start with a version for $\mathbb{Z}/p\mathbb{Z}$ and continue with versions for p-adic integers also covering the multivariate case.

Lemma 2.99 (Hensel's lemma for modular arithmetic). *Let* $f \in \mathbb{Z}[X]$ *and* $k, m \in \mathbb{Z}_{>0}$ *such that* $k \leq m$. *For any* $x \in \mathbb{Z}$ *with*

$$f(x) \equiv 0 \pmod{p^m} \quad \textit{and} \quad f'(x) \not\equiv 0 \pmod{p}$$

there exists $y \in \mathbb{Z}$ *with* $y \equiv x \pmod{p^m}$ *such that*

$$f(y) \equiv 0 \pmod{p^{m+k}}.$$

Furthermore y *is unique modulo* p^{m+k} *and can be computed by the formula*

$$y = x + tp^m \quad \text{where } t = -\frac{f(x)}{p^m}\left(f'(x)^{-1}\right).$$

In this formula, $f(x)$ *is divisible by* p^m*, so the fraction is an ordinary integer division. All other operations — addition, multiplication and additional and multiplicative inversion — are meant as operations performed in* $\mathbb{Z}/p^n\mathbb{Z}$*.*

Proof. Let $t \in \mathbb{Z}$ be arbitrary, set $y = x + tp^m$ and consider the Taylor expansion of f around x:

$$f\left(x + tp^m\right) = f(x) + tp^n f'(x) + \mathcal{O}\left(p^{2m}\right).$$

Since we want $f(y) \equiv 0 \pmod{p^{m+k}}$ we get

$$0 \equiv f\left(x + tp^m\right) \equiv f(x) + tp^n f'(x) \pmod{p^{m+k}}$$

where $\mathcal{O}\left(p^{2m}\right)$ vanishes because $m + k \leq 2m$. Then we note that for some $z \in \mathbb{Z}$ we have $f(x) = zp^m$ and get

$$0 \equiv f(y) \equiv \left(z + tf'(x)\right)p^m \pmod{p^{m+k}} \quad \Longleftrightarrow \quad 0 \equiv z + tf'(x) \pmod{p^k}.$$

We now substitute $z = \frac{f(x)}{p^m}$ and solve the equation for t which is possible since $f'(x)$ is not divisible by p and therefore has an unique inverse in $\mathbb{Z}/p^k\mathbb{Z}$. This implies the existence and uniqueness of y and finishes the proof. □

Lemma 2.100 (Hensel's lemma for p-adic numbers). *Let* $f \in \mathbb{Z}_p[X]$ *and* $m, k \in \mathbb{Z}$ *such that* $0 \leq 2k < m$*. Then for any* $x \in \mathbb{Z}_p$ *with*

$$f(x) \equiv 0 \pmod{p^m} \quad \text{and} \quad \operatorname{ord}_p\left(f'(x)\right) = k$$

there exists $y \in \mathbb{Z}_p$ *with* $y \equiv x \pmod{p^{m-k}}$ *such that*

$$f(y) \equiv 0 \pmod{p^{m+1}} \quad \text{and} \quad \operatorname{ord}_p\left(f'(y)\right) = k.$$

Proof. Let $t \in \mathbb{Z}_p$ be arbitrary, set $y = x + tp^{m-k}$ and consider the Taylor expansion of f around x:

$$f\left(x + tp^{m-k}\right) = f(x) + tp^{m-k}f'(x) + \mathcal{O}(p^{2(m-k)}).$$

Since we want $f(y) \equiv 0 \pmod{p^{m+k}}$ we get

$$0 \equiv f(y) \equiv f\left(x + tp^{m-k}\right) \equiv f(x) + tp^{m-k}f'(x) \pmod{p^{m+1}}$$

where $\mathcal{O}(p^{2(m-k)})$ vanishes because $2k < m \Leftrightarrow 2(m-k) \geq m+1$. By the assumptions we have that there exists $z \in \mathbb{Z}_p$ and $c \in \mathbb{Z}_p^*$ such that $f(x) = zp^m$ and $f'(x) = cp^k$. This gives

$$0 \equiv zp^m + tp^{m-k}cp^k \equiv p^m(z + tc) \pmod{p^{m+1}} \quad \Leftrightarrow \quad 0 \equiv z + tc \pmod{p}.$$

Since c is invertible in $\mathbb{Z}_p$, we can choose $t = -zc^{-1}$ implying that

$$y = x - zc^{-1}p^{m-k}.$$

which satisfies $f(y) \equiv 0 \pmod{p^{m+1}}$. Now apply Taylor's formula to f' at y to obtain

$$f'(y) = f'\left(x - zc^{-1}p^{m-k}\right) = f'(x) + zc^{-1}p^{m-k}f''(x) + \mathcal{O}(p^{2(m-k)}).$$

Reduction modulo p^{m-k} yields that $f'(y) \equiv cp^k \pmod{p^{m-k}}$. Since $m > 2k \Leftrightarrow m - k > k$ we get $\operatorname{ord}_p(f'(y)) = k$. □

Theorem 2.101 (Hensel's multivariate lemma for p-adic numbers). *Let $f \in \mathbb{Z}_p[X_1, \ldots, X_n]$ and $m, k \in \mathbb{Z}, j \in [0:n]$ such that $0 < 2k < m$. Then for any $x \in \mathbb{Z}_p^m$ with*

$$f(x) \equiv 0 \pmod{p^m} \quad \text{and} \quad \operatorname{ord}_p\left(\frac{\partial f}{\partial X_j}(x)\right) = k$$

there exists $y \in \mathbb{Z}_p^m$ with $y \equiv x \pmod{p^{m-k}}$ such that $f(y) = 0$ in $\mathbb{Z}_p$.

Proof. We first cover the univariat case $n = 1$. Apply the above lemma to $x^{(0)} = x$ to obtain $x^{(1)} \in \mathbb{Z}_p$ with $x^{(1)} \equiv x^{(0)} \pmod{p^{n-k}}$ with

$$f\left(x^{(1)}\right) \equiv 0 \pmod{p^{n+1}} \quad \text{and} \quad \operatorname{ord}_p\left(f'\left(x^{(1)}\right)\right) = k.$$

We can repeat this process with $x^{(1)}$ after replacing n by $n + 1$. Arguing inducitvely, we construct a sequence $\left(x^{(i)}\right)_{i \in \mathbb{N}_0}$ such that for every $i \in \mathbb{N}_0$:

$$x^{(i+1)} \equiv x^{(i)} \pmod{p^{n+i-k}} \quad \text{and} \quad f\left(x^{(i)}\right) \equiv 0 \pmod{p^{n+i}}.$$

Since this is a Cauchy sequence we can call its limit y. We have $f(y) = 0$ and $y \equiv x \pmod{p^{n-k}}$, hence the statement for $m = 1$.

Now let $n > 1$. We are going to reduce this case to $n = 1$ by modifying only x_j: Let $g \in \mathbb{Z}_p[X_j]$ obtained from f by plugging in x_i for X_i for every $i \in [0:n] \setminus \{j\}$. The above now gives us a $y_j \in \mathbb{Z}_p$ with $y_j = x_j \pmod{p^{n-k}}$ and $g(y_j) = 0$. Now put $y_i = x_i$ for $i \in [0:n] \setminus \{j\}$. The element $y = (y_i)_{i=1}^n$ now satisfies the desired property. □

The following corollary is the reason Hensel's lemma is often refered to as Hensel's *lifting* lemma:

Corollary 2.102. *Simple zeros modulo p of a polynomial lift to zeros of f with coefficients in $\mathbb{Z}_p$ (simple zeros are zeros where at least one partial derivative is nonzero).*

Proof. This is the special case $m = 1$ and $k = 0$ of Theorem 2.101. □

Corollary 2.103. *Let $a \in \mathbb{Z}_p, f(X) = \sum_{i,j} a_{ij}X_iX_j \in \mathbb{Z}_p[X_1, \ldots, X_n]$ such that $\forall\, i, j\colon a_{ij} = a_{ji}$ or, in other words, the matrix $A := (a_{ij})_{ij}$ is symmetric. It holds that:*

(i). If $p \neq 2$ and $\det(A) \in \mathbb{Z}_p^*$: Every primitive solution $x \in \mathbb{Z}_p^n$ of the equation $f(x) \equiv a \pmod{p}$ lifts to a true solution.

(ii). If $p = 2$: Every primitive solution $x \in \mathbb{Z}_2^n$ of the equation $f(x) \equiv a \pmod{8}$ that does not annihilate all partial derivatives $\frac{\partial f}{\partial X_j}$ modulo 4 lifts to a true solution. This condition is fulfilled if $\det(A)$ is invertible.

$\det(A)$ *is often called the* discriminant of f.

Proof. Since $a_{ij} = a_{ji}$ we have

$$\forall\, k \in [1:n]\colon \frac{\partial f}{\partial X_k} = 2\sum_{j=1}^{n} a_{kj} X_j = 2\left(A \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}\right)_k .$$

As $x = (x_1, \ldots, x_n)$ is primitive, there exists k such that $x_k \not\equiv 0 \pmod{p}$. Since $\det(A) \not\equiv 0 \pmod{p}$ we see that

$$m_k := \left(A \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}\right)_k \not\equiv 0 \pmod{p}.$$

We know that $2m_k = \frac{\partial f}{\partial X_k}$ and distinguish the following cases

Case 1 ($p \neq 2$). As we have $2m_k \not\equiv 0 \pmod{p}$, x does not annhiliate all the partial derivatives of f modulo p and Corollary 2.102 yields the result.

Case 2 ($p = 2$). From $m_k \not\equiv 0 \pmod{2}$ it follows that $2m_k \not\equiv 0 \pmod{4}$ and Theorem 2.101 with $m = 3$ and $k = 1$ yields the statement. □

2.11 The multiplicative group of $\mathbb{Q}_p$

Define for every $n \in \mathbb{N}$

$$U_n := 1 + p^n \mathbb{Z}_p.$$

and

$$\begin{array}{rcl} \varepsilon_n\colon \mathbb{Z}_p^* & \longrightarrow & (\mathbb{Z}/p^n\mathbb{Z})^* \\ (\ldots, x_3, x_2, x_1) & \longmapsto & x_n. \end{array}$$

Observe that $U_n = \ker(\varepsilon_n)$. In particular we have $\mathbb{Z}_p^*/U_1 \cong \mathbb{F}_p^*$, hence it is cyclic of order $p-1$. $(U_n)_{n\in\mathbb{N}}$ is a decreasing sequence of open subgroups of $\mathbb{Z}_p^*$ and

$$\mathbb{Z}_p^* = \varprojlim_{n\in\mathbb{N}_{n\geq 1}} \left(\mathbb{Z}_p^*/U_n\right).$$

For $n \geq 1$, the map $1 + p^n x \mapsto x \pmod{p}$ defines an isomorphism $U_n/U_{n+1} \to \mathbb{Z}/p\mathbb{Z}$ by the following formula:

$$(1 + p^n x)(1 + p^n y) \equiv 1 + p^n x + y \pmod{p^{n+1}}.$$

Induction on n then yields that $|U_1/U_n| = p^{n-1}$.

Lemma 2.104. *Let $0 \to A \to E \to B \to 0$ be an exact sequence of commutative groups (denoted additively) with A and B finite of orders a and b prime to each other. Let $B' := \{ x \in E \mid bx = 0 \}$. Then $E = A \oplus B'$. Moreover B' is the only subgroup of E isomorphic to B.*

Proof. Since a and b are relatively prime to each other Bézout's Lemma yields that there exist $r, s \in \mathbb{Z}$ with $ar + bs = 1$. If $x \in A \cap B'$, we have $ax = bx = 0$, hence $(ar + bs)x = x = 0$ and $A \cap B' = \{ 0 \}$. Moreover all $x \in E$ can be written as $x = arx + bsx$, since $bB' = \{ 0 \}$, we have $bE \subseteq A$, hence $bsx \in A$. On the other hand, from $abE = \{ 0 \}$ follows that $arx \in B'$. Hence we see that $E = A \oplus B'$ and the projection $E \to B$ defines an isomorphism of B' onto B. Conversely, if B'' is a subgroup of E isomorphic to B, we have $bB'' = \{ 0 \}$ hence $B'' \subseteq B'$ and $B'' = B'$ since they have the same order. □

Proposition 2.105. *We have $\mathbb{Z}_p^* = V \times U_1$ where $V = \left\{ x \in \mathbb{Z}_p^* \;\middle|\; x^{p-1} = 1 \right\}$ is the unique subgroup of $\mathbb{Z}_p^*$ isomorphic to $\mathbb{F}_p^*$.*

Proof. Observe that the following sequence is exact

$$1 \to U_1/U_n \to \mathbb{Z}_p^*/U_n \to \mathbb{F}_p^* \to 1,$$

and furthermore that $|U_1/U_n| = p^{n-1}$ and $\left|\mathbb{F}_p^*\right| = p - 1$. Then Lemma 2.104 yields that $\mathbb{Z}_p^*/U_n$ containes a unique subgroup V_n with $V_n \cong \mathbb{F}_p^*$ and the projection

$$\mathbb{Z}_p^*/U_n \to \mathbb{Z}_p^*/U_{n-1}$$

carries V_n isomorphically onto V_{n-1}. Since

$$\mathbb{Z}_p^* = \varprojlim \big(\mathbb{Z}_p^*/U_n \big),$$

we get that V is a subgroup of $\mathbb{Z}_p^*$ isomorphic to $\mathbb{F}_p^*$. Passing to the limit, we have $\mathbb{Z}_p^* = V \times U_1$. The uniqueness of V follows from the uniqueness of V_n. □

Corollary 2.106. *The field $\mathbb{Q}_p$ contains the $(p-1)$th roots of unity.*

Remark 2.107.

(i). The group V is called the group of *multiplicative representatives* of the elements of $\mathbb{F}_p^*$.

(ii). The existence of V can also be proved by applying Corollary 2.102 to the equation $x^{p-1} - 1 = 0$.

Lemma 2.108. *Let $p \in P$ and*

$$n \in \begin{cases} \mathbb{N} & \text{if } p \neq 2 \\ \mathbb{N}_{\geq 2} & \text{else.} \end{cases}$$

For $x \in U_n \setminus U_{n+1}$ we have that $x^p \in U_{n+1} \setminus U_{n+2}$.

Proof. By the definition x has the form $x = 1 + kp^n$ with $k \not\equiv 0 \pmod{p}$. The binomial formula yields

$$x^p = 1 + kp^{n+1} + \ldots + k^p p^{np}$$

where the exponents hidden in the dats are $\geq 2n+1$, hence also $\geq n+2$. Moreover $np \geq n+2$ (since $n \geq 2$ if $p = 2$). This shows

$$x^p \equiv 1 + kp^{n+1} \pmod{p^{n+2}}$$

hence $x^p \in U_{n+1} \setminus U_{n+2}$. □

Proposition 2.109. *For $p \in P$ we have*

(i). If $p \neq 2$: $U_1 \cong \mathbb{Z}_p$.

(ii). If $p = 2$: $U_1 \cong \{\pm 1\} \times U_2$ and $U_2 \cong \mathbb{Z}_2$.

Proof. We distinguish the two cases:

Case 1 ($p \neq 2$). Choose $\alpha \in U_1 \setminus U_2$, for example $\alpha := 1+p$. By Lemma 2.108, we have $\alpha^{p^i} \in U_{i+1} \setminus U_{i+2}$. Define for every $n \in \mathbb{N}$: $\alpha_n \in U_1/U_n$ as the image of α. We have

$$(a_n)^{p^{n-2}} \neq 1 \quad \text{and} \quad (\alpha)^{p^{n-1}} = 1.$$

But U_1/U_n is of order p^{n-1}, hence it is a cyclic group generated by α_n. Now define the following isomorphisms

$$\begin{array}{rccc} \theta_{n,\alpha}\colon & \mathbb{Z}/p^{n-1}\mathbb{Z} & \longrightarrow & U_1/U_n \\ & z & \longmapsto & \alpha_n^z \end{array}$$

and look at the following commutative diagram

$$\begin{array}{ccc} \mathbb{Z}/p^n\mathbb{Z} & \xrightarrow{\theta_{n+1,\alpha}} & U_1/U_{n+1} \\ \downarrow & & \downarrow \\ \mathbb{Z}/p^{n-1}\mathbb{Z} & \xrightarrow{\theta_{n,\alpha}} & U_1/U_n \end{array}$$

From this, we see that the $\theta_{n,\alpha}$ define an isomorphism $\theta\colon \mathbb{Z}_p \to U_1$ since $\mathbb{Z}_p = \varprojlim \big(\mathbb{Z}/p^{n-1}\mathbb{Z}\big)$ and $U_1 = \varprojlim (U_1/U_n)$.

Case 2 ($p = 2$). Choose $\alpha \in U_2 \setminus U_3$. This means $\alpha \equiv 5 \pmod{8}$. Again define isomorphisms

$$\theta_{n,\alpha}\colon \mathbb{Z}/2^{n-2}\mathbb{Z} \to U_2/U_n$$

and argument as above to obtain an isomorphism $\theta_\alpha\colon \mathbb{Z}_2 \to U_2$. On the other hand, the homomorphism

$$U_1 \to U_1/U_2 \cong \mathbb{Z}/2\mathbb{Z}$$

induces an isomorphism $\{\pm 1\} \to \mathbb{Z}/2\mathbb{Z}$. This ultimatley gives

$$U_1 = \{\pm 1\} \times U_2.$$

□

Theorem 2.110. *For $p \in P$ we have*

(i). If $p \neq 2$: $\mathbb{Q}_p^* \cong \mathbb{Z} \times \mathbb{Z}_p \times \mathbb{Z}/(p-1)\mathbb{Z}$.

(ii). If $p = 2$: $\mathbb{Q}_p^* \cong \mathbb{Z} \times \mathbb{Z}_2 \; \mathbb{Z}/2\mathbb{Z}$.

Theorem 2.111. *For $p \in P \setminus \{ 2 \}$ and $x \in \mathbb{Q}_p^*$ write $x = p^n u$ with $n \in \mathbb{Z}$ and $u \in \mathbb{Z}_p^*$. Now it holds that*

$$x \text{ is a square} \quad \Longleftrightarrow \quad n \text{ is even} \quad \text{and} \quad \bar{u} \in \mathbb{F}_p^* = \mathbb{Z}_p^*/U_1 \text{ is a square.}$$

$\bar{u}$ denotes the image of u in $\mathbb{F}_p^*$. The condition $\bar{u} \in \mathbb{F}_p^* = \mathbb{Z}_p^*/U_1$ is a square, means that the Legendre symbol is one i.e. $\left(\frac{\bar{u}}{p}\right) = 1$. In the following, we write $\left(\frac{u}{p}\right)$ instead of $\left(\frac{\bar{u}}{p}\right)$.

Proof. Decompose u as $u = vu_1$ where $v \in V$ and $u_1 \in U_1$. The decomposition

$$\mathbb{Q}_p^* \cong \mathbb{Z} \times V \times U_1$$

given in Theorem 2.110 yields that x is a square if and only if n is even and v and u_1 are squares. But U_1 is isomorphic to $\mathbb{Z}_p$ and 2 is invertible in $\mathbb{Z}_p$, so all elements of U_1 are squares. Since $V \cong \mathbb{F}_p^*$, the statements follows. □

Corollary 2.112. *if $p \in P \setminus \{ 2 \}$, the group $\mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ is a group of type $(2,2)$. It has $\{ 1, p, u, up \}$ as representatives where $u \in \mathbb{Z}_p^*$ such that $\left(\frac{u}{p}\right) = -1$.*

Theorem 2.113. *Let $x \in \mathbb{Q}_2^*$ and write $x = p^n u$ with $n \in \mathbb{Z}$ and $u \in \mathbb{Z}_2^*$. We have*

$$x \text{ is a square} \quad \Longleftrightarrow \quad (n \text{ is even} \quad \text{and} \quad u \equiv 1 \pmod 8) .$$

Proof. Since we have the decomposition $\mathbb{Z}_p^* = \{ \pm 1 \} \times U_2$, we have that u is a square if and only if $u \in U_2$ and it is a square there. The isomorphism $\theta \colon \mathbb{Z}_2 \to U_2$ defined in the proof of Proposition 2.109 carries $2^n\mathbb{Z}_2$ onto U_{n+2}. Now take $n = 1$ and observe that beeing a square in U_2 is the same as beeing a square in U_3. So $u \in \mathbb{Z}_p^*$ is a square if and only if $u \equiv 1 \pmod 8$, hence the statement follows. □

Corollary 2.114. *The group $\mathbb{Q}_2^* \Big/ [\mathbb{Q}_2^*]^2$ is of type $(2,2,2)$. It has*

$$\{ \pm 1, \pm 5, \pm 2, \pm 10 \}$$

as representatives.

Remark 2.115. Theorems 2.111 and 2.113 show that for every $p \in P$ we have that $\left[\mathbb{Q}_p^*\right]^2$ is an open subgroup of $\mathbb{Q}_p^*$.

Chapter 3

The Hilbert Symbol

Abstract

In this chapter, we will introduce the *Hilbert symbol* $(\cdot,\cdot)_v$ and prove some well-known properties that will afterwards be used to lay the foundation for the proof of the upper bound

$$\textsc{HilbertSymbol}_{\mathbb{Q}} \in \mathbf{P}^{\textsc{IntFact}}.$$

On the way, we define the *Hilbert set* for $a, b \in \mathbb{Q}^*$ by

$$\mathcal{H}_{a,b} := \{\, v \in V \mid (a,b)_v = -1 \,\}.$$

We will show that an oracle for IntFact leads to an algorithm that enumerates the Hilbert set in polynomial time.

3.1 Definition

Definition 3.1 (Hilbert equation). For a field $\mathbb{F}$ and $a, b \in \mathbb{F}^*$ we call

$$aX^2 + bY^2 = Z^2$$

a **Hilbert equation** and a solution $(x, y, z) \in \mathbb{F}^3$ to it **trivial** if and only if $x = y = z = 0$, and **nontrivial** otherwise.

Definition 3.2 (Hilbert symbol). For $v \in V$ the **Hilbert symbol relative to** $\mathbb{Q}_v$ is the map

$$\begin{array}{rcl} (\cdot,\cdot)_v : \mathbb{Q}_v^* \times \mathbb{Q}_v^* & \longrightarrow & \{-1, 1\} \\ (a,b) & \longmapsto & \begin{cases} 1 & \text{if } \exists\,(x,y,z) \in \mathbb{Q}_v^3 \setminus \{\mathbf{0}\} : ax^2 + by^2 = z^2 \\ -1 & \text{else.} \end{cases} \end{array}$$

Even though it looses a lot of the nice properties that we will discuss in the next sections, we will generalize the Hilbert symbol to a setting where the coefficients are elements of an arbitrary field and the possible solutions reside in a subring. This notation enables us to easily speak about the existence of a nontrivial solution of Hilbert equations in a subring over arbitrary fields.

Definition 3.3 (Generalized Hilbert symbol). Let $\mathbb{F}$ be a field and $R \subseteq \mathbb{F}$ be a subring. We define the following map that captures the nontrivial R-solvability of a quadratic diagonal equation in three variables with coefficients in $\mathbb{F}$:

$$\begin{array}{rcl} \mathrm{h}_{R,\mathbb{F}}(\cdot,\cdot) : \mathbb{F}^* \times \mathbb{F}^* & \longrightarrow & \{-1,1\} \\ (a,b) & \longmapsto & \begin{cases} 1 & \text{if } \exists\,(x,y,z) \in R^3 \setminus \{\mathbf{0}\} : ax^2 + by^2 = z^2 \\ -1 & \text{else.} \end{cases} \end{array}$$

If we choose $R = \mathbb{F}$, we also write $\mathrm{h}_{\mathbb{F}}(\cdot,\cdot) := \mathrm{h}_{\mathbb{F},\mathbb{F}}(\cdot,\cdot)$.

We cleary have that $(\cdot,\cdot)_v = \mathrm{h}_{\mathbb{Q}_v}(\cdot,\cdot)$. So it is indeed a generalization of the Hilbert symbol. Each property that we prove for $\mathrm{h}_{\cdot,\cdot}(\cdot,\cdot)$, will also hold for any $(\cdot,\cdot)_v$.

Remark 3.4. Since neither $(a,b)_v$ nor $\mathrm{h}_{R,\mathbb{F}}(a,b)$ changes if a or b are multiplied by nonzero squares, they can both be seen as maps

$$\mathbb{F}^* \Big/ \left[\mathbb{F}^*\right]^2 \times \mathbb{F}^* \Big/ \left[\mathbb{F}^*\right]^2 \to \{-1,1\}.$$

Definition 3.5 (Hilbert Problem). Let $\mathbb{F}$ be a field and $R \subseteq \mathbb{F}$ be a subring. We now define the problem of deciding the nontrivial R-solvability of a quadratic diagonal equation in three variables with coefficients in $\mathbb{F}$:

$$\textsc{HilbertSymbol}_{R,\mathbb{F}} := \left\{ (a,b) \in \left(\mathbb{F}^*\right)^2 \;\middle|\; \mathrm{h}_{R,\mathbb{F}}(a,b) = 1 \right\}.$$

If we choose $R = \mathbb{F}$ we also write

$$\textsc{HilbertSymbol}_{\mathbb{F}} := \textsc{HilbertSymbol}_{\mathbb{F},\mathbb{F}}.$$

Fact 3.6. *We have*

$$\mathrm{h}_{\mathbb{Z},\mathbb{Q}}(\cdot,\cdot) = \mathrm{h}_{\mathbb{Q}}(\cdot,\cdot) \quad \textit{and} \quad \textsc{HilbertSymbol}_{\mathbb{Z},\mathbb{Q}} = \textsc{HilbertSymbol}_{\mathbb{Q}}.$$

Proof. Let $a,b \in \mathbb{Q}^*$ and $(x,y,z) \in \mathbb{Q}^3 \setminus \{\mathbf{0}\}$ be a nontrivial solution of the Hilbert equation in a and b. Then one can multiply the solution by all denominators of x, y and z to obtain a solution $(x',y',z') \in \mathbb{Z}^3 \setminus \{\mathbf{0}\}$. On the other hand any solution over $\mathbb{Z}$ is trivially a solution over $\mathbb{Q}$. □

3.2 Properties and Formulas

Fact 3.7 (Simple Properties of the generalized Hilbert symbol). *For any field $\mathbb{F}$ and subring $R \subseteq \mathbb{F}$ we have*

(i). $\forall\, a \in \mathbb{F}^* :\ \mathrm{h}_{R,\mathbb{F}}(a,1) = 1.$

(ii). $\forall\, a,b \in \mathbb{F}^* :\ \mathrm{h}_{R,\mathbb{F}}(a,b) = \mathrm{h}_{R,\mathbb{F}}(b,a).$

(iii). $\forall\, a,c \in \mathbb{F}^* :\ \mathrm{h}_{R,\mathbb{F}}\left(a,c^2\right) = 1.$

(iv). $\forall\, a \in \mathbb{F}^* :\ \mathrm{h}_{R,\mathbb{F}}(a,-a) = 1.$

(v). $\forall\, a \in \mathbb{F}^* \setminus \{1\} :\ \mathrm{h}_{R,\mathbb{F}}(a,1-a) = 1.$

(vi). $\forall\, a,b,c \in \mathbb{F}^*\colon\ \mathrm{h}_{R,\mathbb{F}}(a,b) = \mathrm{h}_{R,\mathbb{F}}\left(a, c^2 b\right)$.

Proof.

(i). $(0,1,1)$ is a solution for $aX^2 + Y^2 = Z^2$.

(ii). This is obvious since $aX^2 + bY^2 = Z^2$ is symmetric in a and b.

(iii). $(0,1,c)$ is a solution for $aX^2 + (cY)^2 = Z^2$.

(iv). $(1,1,0)$ is a solution for $aX^2 - aY^2 = Z^2$.

(v). $(1,1,1)$ is a solution for $aX^2 + (1-a)Y^2 = Z^2$.

(vi). A nontrivial solution (x,y,z) to $aX^2 + bY^2 = Z^2$ can be translated to a nontrivial solution (x, yc^{-1}, z) to $aX^2 + bc^2Y^2 = Z^2$ and vice versa. □

Using Definition 2.61, we can prove the following lemma.

Lemma 3.8. *For any $v \in V$ and $a,b \in \mathbb{Q}_v^*$ define $\mathbb{Q}_{v,b} := \mathbb{Q}_v\left[\sqrt{b}\right]$. Then it holds that*

$$(a,b)_v = 1 \quad \Leftrightarrow \quad a \in N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}\left(\mathbb{Q}_{v,b}^*\right).$$

Proof.

Case 1 (b is a square). If b is the square of an element $c \in \mathbb{Q}_v$, the equation $aX^2 + bY^2 = Z^2$ is always solved by $(0,1,c)$, hence $(a,b)_v = 1$ and the proposition follows since in this case we also have that $\mathbb{Q}_{v,b} = \mathbb{Q}_v$ and $N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}\left(\mathbb{Q}_{v,b}^*\right) = \mathbb{Q}_v^*$.

Case 2 (b is not a square). $\mathbb{Q}_{v,b}$ is quadratic over $\mathbb{Q}_v$. Let $\beta := \sqrt{b} \in \mathbb{Q}_{v,b}$ be a square root in the field extension. Then every element $\xi \in \mathbb{Q}_{v,b}$ can be written as $\xi = z + \beta y$ with $z,y \in \mathbb{Q}_v$. Note that $N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}(\xi) = z^2 - by^2$.

Now if $a \in N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}\left(\mathbb{Q}_{v,b}^*\right)$, there exist $y,z \in \mathbb{Q}_v$ such that $a = z^2 - by^2$. This yields $(1,y,z)$ as a solution for the equation $aX^2 + bY^2 = Z^2$, hence $(a,b)_v = 1$. Conversely, if $(a,b)_v = 1$ we have a solution $(x,y,z) \neq (0,0,0)$ for $aX^2 + bY^2 = Z^2$.

Case 2.1 ($x = 0$ and $y = 0$). Then $z = 0$, which is a contradiction.

Case 2.2 ($x = 0$ but $y \neq 0$). We have that $b = \frac{z^2}{y^2}$ which is a contradiction to the fact that b is not a square

Case 2.3 ($x \neq 0$). From $ax^2 + by^2 = z^2$ it follows that

$$a = \left(\frac{z}{x}\right)^2 - b\left(\frac{y}{x}\right)^2 = N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}\left(\frac{z}{x} + \beta\frac{y}{x}\right) \in N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}\left(\mathbb{Q}_{v,b}^*\right). \qquad \square$$

Proposition 3.9 (Properties of the Hilbert symbol). *For every $v \in V$, we have:*

(i). $\forall\, a,a',b \in \mathbb{Q}_v^*\colon\ (a,b)_v = 1 \ \Rightarrow (aa',b)_v = (a',b)_v$

(ii). $\forall\, a,b \in \mathbb{Q}_v^*\colon\ (a,b)_v = (a,-ab)_v$

(iii). $\forall\, a \in \mathbb{Q}_v^* \setminus \{1\}, b \in \mathbb{Q}_v^*\colon\ (a,b)_v = (a,(1-a)b)_v$

Proof.

(i). If $(a,b)_v = 1$, Lemma 3.8 implies that $a \in N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}\left(\mathbb{Q}_{v,b}^*\right)$. Now we have that $a' \in N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}\left(\mathbb{Q}_{v,b}^*\right) \Leftrightarrow aa' \in N_{\mathbb{Q}_{v,b}/\mathbb{Q}_v}\left(\mathbb{Q}_{v,b}^*\right)$ and again Lemma 3.8 gives the property.

(ii). By fact 3.7(v) we have that for any $a \in \mathbb{Q}_v^*$ it holds that $(a,-a)_v \overset{3.7(ii)}{=} (-a,a)_v = 1$ and we get by (i) that for all $a' \in \mathbb{Q}_v^*$ it holds that $(-aa',a)_v = (a',a)_v$. Replace $a' = b$ and use the symmetry (i.e. fact 3.7(ii)) again to see $(a,-ab)_v = (a,b)_v$.

(iii). By fact 3.7(v) we have that for any $a \in \mathbb{Q}_v^*$ it holds that $(1-a,a)_v \overset{3.7(ii)}{=} (a,1-a)_v = 1$ and we get by (i) that for all $a' \in \mathbb{Q}_v^*$ it holds that $((1-a)a',a)_v = (a',a)_v$. Replace $a' = b$ and use the symmetry (i.e. fact 3.7(ii)) again to see $(a,(1-a)b)_v = (a,b)_v$. □

Lemma 3.10. *Let $p \in P$ and $v \in \mathbb{Z}_p^*$ be a p-adic unit. If the equation*

$$Z^2 - pX^2 - vY^2 = 0$$

has a nontrivial solution in $\mathbb{Q}_p$, it also has a solution (x,y,z) such that $x,y \in \mathbb{Z}_p^*$ and $z \in \mathbb{Z}_p$.

Proof. By Proposition 2.98 every nontrivial solution in $\mathbb{Q}_p^3$ gives a primitive solution $(x,y,z) \in \mathbb{Z}_p^3$ (that is not all of the x,y,z are divisible by p). We will show that this solution already has the desired properties. First note that

$$0 = z^2 - vy^2 - px^2 \equiv z^2 - vy^2 \pmod{p},$$

yielding

$$z^2 \equiv vy^2 \pmod{p}. \tag{3.2.1}$$

Let us now assume the contrary:

$$x \notin \mathbb{Z}_p \quad \text{or} \quad y \notin \mathbb{Z}_p^* \quad \text{or} \quad z \notin \mathbb{Z}_p^*.$$

The first is not possible since $(x,y,z) \in \mathbb{Z}_p^3$. So let us assume that one of the latter two is true. This means

$$y \equiv 0 \pmod{p} \quad \text{or} \quad z \equiv 0 \pmod{p}.$$

Together with (3.2.1), the first one also gives $z \equiv 0 \pmod{p}$ and the second one $vy^2 \equiv 0 \pmod{p}$. Since $v \in \mathbb{Z}_p^*$ it follows that $y \equiv 0 \pmod{p}$. So in both cases we have $y \equiv 0 \pmod{p}$ *and* $z \equiv 0 \pmod{p}$ and get

$$0 = z^2 - px^2 - vy^2 \equiv -px^2 \pmod{p^2}$$

which ultimately gives $x \equiv 0 \pmod{p}$, a contradiction to the primitive character of (x,y,z). □

Theorem 3.11. *Let $v \in V$ and $a, b \in \mathbb{Q}_v^*$. Then we can compute the Hilbert symbol as follows:*

If $v = \infty$: *It holds that $(a,b)_\infty = -1$ if and only if $a, b < 0$.*

If $v \in P$: *Write $a = v^\alpha u$ and $b = v^\beta w$ where u and w are v-adic units, we have*

If $v = 2$: $(a,b)_v = (-1)^{\epsilon(u)\,\epsilon(w) + \alpha\,\omega(w) + \beta\,\omega(u)}$

If $v \neq 2$: $(a,b)_v = (-1)^{\alpha\beta\,\epsilon(v)} \left(\frac{u}{v}\right)^\beta \left(\frac{w}{v}\right)^\alpha$

Where $\left(\frac{u}{v}\right)$ denotes the Legendre symbol from Section 2.6 reduced modulo v — i.e. $\left(\frac{\pi(u)}{v}\right)$ where $\pi\colon \mathbb{Z}_v^* \to \mathbb{F}_v^*$ is the canonical projection — and with $\epsilon(u) = \frac{u-1}{2}$ and $\omega(u) = \frac{u^2-1}{8}$.

Proof. We distinguish the three cases $v = \infty$, $v = 2$ and $v \in P \setminus \{2\}$:

Case 1 ($v = \infty$). If $a, b \in \mathbb{R}_{<0}$ the equation $ax^2 + by^2 = z^2$ directly implies $x = y = z = 0$ and therefore $(a,b)_\infty = -1$. For the converse we can assume that $a \in \mathbb{R}_{>0}$ since the Hilbert symbol is symmetric in a and b. Then for example $\left(\frac{1}{\sqrt{a}}, 0, 1\right)$ is a nontrivial solution for $ax^2 + by^2 = z^2$.

Case 2 ($p := v \notin \{\infty, 2\}$). The exponents α and β come in only by their residue (mod 2). Since the Hilbert symbol is symmetric, we only have to distinguish three cases:

Case 2.1 ($\alpha = 0, \beta = 0$). In this case it holds that $a = u$, $b = w$, and the right hand side is 1 so we have to show that $(u,w)_p = 1$. Consider the equation

$$Z^2 - uX^2 - wY^2 = 0.$$

It is a form in 3 variables and by Corollary 2.72 it therefore has a nontrivial solution modulo p. Since the discriminant of this equation is $1 \cdot u \cdot w$, which is as a product of p-adic units a p-adic unit and therefore primitive, the above solution lifts to a p-adic solution by Corollary 2.103, hence $(u,w)_p = 1$.

Case 2.2 ($\alpha = 1, \beta = 0$). We have to check that $(pu,w)_p = \left(\frac{w}{p}\right)$. By the previous step, we have that $(u,w)_p = 1$ which by Proposition 3.9(i) yields that $(pu,w)_p = (p,w)_p$. So its left to show that $(p,w)_p = \left(\frac{w}{p}\right)$. If w is a square, this is clear since then both terms are 1. Otherwise, we have $\left(\frac{w}{p}\right) = -1$ by Theorem 2.111. So w is not a square modulo p. If we now assume that (x,y,z) is a nontrivial solution of $pX^2 + wY^2 = Z^2$ by Lemma 3.10 we can without loss of generality assume that $y, z \in \mathbb{Z}_p^*$ and $x \in \mathbb{Z}_p$. Looking at this equation as an equation in $\mathbb{F}_p$, we get $wy^2 = z^2$ and since y and z are invertible in $\mathbb{Z}_p$, that w is a square modulo p. This contradicts the assumption and therefore yields that there is no nontrivial solution i.e. $(p,w)_p = -1$.

Case 2.3 ($\alpha = 1, \beta = 1$)**.** We calculate

$$\begin{aligned}(pu, pw)_p &\overset{3.9(ii)}{=} \left(pu, -p^2uw\right)_p \\ &\overset{3.7(vi)}{=} (pu, -uw)_p \\ &\overset{(*)}{=} \left(\frac{-uw}{p}\right) \\ &\overset{2.50(ii)}{=} \left(\frac{-1}{p}\right)\left(\frac{u}{p}\right)\left(\frac{w}{p}\right) \\ &\overset{2.54(ii)}{=} (-1)^{\frac{p-1}{2}}\left(\frac{u}{p}\right)\left(\frac{w}{p}\right)\end{aligned}$$

where at the equality (*) we apply the same argument as in Case 2.2 with w replaced by $-uw$.

Case 3 ($p := v = 2$)**.** Again, only the residual classes modulo 2 of α and β play a role and we distinguish the different cases

Case 3.1 ($\alpha = 0$ **and** $\beta = 0$)**.** We have to check that

$$(u, w)_2 = \begin{cases} 1 & u \equiv 1 \pmod 4 \text{ or } w \equiv 1 \pmod 4 \\ -1 & \text{otherwise.} \end{cases}$$

By symmetry we only have to distinguish the following two cases

Case 3.1.1 ($u \equiv 1 \pmod 4$)**.** We can distinguish two cases

Case 3.1.1.1 ($u \equiv 1 \pmod 8$)**.** By Theorem 2.113 we have that u is a square implying that $(u, w)_2 = 1$.

Case 3.1.1.2 ($u \equiv 5 \pmod 8$)**.** We have $u + 4w \equiv 1 \pmod 8$ and Theorem 2.113 again implies that there exists $z \in \mathbb{Z}_p^*$ such that $z^2 = u + 4w$. Thus the form $uX^2 + wY^2 = Z^2$ has $(1, 2, z)$ as a solution implying that $(u, w)_2 = 1$.

Case 3.1.2 ($u \equiv w \equiv -1 \pmod 4$)**.** If (x, y, z) is a primitive solution for $uX^2 + vY^2 = Z^2$, then $x^2 + y^2 + z^2 = 0 \pmod 4$ by assumption. But the squares of $\mathbb{Z}/4\mathbb{Z}$ are 0 and 1 implying that $x \equiv y \equiv z \equiv 0 \pmod 2$ which contradicts the assumption of primitivity. Thus $(u, w)_2 = -1$.

Case 3.2 ($\alpha = 1$ **and** $\beta = 0$)**.** We have to check that

$$(2u, w)_2 = (-1)^{\epsilon(u)\,\epsilon(w)+\omega(w)}$$

For this, we will first prove that $(2, w)_2 = (-1)^{\omega(w)}$ or in other words that

$$(2, w)_2 = 1 \quad \Longleftrightarrow \quad w \equiv \pm 1 \pmod 8.$$

Step 1 ("⇒"). By Lemma 3.10 there exist $x, y, z \in \mathbb{Z}_2$ such that

$$2x^2 + wy^2 = z^2 \tag{3.2.2}$$

and $y, z \not\equiv 0 \pmod 2$. For all such y and z it holds that $y^2 \equiv z^2 \equiv 1 \pmod 8$. So we can reduce (3.2.2) to

$$2x^2 + w \equiv 1 \pmod 8.$$

The only squares modulo 8 are 0, 1 and 4 implying $w \equiv \pm 1 \pmod 8$.

Step 2 ("⇐").

Case 3.2.1 ($w \equiv 1 \pmod 8$)**.** By Theorem 2.113 w is a square and therefore $(2,w)_2 = 1$.

Case 3.2.2 ($w \equiv -1 \pmod 8$)**.** The equation

$$Z^2 - 2X^2 - vY^2 \equiv 0 \pmod 8$$

has $(1,1,1)$ for a solution. By Corollary 2.103 this solution lifts to a true solution, thus we have $(2,w)_2 = 1$.

Now we will show that $(2u,w)_2 = (2,w)_2 (u,w)_2$. For this, we distinguish three cases:

Case 3.2.1 ($(2,w)_2 = 1$)**.** $(2u,w)_2 \overset{3.9(\mathrm{i})}{=} (u,w)_2 = (2,w)_2 (u,w)_2$.

Case 3.2.2 ($(u,w)_2 = 1$)**.** $(2u,w)_2 \overset{3.9(\mathrm{i})}{=} (2,w)_2 = (2,w)_2 (u,w)_2$.

Case 3.2.3 ($(2,w)_2 = (u,w)_2 = -1$)**.** We are in the case that

$$w \equiv 3 \pmod 8 \quad \text{and} \quad (u \equiv 3 \pmod 8 \text{ or } u \equiv -1 \pmod 8).$$

Since we are free to multiply w and u by squares, we can assume that

$$(u = -1 \text{ and } w = 3) \quad \text{or} \quad (u = 3 \text{ and } w = -5).$$

But the equations

$$Z^2 + 2X^2 - 3Y^2 = 0 \quad \text{and} \quad Z^2 - 6X^2 + 5Y^2 = 0$$

have for solution $(1,1,1)$ implying that $(2u,w)_v = 1$.

Case 3.3 ($\alpha = 1$ **and** $\beta = 1$)**.** We calculate

$$(2u,2w)_v \overset{3.9(\mathrm{ii})}{=} (2u,-4uw)_v \overset{3.7(\mathrm{vi})}{=} (2u,-uw)_v .$$

By the above we therefore have

$$(2u,2w)_v = (-1)^{\epsilon(u)\,\epsilon(-uw)+\omega(-uw)} .$$

Now since $\epsilon(-1) = 1, \omega(-1) = 0$ and $\epsilon(u)(1+\epsilon(u)) = 0$ this means

$$(2u,2w)_v = (-1)^{\epsilon(u)\,\epsilon\, w+\omega\, u+\omega\, w} . \qquad \square$$

Definition 3.12. Let $v \in V$. We say that the **Hilbert symbol is nondegenerate** if and only if for $b \in \mathbb{Q}_v^*$ it holds that

$$\forall\, a \in \mathbb{Q}_v^* \colon\ (a,b)_v = 1 \quad \Rightarrow \quad b \in \left[\mathbb{Q}_v^*\right]^2 .$$

Theorem 3.13. *The Hilbert symbol is a nondegenerate bilinear form on the* $\mathbb{F}_2$*-vector space* $\mathbb{F}^* \Big/ \left[\mathbb{F}^*\right]^2$.

Proof. The bilinearity follows from the formulas in Theorem 3.11, Lemma 2.50(ii) and the fact that ϵ and ω are homomorphisms. Nondegeneracy can be checked on the representatives $\{1, 2u\}$ where $u \in \{1,5,-1,-5\}$. Indeed, we have $(5,2u)_v = -1$ and

$$(-1,-1)_v = (-1,-5)_v = -1. \qquad \square$$

Remark 3.14. One can view the Hilbert symbol as a form $(-1)^{[a,b]}$ where $[a,b]$ is a symmetric bilinear form on $\mathbb{F}^* \Big/ [\mathbb{F}^*]^2$ with values in $\mathbb{Z}/2\mathbb{Z}$. Theorem 3.11 then gives its matrix with respect to some basis of $\mathbb{F}^* \Big/ [\mathbb{F}^*]^2$:

Case 1 ($\mathbb{F} = \mathbb{R}$). Independent of the choice of basis the matrix $\big(1\big)$.

Case 2 ($\mathbb{F} = \mathbb{Q}_p, p \in P \setminus \{2\}$). Take a basis $\{p, u\}$ where $\left(\frac{u}{p}\right) = -1$, then

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ if } p \equiv 1 \pmod 4 \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \text{ if } p \equiv 3 \pmod 4.$$

Case 3 ($\mathbb{F} = \mathbb{Q}_2$). With the basis $\{2, -1, 5\}$ it is the matrix

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

3.3 The Hilbert set

A theorem due to Hilbert tells us that for given $a, b \in \mathbb{Q}$, the Hilbert symbol is -1 only for finitely many places (i.e. elements of V). This means that the set $\mathcal{H}_{a,b}$ in the following definition is finite. Since we want to prove statements about the complexity of the Hilbert symbol, we will not only prove that $\mathcal{H}_{a,b}$ is finite but exactly determine its cardinality. On our way, we will not make use of the finiteness, but later obtain it as a corollary.

Definition 3.15 (Hilbert set). For $a, b \in \mathbb{Q}^*$ define

$$\mathcal{H}_{a,b} := \{\, v \in V \mid (a,b)_v = -1 \,\}.$$

We will first see, that if we take a and b as -1 or as prime numbers, the Hilbert set is not only finite but has a constant number of elements, meaning that its cardinality is asymptotically independent of a and b. For this, we will use Theorem 3.11 repeatedly.

Lemma 3.16. *For $a, b \in P \cup \{-1\}$ the set $\mathcal{H}_{a,b}$ is as follows:*

(i). $\mathcal{H}_{-1,-1} = \{\infty, 2\}$

(ii). For $p \in P$: $\mathcal{H}_{p,p} = \mathcal{H}_{-1,p} = \begin{cases} \varnothing & \text{if } p = 2 \text{ or } \epsilon(p) \text{ is odd} \\ \{2, p\} & \text{else.} \end{cases}$.

(iii). For $p \in P \setminus \{2\}$: $\mathcal{H}_{p,2} = \begin{cases} \{2, p\} & \text{if } \omega(p) \text{ is even} \\ \varnothing & \text{else.} \end{cases}$

(iv). For distinct $p, q \in P \setminus \{2\}$: $\mathcal{H}_{p,q} \subseteq \{2, p, q\}$*, more precisely:*

- $2 \in \mathcal{H}_{p,q} \iff \epsilon(p)\,\epsilon(q)$ *is odd*
- $p \in \mathcal{H}_{p,q} \iff \left(\frac{p}{q}\right) = -1$
- $q \in \mathcal{H}_{p,q} \iff \left(\frac{q}{p}\right) = -1$

Proof. For $v \in V$, $p, q \in P \cup \{-1\}$ Theorem 3.11 gives the value of $(p,q)_v$:

Case 1 $(p, q = -1)$**.** We have $(-1,-1)_\infty = (-1,-1)_2 = -1$ and $(-1,-1)_v = 1$ for $v \notin \{2, \infty\}$.

Case 2 $(p = -1, q \in P)$**.**

Case 2.1 $(q = 2)$**.** $(-1,2)_v = 1$ for every $v \in V$.

Case 2.2 $(q \neq 2)$**.**

Case 2.2.1 $(v \in \{2, q\})$**.** $(-1,q)_v = (-1)^{\epsilon(q)}$.

Case 2.2.2 $(v \notin \{2, q\})$**.** $(-1,q)_v = 1$.

Case 3 $(p \in P, q \in P)$**.**

Case 3.1 $(p = q)$**.** We calculate for any $v \in V$:

$$(p,p)_v \overset{3.9(ii)}{=} \left(p, -p^2\right)_v = (p,-1)_v \cdot \left(p, p^2\right)_v \overset{3.9(iii)}{=} (p,-1)_v .$$

By Case 2 we are done.

Case 3.2 $(p \neq q)$**.**

Case 3.2.1 $(q = 2)$**.**

Case 3.2.1.1 $(v \in \{2, p\})$**.** $(p,2)_v = (-1)^{\omega(p)}$ (by 2.54(iii)).

Case 3.2.1.2 $(v \notin \{2, p\})$**.** $(p,2)_v = 1$

Case 3.2.2 $(q \neq 2)$**.**

Case 3.2.2.1 $(v \notin \{2, p, q\})$**.** $(p,q)_v = 1$.

Case 3.2.2.2 $(v = 2)$**.** $(p,q)_2 = (-1)^{\epsilon(p)\,\epsilon(q)}$

Case 3.2.2.3 $(v = p)$**.** $(p,q)_p = \left(\frac{q}{p}\right)$

Case 3.2.2.4 $(v = q)$**.** $(p,q)_q = \left(\frac{p}{q}\right)$ □

Next, the linearity of the Hilbert symbol will enable us to calculate the Hilbert set not only for -1 and primes but for every pair of rational numbers.

Theorem 3.17. *Given $a = \pm\frac{a_1}{a_2}, b = \pm\frac{b_1}{b_2} \in \mathbb{Q}^*$ and the factorizations of their numerator and denominator, the set $\mathcal{H}_{a,b}$ can be determined in polynomial time in the bitsizes of the numerators and denominators of a and b. More precisely the complexity is in $\mathcal{O}\left(n^2\right)$ where $n = \log\left(\max\{a_1, a_2, b_1, b_2\}\right)$.*

Proof. The factorization of a number $d \in \mathbb{Z}$ can be written as

$$d = \pm\prod_{i=1}^{r} p_i \text{ with } r \in \mathcal{O}\left(\log\left(d\right)\right), \forall\, i \in [1:r]\colon p_i \in P$$

Let $a = \varepsilon\frac{a_1}{a_2}$ and $b = \eta\frac{b_1}{b_2}$ with $a_1, a_2, b_1, b_2 \in \mathbb{N}$ and $\varepsilon, \eta \in \{-1, 1\}$. Then by

the bilinearity of the Hilbert symbol (see Theorem 3.13), we calculate

$$\begin{aligned}
&\left(\varepsilon\frac{a_1}{a_2}, \eta\frac{b_1}{b_2}\right)_v \\
&= (\varepsilon,\eta)_v \left(\varepsilon, \frac{b_1}{b_2}\right)_v \left(\frac{a_1}{a_2}, \eta\right)_v \left(\frac{a_1}{a_2}, \frac{b_1}{b_2}\right)_v \\
&= (\varepsilon,\eta)_v (\varepsilon,b_1)_v \left(\varepsilon, \frac{1}{b_2}\right)_v (a_1,\eta)_v \left(\frac{1}{a_2}, \eta\right)_v \left(a_1, \frac{b_1}{b_2}\right)_v \left(\frac{1}{a_2}, \frac{b_1}{b_2}\right)_v \\
&= (\varepsilon,\eta)_v (\varepsilon,b_1)_v \left(\varepsilon, \frac{1}{b_2}\right)_v (a_1,\eta)_v \left(\frac{1}{a_2}, \eta\right)_v \\
&\quad \cdot (a_1,b_1)_v \left(a_1, \frac{1}{b_2}\right)_v \left(\frac{1}{a_2}, b_1\right)_v \left(\frac{1}{a_2}, \frac{1}{b_2}\right)_v .
\end{aligned}$$

Using bilinearity and symmetry of the Hilbert symbol and the factorizations of a_1, a_2, b_1, b_2, we end up in the situation where we have to calculuate an amount of

$$\mathcal{O}\left(\log(a_1)\log(b_1) + \log(a_1)\log(b_2) + \log(a_2)\log(b_1) + \log(a_2)\log(b_2)\right)$$

Hilbert symbols of the form

$$(\pm 1, \pm 1)_v \quad (\pm 1, p)_v \quad (p,q)_v \quad \left(p, \frac{1}{q}\right)_v \quad \left(\frac{1}{p}, \frac{1}{q}\right)_v \tag{3.3.1}$$

where p and q are prime numbers. Since

$$\begin{aligned}
&\log(a_1)\log(b_1) + \log(a_1)\log(b_2) + \log(a_2)\log(b_1) + \log(a_2)\log(b_2) \\
\leq &4\log\left(\max\{a_1,a_2\}\right)\log\left(\max\{b_1,b_2\}\right) \leq 4\log^2\left(\max\{a_1,a_2,b_1,b_2\}\right)
\end{aligned}$$

these are only polynomial many in the bitsizes of a_1, a_2, b_1 and b_2 — more precisely we have to calculate

$$\mathcal{O}\left(n^2\right), \text{ where } n := \log\left(\max\{a_1,a_2,b_1,b_2\}\right)$$

many Hilbert symbols. Whenever we encounter 1 as an argument, Fact 3.7(i) yields that $(x,1)_v = 1$. Now let $c \in \left\{p, \frac{1}{p}\right\}$ and use the following further reductions for the last two cases in (3.3.1):

$$\begin{aligned}
\left(c, \frac{1}{q}\right)_v = 1 &\Leftrightarrow cx^2 + \frac{1}{q}y^2 = z^2 \text{ is nontrivially solvable over } \mathbb{Q}_v \\
&\Leftrightarrow cqx^2 + y^2 = qz^2 \text{ is nontrivially solvable over } \mathbb{Q}_v \\
&\Leftrightarrow qz^2 - cqx^2 = y^2 \text{ is nontrivially solvable over } \mathbb{Q}_v \\
&\Leftrightarrow qx^2 - cqy^2 = z^2 \text{ is nontrivially solvable over } \mathbb{Q}_v \\
&\Leftrightarrow (q,-cq)_v = (q,-1)_v (q,c)_v (q,q)_v = 1
\end{aligned}$$

So for $c = \frac{1}{p}$ we reduce to the case where not both of the arguments are fractions and the case $c = p$ covers exactly this.

All these reductions only increase the number of Hilbert symbols that we have to calculate by a constant factor and ultimately show that we only have to

look at the cases where $a, b \in P \cup \{-1\}$. Lemma 3.16 shows that in this case, $|\mathcal{H}_{a,b}| \leq 3$ and the onyl thing that is left to calculate are the values of $\epsilon(p)\,\epsilon(q)$ and $\left(\frac{p}{q}\right)$ for $p, q \in P$. Determine $\epsilon(p) = \frac{p-1}{2}$ is easy and by Theorem 2.56 we know that $\left(\frac{p}{q}\right)$ can be calculated in linear time. □

This theorem also tells us, that the Hilbert set is finite for all a and b. In addition, we have the following nice formula for the Hilbert set where one of the parameters is a product of two numbers:

Fact 3.18. *For $a, b, c \in \mathbb{Q}^*$ it holds that*

$$\mathcal{H}_{a,bc} = \mathcal{H}_{a,b} \triangle \mathcal{H}_{a,c}.$$

Proof. Let $v \in V$. By Theorem 3.13, we know that $(a, bc)_v = (a, b)_v (a, c)_v$. Since the Hilbert symbol only takes values in $\{\pm 1\}$, this means that $-1 = (a, bc)_v$ is equivalent to $(a, b)_v \neq (a, c)_v$. So

$$v \in \mathcal{H}_{a,bc} \quad \Leftrightarrow \quad v \in (\mathcal{H}_{a,b} \setminus \mathcal{H}_{a,c}) \cup (\mathcal{H}_{a,c} \setminus \mathcal{H}_{a,b}) \quad \Leftrightarrow \quad v \in \mathcal{H}_{a,b} \triangle \mathcal{H}_{a,c}. \qquad \square$$

Properties of the symmetric difference enable us to prove the statement that the cardinality of the Hilbert set is even. This translates to a classical theorem by Hilbert, given after the proof.

Lemma 3.19. *The cardinality of the Hilbert set is even.*

Proof. By Fact 3.18, Theorem 3.13 and the fact that the cardinality of a symmetric difference of sets with an even number of elements is even, it suffices to show the statement for $a, b \in P \cup \{-1\}$. Lemma 3.16 gives the elements of $\mathcal{H}_{a,b}$. In the first three cases, we have $|\mathcal{H}_{a,b}| \in \{0, 2\}$. So it is left to show that for distinct and odd $a, b \in P$ exactly two of the following relations hold:

- $\epsilon(a)\,\epsilon(b)$ is odd
- $\left(\frac{a}{b}\right) = -1$
- $\left(\frac{b}{a}\right) = -1$

Since $a, b \neq 2$ the Quadratic reciprocity law (Theorem 2.57) states that

$$\left(\frac{a}{b}\right) = (-1)^{\epsilon(a)\,\epsilon(b)} \left(\frac{b}{a}\right).$$

If $\epsilon(a)\,\epsilon(b)$ is odd, $\left(\frac{a}{b}\right)$ and $\left(\frac{b}{a}\right)$ have different signs. If both $\left(\frac{a}{b}\right) = -1$ and $\left(\frac{b}{a}\right) = -1$ we have that $\epsilon(a)\,\epsilon(b)$ is even. And finally if for example $\left(\frac{a}{b}\right) = -1$ and $\epsilon(a)\,\epsilon(b)$ is odd, $\left(\frac{b}{a}\right) = 1$. □

This lemma gives the classical theorem about the Hilbert symbol as a corrollary:

Theorem 3.20 (Hilbert Theorem). *For $a, b \in \mathbb{Q}^*$ we have $(a, b)_v = 1$ for almost all $v \in V$ and*

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

Proof. This is a corollary to Lemma 3.19. □

Lemma 3.21 (Approximation Theorem). *Let $S \subseteq V$ be a finite set. The image of $\mathbb{Q}$ in $\prod_{v\in S} \mathbb{Q}_v$ is dense in this product.*

Proof. Since we are free to enlarge S, assume without loss of generality, that $S = \{\infty, p_1, \ldots, p_n\}$ where $p_i \in P$ are pairwise different. Its now left to show that $\mathbb{Q}$ is dense in

$$A := \mathbb{R} \times \mathbb{Q}_{p_1} \times \cdots \times \mathbb{Q}_{p_n}.$$

We will now show that any point $p := (x_\infty, x_1, \ldots, x_n) \in A$ is adherent to $\mathbb{Q}$: Since we are free to multiply p by some integer, we may suppose that $\forall\, i \in [1:n]\colon x_i \in \mathbb{Z}_{p_i}$. We are left to show that $\forall\, \varepsilon \in \mathbb{R}_{>0}, N \in \mathbb{N}\colon \exists\, x \in \mathbb{Q}$:

$$|x - x_\infty| \leq \varepsilon \quad \text{and} \quad \forall\, i \in [1:n]\colon \operatorname{ord}_{p_i}(x - x_i) \geq N.$$

So let $\varepsilon \in \mathbb{R}_{>0}$ and $N \in \mathbb{N}$ be arbitrary and define $\forall\, i \in [1:n]$ the number $m_i := p_i^N$. By the Chinese Remainder Theorem there exists $x_0 \in \mathbb{Z}$ with

$$\forall\, i \in [1:n]\colon \operatorname{ord}_{p_i}(x_0 - x_1) \geq N.$$

Now choose an integer $q \geq 2$ that is relatively prime to all the p_i (one can for example choose any $q \in P \setminus S$) and observe that

$$\left\{ \frac{a}{q^m} \;\middle|\; a \in \mathbb{Z}, m \in \mathbb{N} \right\}$$

is dense in $\mathbb{R}$ since $q^m \to \infty$ if $m \to \infty$ and choose $u = \frac{a}{q^m}$ with

$$\left|x_0 - x_\infty + u p_1^N \cdot \cdots \cdot p_n^N\right| \leq \varepsilon.$$

Finally $x := x_0 u p_1^N \cdot \cdots \cdot p_n^N$ has the desired property. □

Theorem 3.22. *Let $(a_i)_{i\in I}$ with $a_i \in \mathbb{Q}^*$ be a finite family and let $(\varepsilon_{i,v})_{i\in I, v\in V}$ be a family of numbers equal to ± 1. In order that there exists $x \in \mathbb{Q}^*$ such that $(a_i, x)_v = \varepsilon_{i,v}$ for all $i \in I$ and $v \in V$, it is necessary and sufficient that the following conditions are satisfied:*

(i). Almost all $\varepsilon_{i,v}$ are equal to 1.

(ii). $\forall\, i \in I\colon \prod_{v\in V} \varepsilon_{i,v} = 1$.

(iii). $\forall\, v \in V\colon \exists\, x_v \in \mathbb{Q}_v^*\colon \forall\, i \in I\colon (a_i, x_v)_v = \varepsilon_{i,v}$.

Proof. The necessity of (i) and (ii) follows from the Hilbert Theorem (Theorem 3.20) that of (iii) is trivial (take $x_v = x$). To prove sufficiency, we will need the Approximation Theorem (Lemma 3.21), Dirichlet Theorem (Theorem 2.11) and Chinese Remainder Theorem (Theorem 2.10):

Let $(\varepsilon_{i,v})$ be a family of numbers equal to ± 1 that satisfy (i) to (iii). We are free to multiply to a_i by squares of integers giving that without loss of generality we have $a_i \in \mathbb{Z}$. Now define S to be the set $\{\infty, 2\}$ together with all prime factors of the a_i. Additionally define

$$T := \{\, v \in V \mid \exists\, i \in I\colon \varepsilon_{i,v} = -1 \,\}.$$

Both S and T are finite. Now distinguish the following two cases

Case 1 ($S \cap T = \varnothing$). Define two numbers

$$a := \prod_{l \in T \setminus \{\infty\}} l \quad \text{and} \quad m := 8 \prod_{l \in S \setminus \{2, \infty\}} l.$$

By assumption $S \cap T = \varnothing$ and therefore $\gcd(a, m) = 1$. By the Dirichlet Theorem there exists a $p \in P \setminus (S \cup T)$ such that $p \equiv a \pmod{m}$. We will show that $x := ap$ has the desired property i.e. that

$$\forall\, i \in I, v \in V\colon\ (a_i, x)_v = \varepsilon_{i,v}.$$

We again distinguish different cases:

Case 1.1 ($v \in S$). We have that $\varepsilon_{i,v} = 1$ since $S \cap T = \varnothing$ and we have to check that $(a_i, x)_v = 1$. If $v = \infty$ this follows from $x > 0$. So let $v \in P$. This means that $x \equiv a^2 \pmod{m}$, hence

$$x \equiv \begin{cases} a^2 \pmod 8 & \text{if } v = 2 \\ a^2 \pmod v & \text{else.} \end{cases}$$

So, since x and a are v-adic units, x is a square in $\mathbb{Q}_v^*$ by Theorem 2.111 and following. So we have $(a_i, x)_v = 1$.

Case 1.2 ($v \notin S$). In this case, a_i is a v-adic unit and calculate the Hilbert symbol with Theorem 3.11.

$$\forall\, b \in \mathbb{Q}_v^*\colon\ (a_i, b)_v = \left(\frac{a_i}{v}\right)^{\operatorname{ord}_v(b)} \tag{3.3.2}$$

since $v \neq 2$.

Case 1.2.1 ($v \notin T \cup \{p\}$). x is a v-adic unit, hence $\operatorname{ord}_v(x) = 0$ and (3.3.2) yields $(a_i, x)_v = 1$. But since $v \notin T$ we have $\varepsilon_{i,v} = 1$.

Case 1.2.2 ($v \in T$). This means that $\operatorname{ord}_v(x) = 1$ and (iii) yields that $x_v \in \mathbb{Q}_v^*$ such that $\forall\, i \in I\colon\ (a_i, x_v)_v = \varepsilon_{i,v}$. Since $v \in T$ one of the $\varepsilon_{i,v}$ is -1 and we have $\operatorname{ord}_v(x_v) \equiv 1 \pmod 2$ and hence

$$\forall\, i \in I\colon\ (a_i, x)_v = \left(\frac{a_i}{v}\right) = (a_i, x_v)_v = \varepsilon_{i,v}.$$

Case 1.2.3 ($v = p$). We reduce this case to the preceding ones by the Hilbert Theorem (Theorem 3.20):

$$(a_i, x)_p \overset{3.20}{=} \prod_{v \neq p} (a_i, x)_v = \prod_{v \neq p} \varepsilon_{i,v} = \varepsilon_{i,p}.$$

Case 2 ($S \cap T \neq \varnothing$). Since the square $[\mathbb{Q}_v^*]^2$ form an open subgroup of $\mathbb{Q}_v^*$ by Lemma 3.21 there exists $x' \in \mathbb{Q}^*$ such that

$$\forall\, v \in V\colon \frac{x'}{x_v} \in [\mathbb{Q}_v^*]^2.$$

In particular this means that

$$\forall\, v \in S\colon\ (a_i, x')_v = (a_i, x_v)_v = \varepsilon_{i,v}.$$

Now set

$$\forall\, i \in I, v \in V\colon \eta_{i,v} := \varepsilon_{i,v} \left(a_i, x'\right)_v .$$

The family $(\eta_{i,v})$ satisfies condition (i) to (iii) and moreover $\forall\, v \in S\colon \eta_{i,v} = 1$. Be the preceding case, there exists $y \in \mathbb{Q}^*$ such that

$$\forall\, i \in I, v \in V\colon \ (a_i, y)_v = \eta_{i,v}.$$

The element $x := yx'$ satisfies the desired properties. □

3.4 Exponential Upper bound

In this section we describe the classical way to calculate the Hilbert symbol according to a method by Legendre with elementary methods (even avoiding p-adic numbers). For this, we will explicitly construct a solution to a quadratic diagonal equation in three variables (if it exists). This will lead to an exponential running time, an upper bound for the complexity of the HilbertSymbol$_{\mathbb{Q}}$ with the additional restriction to integral instances. In Section 3.5 we will improve this upper bound by using a different approach that uses an oracle for IntFact.

Notation 3.23 (Square decomposition). For $a \in \mathbb{Z}$, denote by $\widetilde{a} \in \mathbb{N}$ the maximal number such that for some $\overline{a} \in \mathbb{Z}$ we can write

$$a = \widetilde{a}^2 \overline{a}.$$

Fact 3.24. *For $a \in \mathbb{Z}$, $\overline{a}$ is squarefree.*

Lemma 3.25 (Rational Equation to Integral Equation). *For every $\alpha = \frac{\alpha_1}{\alpha_2}, \beta = \frac{\beta_1}{\beta_2}, \gamma = \frac{\gamma_1}{\gamma_2} \in \mathbb{Q}, \gamma \neq 0$ there exist squarefree $a, b \in \mathbb{Z}$ such that*

$$\alpha x^2 + \beta y^2 = \gamma \tag{3.4.1}$$

is solvable over rationals if and only if

$$ax^2 + by^2 = z^2 \tag{3.4.2}$$

is solvable over integers with pairwise coprime x, y, z.
In great detail we have with

$$A := \alpha_1\beta_2\gamma_2 \qquad B := \alpha_2\beta_1\gamma_2 \qquad C := \alpha_2\beta_2\gamma_1$$

that

$$a = \overline{A} \cdot \overline{C} \qquad b = \overline{B} \cdot \overline{C}$$

such that one can obtain a solution $(x, y) = \left(\frac{u}{w}, \frac{v}{w}\right) \in \mathbb{Q}^2$ of Equation (3.4.1) from a solution $(\hat{x}, \hat{y}, \hat{z}) \in \mathbb{Z}^3$ of Equation (3.4.2) and vice versa by the following relations:

$$x = \frac{\hat{x}\overline{C}\widetilde{C}}{\hat{z}\widetilde{A}} \qquad y = \frac{\hat{y}\overline{C}\widetilde{C}}{\hat{z}\widetilde{B}}$$

$$\hat{x} = \frac{\widetilde{A}}{\widetilde{C}}u \qquad \hat{y} = \frac{\widetilde{B}}{\widetilde{C}}v \qquad \hat{z} = \overline{C}w.$$

Proof. Equation (3.4.1) is solvable over rationals if and only if

$$Ax^2 + By^2 = C$$

is solvable over rationals. The solvability of the last equation is equivalent to the solvability of the following equation:

$$\overline{A}\left(\frac{\widetilde{A}}{\widetilde{C}}x\right)^2 + \overline{B}\left(\frac{\widetilde{B}}{\widetilde{C}}y\right)^2 = \overline{C}. \tag{3.4.3}$$

Hence with the substitution $\hat{x} := \frac{\widetilde{A}}{\widetilde{C}}x$ and $\hat{y} := \frac{\widetilde{B}}{\widetilde{C}}y$ we want to solve

$$\overline{A}\hat{x}^2 + \overline{B}\hat{y}^2 = \overline{C} \tag{3.4.4}$$

over rationals. Use homogenization to switch to integers: Having a solution for Equation (3.4.4) means that there exists $u, v, w \in \mathbb{Z}$ with $w \neq 0$ such that

$$\overline{A}\left(\frac{u}{w}\right)^2 + \overline{B}\left(\frac{v}{w}\right)^2 = \overline{C}$$

or, equivalently

$$\overline{A}u^2 + \overline{B}v^2 = \overline{C}w^2.$$

Multiplying by $\overline{C} \neq 0$ yields

$$(\overline{C} \cdot \overline{A})u^2 + (\overline{C} \cdot \overline{B})v^2 = \overline{C}^2 w^2. \tag{3.4.5}$$

Now Equation (3.4.5) has an integer solution with $w \neq 0$ if and only if

$$(\overline{C} \cdot \overline{A})u^2 + (\overline{C} \cdot \overline{B})v^2 = z^2$$

has an integer solution with $z \neq 0$. Since $\overline{C}(\overline{a}u^2 + \beta v^2) = z^2$ we have that $\overline{C} \mid z^2$ and since $\overline{C}$ is squarefree $\overline{C} \mid z$ and therefore $w^2 = \frac{z^2}{\overline{C}^2} \in \mathbb{Z}$. □

Proposition 3.26. *For $x, y, z, d \in \mathbb{Z}$ with $x + y = z$ we have: If d divides two elements of the set $\{x, y, z\}$, then d divides all three elements of $\{x, y, z\}$.*

Proof. Without loss of generality assume that $d \mid x$ and $d \mid y$. Then there exist elements $u, v \in \mathbb{Z}$ such that $x = du$ and $y = dv$, hence we have:

$$z = du + dv = d(u + v) \quad \Rightarrow \quad d \mid z.$$ □

Proposition 3.27. *If a prime divides a product of integrals, it divides at least one of the factors.*

Proof. Let p be a prime and $a, b \in \mathbb{Z}$ with $p \mid ab$. We want to prove:

$$p \nmid a \Rightarrow p \mid b.$$

Set $g := \gcd(a, p)$. Then of course $g \mid p$. Since p is prime, we have that $g = 1$ or $g = p$. If $g = p$, since $g \mid a$ too, $p \mid a$ which is a contradiction. So $g = 1$. By the euclidean algorithm, there exist

$$x, y \in \mathbb{Z} : px + ay = 1.$$

Multiplying by b gives:

$$bpx + bay = b.$$

Observe that $p \mid pxb$ and $p \mid aby$ (since it is assumed that $p \mid ab$). Therefore, by Proposition 3.26 it follows that $p \mid b$. □

Definition 3.28 (Norm of an element in a number field). Note that for $a \in \mathbb{Q}$ an element of the number field $\mathbb{Q}[\sqrt{a}]$ can be written as $\alpha + \beta\sqrt{a}$ with $\alpha, \beta \in \mathbb{Q}$. Now define the **norm** by

$$\begin{array}{rcl} N : \mathbb{Q}[\sqrt{a}] & \longrightarrow & \mathbb{Q} \\ \alpha + \beta\sqrt{a} & \longmapsto & \alpha^2 - a\beta^2. \end{array}$$

Fact 3.29. *The norm defined above is a multiplicative function.*

Proof. One can note that the map above is a norm as in Section 2.7 and by Fact 2.63 therefore is multiplicative, or explicitly calculated with $\alpha + \beta\sqrt{a}, \alpha' + \beta'\sqrt{a} \in \mathbb{Q}[\sqrt{a}]$:

$$\begin{aligned} & N\left(\alpha + \beta\sqrt{a}\right) N\left(\alpha' + \beta'\sqrt{a}\right) = \left(\alpha^2 - a\beta^2\right)\left(\alpha'^2 - a\beta'^2\right) \\ &= \alpha^2\alpha'^2 - a\alpha^2\beta'^2 - a\beta^2\alpha'^2 + a^2\beta^2\beta'^2 = \alpha^2\alpha'^2 + a^2\beta^2\beta'^2 - a\left(\alpha^2\beta'^2 + \beta^2\alpha'^2\right) \\ &= \alpha^2\alpha'^2 - 2\alpha\alpha'\beta\beta'a + a^2\beta^2\beta'^2 - a\left(\alpha^2\beta'^2 - 2\alpha\beta'\alpha'\beta + \alpha'^2\beta^2\right) \\ &= \left(\alpha\alpha' + a\beta\beta'\right)^2 - a\left(\alpha\beta' + \alpha'\beta\right)^2 = N\left(\alpha\alpha' + a\beta\beta' + (\alpha\beta' + \alpha'\beta)\sqrt{a}\right) \\ &= N\left(\left(\alpha + \beta\sqrt{a}\right)\left(\alpha' + \beta'\sqrt{a}\right)\right). \end{aligned}$$

□

Lemma 3.30 (Integral Equation Reduction). *Let $a, b \in \mathbb{Z}$ be squarefree with $|a| < |b|$ and $1 < |b|$. Then there exists $b' \in \mathbb{Z}$ with $|b'| < |b|$ such that*

$$ax^2 + by^2 = z^2 \tag{3.4.6}$$

has a solution if and only if

$$ax^2 + b'y^2 = z^2 \tag{3.4.7}$$

has a solution and these solutions can be converted into each other in a constant number of operations.

Proof. If Equation (3.4.6) has a solution, then for any $p \mid b$ we have that p cannot divide x, since otherwise:

$$\begin{array}{ccccccc} & p \mid b \text{ and } p \mid x & \Rightarrow & 0 \equiv_p ax^2 + by^2 = z^2 & \Rightarrow & p \mid z^2 \\ \overset{3.27}{\Rightarrow} & p \mid z & \Rightarrow & p^2 \mid z^2 = ax^2 + by^2 & \overset{p \mid x}{\Rightarrow} & p^2 \mid by^2 \end{array}$$

This means that $\exists\, n \in \mathbb{Z} : np^2 = by^2$. We furthermore know that $p \mid b$, meaning that $\exists\, m \in \mathbb{Z} : mp = b$ with the additional property that $p \nmid m$ since b is squarefree. Putting this together, we get

$$\begin{array}{ccccccc} & np^2 = by^2 = mpy^2 & \Leftrightarrow & np = my^2 & \Rightarrow & p \mid my^2 \\ \overset{3.27}{\Rightarrow} & p \mid m \text{ or } p \mid y^2 & \overset{p \nmid m}{\Rightarrow} & p \mid y^2 & \overset{3.27}{\Rightarrow} & p \mid y \end{array}$$

which is a contradiction since x, y, z were assumed to be coprime. So $p \nmid x$ which means that $x \in \mathbb{F}_p^*$ is invertible and therefore we get

$$z^2 = ax^2 + by^2 \equiv_p ax^2 \quad \Leftrightarrow \quad z^2 \left(x^{-1}\right)^2 \equiv_p a$$

i.e. a is a square modulo p. By the Chinese Remainder Theorem $a \in \mathbb{Z}/(b)$ is a square i.e. there is a $t \in \mathbb{Z}$ such that $|t| \leq \frac{|b|}{2}$ and $a \equiv_b t^2$. Let $b' \in \mathbb{Z}$ be such that

$$t^2 = a + bb'. \tag{3.4.8}$$

Now we prove that $ax^2 + by^2 = z^2$ has a solution if and only if $ax^2 + b'y^2 = z^2$ has a solution: If $ax^2 + by^2 = z^2$ has a solution then

$$\begin{array}{rcrcl} N\left(\frac{z+x\sqrt{a}}{y}\right) = \frac{z^2}{y^2} - a\frac{x^2}{y^2} & \iff & y^2 N\left(\frac{z+x\sqrt{a}}{y}\right) = z^2 - ax^2 \\ \iff \quad ax^2 + y^2 N\left(\frac{z+x\sqrt{a}}{y}\right) = z^2 & \implies & N\left(\frac{z+x\sqrt{a}}{y}\right) = b. \end{array}$$

From Equation (3.4.8) we furthermore get:

$$\begin{aligned} & bb' = t^2 - a = N\left(t + \sqrt{a}\right) \\ \implies & b' = \frac{N\left(t+\sqrt{a}\right)}{b} = \frac{N\left(t+\sqrt{a}\right)}{N\left(\frac{z+x\sqrt{a}}{y}\right)} \overset{\text{Fact 3.29}}{=} N\left(\frac{yt + y\sqrt{a}}{z + x\sqrt{a}}\right). \end{aligned}$$

Which, by expanding the fraction by $z - x\sqrt{a}$ i.e. by rationalizing the denominator, effectively gives an integral solution of eq. (3.4.7). Since the argument is symmetric in b and b' we get a solution of eq. (3.4.6) from a solution of eq. (3.4.7).

We are left to show that $|b'| < |b|$:

$$\begin{array}{ccccccc} & |a| + |b'| & = & |a| + \left|\frac{t^2-a}{b}\right| & \leq & |a| + \left|\frac{t^2}{b}\right| + \left|\frac{a}{b}\right| & \\ \overset{|a|<|b|}{\leq} & |a| + \left|\frac{t^2}{b}\right| + 1 & \overset{|t|\leq\frac{|b|}{2}}{\leq} & |a| + \frac{|b|}{4} + 1 & \overset{1<|b|}{<} & |a| + |b| & \end{array}$$

□

Corollary 3.31. *Let $a, b \in \mathbb{Z}$ be squarefree with $|a| < |b|$ and $1 < |b|$. Then determining if a solution of*

$$ax^2 + by^2 = z^2$$

exists and calculating it, can be done in $O(|a| + |b|)$ steps, which is exponential in the bitsizes of a and b.

Proof. Repeatedly apply the reduction in Lemma 3.30 to end up with one of the following equations:

$$\pm x^2 \pm y^2 = z^2 \text{ or } \pm x^2 = z^2.$$

where x, y, z (resp. x, z) are coprime. In detail we have:

$x^2 + y^2 = z^2$	has $(1, 0, 1)$ as a solution.
$x^2 - y^2 = z^2$	has $(1, 0, 1)$ as a solution.
$-x^2 + y^2 = z^2$	has $(0, 1, 1)$ as a solution.
$-x^2 - y^2 = z^2$	has $(0, 0, 0)$ as the only possible solution but 0 is not coprime to 0 so it has no solution with pairwise coprime x, y and z.
$x^2 = z^2$	has $(1, 1)$ as a solution.
$-x^2 = z^2$	has $(0, 0)$ as the only possible solution but 0 is not coprime to 0 so it has no solution with coprime x and y.

By Lemma 3.30 in every step either $|a|$ or $|b|$ is reduced by at least 1 each of which needs constant time. So we need at most $|a| + |b|$ steps. □

One could summarize the result of this section as

$$\textsc{HilbertSymbol}^*_{\mathbb{Q}} \in \mathbf{EXP}$$

Where $\textsc{HilbertSymbol}^*_{\mathbb{Q}}$ denotes the special case of $\textsc{HilbertSymbol}_{\mathbb{Q}}$ where only integral instances are allowed:

$$\textsc{HilbertSymbol}^*_{\mathbb{Q}} := \left\{ (a, b) \in (\mathbb{Z} \setminus \{ 0 \})^2 \;\middle|\; \mathrm{h}_{\mathbb{Q}}(a, b) = 1 \right\}.$$

3.5 Upper bound using an oracle for IntFact

In this section, we will see that, using an oracle for IntFact, one can calculate the Hilbert symbol in polynomial time, ultimately yielding that its complexity is in $\mathbf{NP} \cap \mathbf{coNP}$.

Algorithm 6 will show that the hardest step in deciding quadratic rational form equivalence is to decide for $a, b \in \mathbb{Q}^*$ the nontrivial integral solvability of

$$aX^2 + bY^2 = Z^2.$$

More precisely, we have

$$\textsc{QuadFormEquiv}_{\mathbb{Q}} \in \mathbf{P}^{\textsc{HilbertSymbol}_{\mathbb{Q}}}$$

This means the complexity of the Hilbert symbol is of great interest for the complexity of the rational quadratic form equivalence problem. In this section, we will see that

$$\textsc{HilbertSymbol}_{\mathbb{Q}} \in \mathbf{P}^{\textsc{IntFact}}$$

which ultimately will give

$$\textsc{QuadFormEquiv}_{\mathbb{Q}} \in \mathbf{P}^{\textsc{IntFact}}$$

a main result of this work. Remember that FuncIntFact is as hard as IntFact by Theorem 2.34 so whenever we are allowed to use an oracle for IntFact, we may also use an oracle for FuncIntFact i.e. obtain the prime factorization of an integral number.

The classical way of finding a solution for a Hilbert equation is described in Section 3.4. But this methods, despite the fact that it only allows integral instances, leads to an upper bound that is exponential in the bit sizes of a and b. We will now improve this complexity bound by using an oracle for IntFact.

Armed with Theorem 3.17, the following algorithm decides the solvability of

$$aX^2 + bY^2 = Z^2 \quad \text{for } a, b \in \mathbb{Q}^*$$

efficiently when given an oracle for IntFact.

Algorithm 4 Hilbert-Symbol

Input: $a, b \in \mathbb{Q}^*$.
Output:
true if $ax^2 + by^2 = z^2$ has a nontrivial solution in $\mathbb{Q}^3$
false otherwise.

1: **calculate** $\mathcal{H}_{a,b}$ using Theorem 3.17.
2: **assert** $\mathcal{H}_{a,b} = \varnothing$
3: **return true**

Lemma 3.32. *Algorithm 4 is correct.*

Proof. $\mathcal{H}_{a,b}$ is the set of $v \in V$ such that the equation $aX^2+bY^2 = Z^2$ has a nontrivial solution over $\mathbb{Q}_v$. The famous Local-Global-Principle / Hasse-Minkowski Theorem (Theorem 4.78) for quadratic forms applied to $aX^2 + bY^2 - Z^2$, yields that the equation does not have a solution over $\mathbb{Q}$ if and only if for all $v \in V$ it has no solution over $\mathbb{Q}_v$.

We will see a proof of Local-Global-Principle / Hasse-Minkowski Theorem in Section 4.8. The reader can check that we will not use this lemma or any of its corollaries for the proof of Theorem 4.78. □

Lemma 3.33. *Using an oracle for* IntFact*, Algorithm 4 runs in polynomial time.*

Proof. This is clear by Theorem 3.17. □

We summarize Lemmas 3.32 and 3.33 in the following theorem:

Theorem 3.34.

$$\textsc{HilbertSymbol}_{\mathbb{Q}} \in \mathbf{P}^{\textsc{IntFact}}$$

Corollary 3.35.

$$\textsc{HilbertSymbol}_{\mathbb{Q}} \in \mathbf{NP} \cap \mathbf{coNP}.$$

Proof. This is clear since

$$\textsc{IntFact} \in \mathbf{NP} \cap \mathbf{coNP}.$$ □

3.6 Lower bound

In this section, we will establish a lower bound for HilbertSymbol$_{\mathbb{Q}}$: We will show that an oracle for HilbertSymbol$_{\mathbb{Q}}$ yields an efficient algorithm to decide QuadResidue for the following special case:

- The modulus n is a composite number of two primes p and q for which it holds that $p, q \equiv 1 \pmod 4$
- The set $\{ p, q \}$ "quadratically agrees" on the instances r:

Definition 3.36 (Quadratic agreement). Let $r \in \mathbb{Z}_{>0}$ be a natural number and $Q \subseteq \mathbb{Z}_{>0}$ be a finite set of natural numbers. We say that Q **quadratically agrees on** r and write $Q \vdash r$ if every factor of r is either a quadratic residue modulo every or no element of Q. More formally write $Q = \{ n_1, \ldots, n_l \} \subseteq \mathbb{Z}_{>0}$ and the prime factorization of r:

$$r = \prod_{i=1}^{k} p_i \quad k \in \mathbb{N}, \forall\, i \in [1:k] \colon p_i \in P.$$

Then Q quadratically agrees on r if and only if for every $i \in [1:k]$ it holds that

$$\forall\, j \in [1:l] \colon p_i \mathrm{R} n_j \quad \text{or} \quad \forall\, j \in [1:l] \colon p_i \mathrm{N} n_j.$$

We denote this special case by QuadResidue*. QuadResidue in general is, without IntFact, believed to be a hard problem, see for example Problem 11 in [AM]. Its functional version, i.e. taking square roots modulo n, yields an efficient randomized algorithm for IntFact as shown in Section 5.3.2. This is also a hint to the hardness of QuadResidue.

The assumptions in the following lemmas seem very artifical. This is because their sole purpose is to successively establish conditions for the HilbertSymbol$_{\mathbb{Q}}$ oracle algorithm that decides QuadResidue*.

Lemma 3.37. *For distinct odd primes p and q define $n := pq$. For every $r \in \mathbb{Z}_{>0}$ with $\gcd(r, n) = 1$ and $\{ p, q \} \vdash r$ the following statements hold:*

(i). $(r, n)_p = \left(\frac{r}{p}\right)$.

(ii). if additionally $p, q \equiv 1 \pmod 4$ it holds that

$$\mathrm{h}_{\mathbb{Q}}(r, n) = 1 \quad \Leftrightarrow \quad (r, n)_p = 1 \text{ and } (r, n)_q = 1.$$

Proof.

(i). We compute the value with Theorem 3.11:

$$(r, n)_p \overset{3.13}{=} (r, p)_p (r, q)_p = \left(p^0 r, p^1 \cdot 1\right)_p \left(p^0 r, p^0 q\right)_p \overset{3.11}{=} \left(\frac{r}{p}\right) \cdot 1.$$

(ii). First observe that by the Local-Global-Principle / Hasse-Minkowski Theorem (Theorem 4.78) applied to the form $rX^2 + nY^2 - Z^2$, we have that $\mathrm{h}_{\mathbb{Q}}(r, n) = 1$ if and only if

$$\forall\, v \in V \colon (r, n)_v = 1.$$

So the implication "⇒" is clear since $p, q \in V$. For the other direction write down the prime factorization of r:

$$r = \prod_{i=1}^{k} p_i \quad k \in \mathbb{N}, \forall\, i \in [1:k] \colon p_i \in P.$$

We can calculate the value of $(r,n)_v$ for any $v \in V$ with Theorem 3.11:

Case 1 ($v = \infty$). Since $r, n \geq 0$, we have $(r,n)_\infty \overset{3.11}{=} 1$.

Case 2 ($v = 2$). $(r,n)_2 \overset{3.11}{=} (-1)^{\epsilon(r)\,\epsilon(n)} = 1$ since $p, q \equiv 1 \pmod 4$ which implies $\epsilon(n) = \epsilon(pq) = \epsilon(1) = 0$.

Case 3 ($v \in \{p, q\}$). By assumption $(r,n)_p = (r,n)_q = 1$.

Case 4 ($v \notin \{p_i\}_{i=1}^k \cup \{\infty, 2, p, q\}$). $(r,n)_v = \left(v^0 r, v^0 n\right)_v \overset{3.11}{=} 1$.

Case 5 ($v \in \{p_i\}_{i=1}^k$). Let $i \in [1:k]$ such that $v = p_i$ and calculate

$$(n,r)_{p_i} = \left(pq, \prod\nolimits_{j=1}^{k} p_j\right)_{p_i} \overset{3.13}{=} \prod_{j=1}^{k} (p, p_j)_{p_i} (q, p_j)_{p_i} = \left(\frac{p}{p_i}\right)\left(\frac{q}{p_i}\right)$$

where the last equality holds by Theorem 3.11: If $j \neq i$, we have

$$\left(p_i^0 p, p_i^0 p_j\right)_{p_i} \left(p_i^0 q, p_i^0 p_j\right)_{p_i} = 1$$

and if $i = j$, we have

$$\left(p_i^0 p, p_i^1 \cdot 1\right)_{p_i} \left(p_i^0 q, p_i^1 \cdot 1\right)_{p_i} = \left(\frac{p}{p_i}\right)\left(\frac{q}{p_i}\right).$$

By the Quadratic reciprocity law (Theorem 2.57) we then get

$$\left(\frac{p}{p_i}\right)\left(\frac{q}{p_i}\right) \overset{2.57}{=} (-1)^{\epsilon(p)\,\epsilon(p_i)} \left(\frac{p_i}{p}\right) (-1)^{\epsilon(q)\,\epsilon(p_i)} \left(\frac{p_i}{q}\right) = \left(\frac{p_i}{p}\right)\left(\frac{p_i}{q}\right)$$

where the last equality follows from the fact that $p, q \equiv 1 \pmod 4$ which implies that $\epsilon(p)$ and $\epsilon(q)$ are even. So in total we have

$$(n,r)_{p_i} = \left(\frac{p_i}{p}\right)\left(\frac{p_i}{q}\right).$$

That means $(n,r)_{p_i} = 1$ if and only if $\left(\frac{p_i}{p}\right) = \left(\frac{p_i}{q}\right)$, which is true by the assumption that $\{p, q\}$ quadratically agrees on r.

Now we have shown that $\forall\, v \in V \colon (r,n)_v = 1$ which, by Theorem 4.78, implies that $\mathrm{h}_{\mathbb{Q}}(r,n) = 1$. □

Lemma 3.38. *For distinct odd primes p and q and every $r \in \mathbb{Z}$ we have that*

$$r \mathrm{R} pq \quad \Leftrightarrow \quad r \mathrm{R} p \text{ and } r \mathrm{R} q$$

or in other words

$$r \mathrm{R} pq \quad \Leftrightarrow \quad \left(\frac{r}{p}\right) = 1 \text{ and } \left(\frac{r}{q}\right) = 1.$$

Proof. This directly follows from the Chinese Remainder Theorem: Since p and q are distinct, we have that

$$\mathbb{Z}/pq\mathbb{Z} \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}\,.$$

This implies that $r\mathrm{R}pq$ if and only if $r\mathrm{R}p$ and $r\mathrm{R}q$, and the latter two statements are equivalent to $\left(\frac{r}{p}\right) = 1$ and $\left(\frac{r}{q}\right) = 1$. □

Lemma 3.39. *For $p \in P \setminus \{2\}$ and $r \in [1 : 2p-1]$ we have that*

$$r\mathrm{R}2p \quad \Leftrightarrow \quad r\mathrm{R}p.$$

Proof. Again by the Chinese Remainder Theorem: Since $p \neq 2$, we have that

$$\mathbb{Z}/2p\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}\,.$$

This implies that $r\mathrm{R}2p$ if and only if $r\mathrm{R}2$ and $r\mathrm{R}p$. But $r\mathrm{R}2$ is always true. □

Lemma 3.40. *For $p \in P \setminus \{2\}$ and $r \in \left[1 : p^2-1\right]$ with $\gcd(p,r) = 1$ it holds that*

$$r\mathrm{R}p \Leftrightarrow r\mathrm{R}p^2.$$

Proof.

Case 1 ("⇐"). This is elementary: By definition $r\mathrm{R}p^2$ means that

$$\begin{aligned} & \exists\, x \in \mathbb{Z}/p^2\mathbb{Z} : x^2 \equiv r \pmod{p^2} \\ \Leftrightarrow\ & \exists\, x \in \mathbb{Z}/p^2\mathbb{Z}\,, k \in \mathbb{Z} \colon x^2 = r + (kp)p \\ \Rightarrow\ & \exists\, x \in \mathbb{Z}/p\mathbb{Z} : x^2 \equiv r \pmod{p} \end{aligned}$$

which is the definition of $r\mathrm{R}p$.

Case 2 ("⇒"). This is an application of Hensel's lemma for modular arithmetic (Lemma 2.99): We simply lift a root modulo p to a root modulo p^2. In detail, we define $f = X^2 - r$. If $r\mathrm{R}p$ we have that

$$\exists\, x \in \mathbb{Z}/p\mathbb{Z} : f(x) \equiv 0 \pmod{p}.$$

But since $\gcd(r,p) = 1$ we have that $r \not\equiv 0 \pmod{p}$ and since $x^2 \equiv r \pmod{p}$ it follows that $x \not\equiv 0 \pmod{p}$. With $p \neq 2$ this implies $f'(x) = 2x \not\equiv 0 \pmod{p}$. So the conditions for Lemma 2.99 are met and we find a $y \in \mathbb{Z}$ such that $f(y) \equiv 0 \pmod{p^2}$ which means that $r\mathrm{R}p^2$. □

I recommend reading the proof of the following algorithm in parallel to the algorithm itself.

Algorithm 5 QuadraticResidue-With-HilbertSymbol-Oracle

Input: $n \in \mathbb{N}$ with $n = pq$, $p, q \in P$ and $p, q \equiv 1 \pmod 4$, $r \in [1:n]$ with $\{p, q\} \vdash r$

Output: **true** if r is a quadratic residue modulo n and **false** else.

1: **if** $r = n$ **then**
2: **return true**
3: **end if**
4: **if** $2 \mid n$ **then**
5: **if** $n = 4$ **then**
6: **return** $r \equiv 1 \pmod 2$
7: **else**
8: **return** $\left(\frac{r}{\frac{n}{2}}\right) = 1$
9: **end if**
10: **end if**
11: $g := \gcd(r, n)$.
12: **if** $g \neq 1$ **then**
13: **return** $\left(\left(\frac{r}{g}\right) = 1 \text{ and } \left(\frac{r}{\frac{n}{g}}\right) = 1\right)$
14: **end if**
15: **if** $\sqrt{n} \in \mathbb{Z}$ **then**
16: **return** $\left(\frac{r}{\sqrt{n}}\right) = 1$.
17: **end if**
18: **return** $\mathrm{h}_{\mathbb{Q}}(r, n) = 1$

Theorem 3.41. *Algorithm 5 is correct and runs in polynomial time in the bit-sizes of the input given an oracle for* HilbertSymbol$_{\mathbb{Q}}$.

Proof. Every step in the algorithm, except for the last one, establishes a situation where we can use Lemma 3.37.

Steps 1 to 3 If $r = n$ we have that $r \equiv 0 \pmod n$ meaning that it is trivially a square modulo n. So after this step we are in the situation that $1 \leq r < n$.

Steps 4 to 10 Because of

$$1^2 \equiv 1 \pmod 4, \quad 2^2 \equiv 0 \pmod 4, \quad 3^2 \equiv 1 \pmod 4$$

we know that the only nonzero square in $\mathbb{Z}/4\mathbb{Z}$ is 1, justifying step 6. If $n \neq 4$ Lemma 3.39 and the fact that $r \mathrm{R} p \Leftrightarrow \left(\frac{r}{p}\right) = 1$ yield the correctness of step 8

Steps 12 to 14 In this case, we have found a factor of n, which makes everything very simple: We have that $g \neq 1, g \mid r$ and $g \mid n$. By steps 1 to 3 we have that $r \neq n$ in addition to $n = pq$, which together implies that $g \in \{p, q\}$. So g and $\frac{n}{g}$ are the two factors of n. Then we can apply Lemma 3.38, if these two factors are distinct, and Lemma 3.40, if they are even, to obtain that we can simply check $\left(\frac{r}{g}\right) = 1$ and $\left(\frac{g}{\frac{n}{g}}\right) = 1$ to check if $r \mathrm{R} n$. This is possible in polynomial time by Theorem 2.56. After these steps, we have that $\gcd(r, n) = 1$.

Steps 15 to 17 Remember that $\sqrt{n} \in \mathbb{Z}$ can be checked in polynomial time by Algorithm 3. Furthermore, $\sqrt{n} \in \mathbb{Z}$ is equivalent to $p = q$ which means that $p = q = \sqrt{n} \in P$. Then the question of whether r is a quadratic residue modulo n can be answered by calculating $\left(\frac{r}{\sqrt{n}}\right)$: Since $1 \leq r < n$ we can apply Lemma 3.40. By Theorem 2.56 we know that $\left(\frac{r}{\sqrt{n}}\right)$ can be calculated in polynomial time. After these steps, we are in the situation that $p \neq q$.

Step 18 The oracle for HilbertSymbol$_{\mathbb{Q}}$ is called and Lemma 3.37(ii) yields the correctness of the algorithm. □

Chapter 4

Classification of Quadratic Forms

Abstract

The goal of this chapter is to understand all preliminaries required for designing an algorithm which outputs a quadratic form equivalence. This case is significantly easier than equivalences for higher degree forms, because quadratic modules, which are closely connected to quadratic forms, are well-studied and have a lot of structure. We will define a quadratic module associated to a quadratic form and the isomorphy classes of these modules will correspond to the equivalence classes of the associated forms. Ultimately we will prove Witt's theorem yielding two useful corollaries:

(i). Every quadratic form is equivalent to a form $\sum_{i=1}^{n} a_i X_i^2$.

(ii). We have a cancellation rule for quadratic forms (the $\ominus$ will be defined on the way):

$$f \ominus h \sim g \ominus h \quad \Longrightarrow \quad f \sim g$$

This chapter is based on [Ser73].

4.1 Definitions

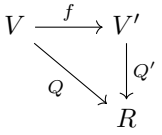
Definition 4.1 (The category of quadratic modules). Let V be a module over a commutative ring R. A function $Q : V \to R$ is called a quadratic form on V if the following two conditions hold:

(i). $\forall\, a \in R, x \in V \colon Q(ax) = a^2 Q(x)$.

(ii). $\Theta_Q \colon \begin{array}{ccc} V \times V & \longrightarrow & R \\ (x,y) & \longmapsto & Q(x+y) - Q(x) - Q(y) \end{array}$ is bilinear.

The pair (V, Q) is called a **quadratic module**. For a fixed ring, the set of all quadratic forms on V is denoted by $\mathrm{Quad}(V)$. Let (V', Q') be another quadratic module. A linear map $f : V \to V'$ is called a **morphism of quadratic**

modules or **metric morphism** if $Q' \circ f = Q$, which means the following diagram commutes:



We write $f : (V, Q) \to (V', Q')$ if f is a morphism of quadratic modules.

Remark 4.2. A form $\phi : V^2 \to R$ is called bilinear if the following two conditions hold:

(i). $\forall\, a, b, c, d \in V : \phi(a + b, c + d) = \phi(a, c) + \phi(a, d) + \phi(b, c) + \phi(b, d)$

(ii). $\forall\, \lambda \in R, a, b \in V : \phi(\lambda a, b) = \lambda\phi(a, b) = \phi(a, \lambda b)$

In our situation, the ring R will always be a field $\mathbb{F}$ with $\mathrm{char}\,(\mathbb{F}) \neq 2$ and the module V will therefore be a vector space. We will furthermore assume that V is finite-dimensional.

Definition 4.3. Let (V, Q) be a quadratic vector space, define $\forall\, x, y \in V$:

$$x.y := \frac{\Theta_Q(x, y)}{2}.$$

Definition/Proposition 4.4. *Let (V, Q) be a quadratic vector space. Then:*

*(i). $(x, y) \mapsto x.y$ is a symmetric bilinear form on V called the **scalar product associated to** Q.*

(ii). $\forall\, x \in V : Q(x) = x.x$, and therefore there is a bijective correspondence between quadratic forms and symmetric bilinear forms:

$$\begin{array}{rcl} \{\, Q : V \to R \mid Q \text{ quadratic form} \} & \longleftrightarrow & \left\{\, \phi \colon V^2 \to R \;\middle|\; \phi \text{ bilinear} \right\} \\ Q & \overset{f}{\longmapsto} & \frac{\Theta_Q}{2} \\ (x \mapsto \phi\,(x, x)) & \overset{g}{\longmapsto} & \phi. \end{array}$$

(iii). For a metric morphism $f : (V, Q) \to (V', Q')$ it holds that $\forall\, x, y \in V : f(x).f(y) = x.y$.

Proof.

(i). Symmetry is obvious and linearity follows from the properties of Q.

(ii). Note that for every $x \in V$:

$$x.x = \frac{Q(2x) - Q(x) - Q(x)}{2} = \frac{4Q(x) - Q(x) - Q(x)}{2} = Q(x), \quad (4.1.1)$$

and calculate for a quadratic form Q and an arbitrary $x \in V$:

$$g(f(Q))(x) = g\left(\frac{\Theta_Q}{2}\right)(x) = \frac{\Theta_Q(x, x)}{2} \overset{4.3}{=} x.x \overset{(4.1.1)}{=} Q(x)$$

For the other direction let $\phi : V^2 \to R$ be bilinear and define

$$\begin{array}{rccc} P\colon & V & \longrightarrow & R \\ & x & \longmapsto & \phi(x,x)\,. \end{array}$$

Let now $x, y \in V$ be arbitrary and simply calculate

$$\begin{aligned} f(g(\phi))(x,y) = f(P)(x,y) = \frac{\Theta_P}{2}(x,y) &= \frac{P(x+y) - P(x) - P(y)}{2} \\ &= \frac{\phi(x+y,x+y) - \phi(x,x) - \phi(y,y)}{2} \\ &= \frac{\phi(x,x) + \phi(x,y) + \phi(y,x) + \phi(y,y) - \phi(x,x) - \phi(y,y)}{2} \\ &= \frac{\phi(x,y) + \phi(y,x)}{2} = \phi(x,y) \end{aligned}$$

(iii). This follows directly from $Q' \circ f = Q$. □

Remark 4.5. Even though $(x,y) \mapsto x.y$ is called scalar product, there is no such thing as positive-definiteness since in general $\mathbb{F}$ is not ordered.

Notation 4.6. For a basis $B = \{ b_1, \ldots, b_n \}$ of V and $x \in V$, one can write $x = \sum_{i=1}^n x_i b_i$ where $\forall\, i \in [n]\colon x_i \in \mathbb{F}$. We denote by $\overline{x}$ the vector of coefficients $(x_1, \ldots, x_n)^T$ with respect to a given basis B.

Definition 4.7 (Matrix associated to a quadratic form). Let (V, Q) be a quadratic vector space and $B = \{ b_1, \ldots, b_n \}$ be a basis of V. The **matrix of Q with respect to** B is defined as $(a_{ij})_{i,j\in[n]}$ where for $i, j \in [n]$ we define $a_{ij} := b_i.b_j$.

Remark 4.8. The matrix associated A to Q as above is symmetric and we have:

$$Q(x) \overset{4.4(ii)}{=} x.x = \left(\sum_{i=1}^n \overline{x}_i b_i\right).\left(\sum_{j=1}^n \overline{x}_j b_j\right) = \sum_{i,j=1}^n \overline{x}_i \overline{x}_j (b_i.b_j) = \sum_{i,j=1}^n a_{ij} \overline{x}_i \overline{x}_j.$$

Hence Q is a quadratic form in the variables $\overline{x}_1, \ldots, \overline{x}_n$ in the usual sense. Furthermore we can calculate for the coefficient vectors:

$$\begin{aligned} \overline{x.y} &= \frac{1}{2}\overline{(Q(x+y) - Q(x) - Q(y))} \\ &= \frac{1}{2}\left(\left(\overline{x} + \overline{y}\right)^T A \left(\overline{x} + \overline{y}\right) - \overline{x}^T A \overline{x} - \overline{y}^T A \overline{y}\right) \\ &= \frac{1}{2}\left(\overline{x}^T A \overline{x} + \overline{x}^T A \overline{y} + \overline{y}^T A \overline{x} + \overline{y}^T A \overline{y} - \overline{x}^T A \overline{x} - \overline{y}^T A \overline{y}\right) \\ &= \frac{1}{2}\left(\overline{x}^T A \overline{y} + \overline{y}^T A \overline{x}\right) \\ &= \overline{x}^T A \overline{y}. \end{aligned}$$

Definition 4.9 (Discriminant of a quadratic form). Let (V, Q) be a quadratic vector space and let A be a matrix associated to Q. Denote the projection

$$\pi\colon \mathbb{F} \to \mathbb{F}\Big/\left[\mathbb{F}^*\right]^2$$

and define the **discriminant of** Q by $\operatorname{disc}(Q) := \pi(\det(A))$.

Remark 4.10. If one changes the basis that defined A by $X \in \mathrm{Gl}_n(\mathbb{F})$, the matrix A' with respect to this new basis is $X \cdot A \cdot X^T$, which means

$$\det(A') = \det(A) \det(X)^2 .$$

Therefore $\det(A)$ is determined up to multiplication by a square in $\mathbb{F}^*$, hence $\mathrm{disc}(Q)$ is independent of the choice of a basis.

4.2 Orthogonality

Definition 4.11 (Orthogonality). Two elements x and y of V are called **orthogonal** if $x.y = 0$. For a subset $H \subseteq V$, we define the **orthogonal complement of** H by

$$H^\perp := \{\, x \in V \mid \forall\, y \in H \colon x.y = 0 \,\} .$$

Two subspaces $U, W \subseteq V$ are called **orthogonal** if $U \subseteq W^\perp$ i.e. whenever $x \in U, y \in W$ implies $x.y = 0$. The orthogonal complement $V^\perp$ of the whole space V is called the **radical** or **kernel of** V and is denoted by $\mathrm{rad}(V)$. Its codimension i.e. $\dim(V) - \dim(\mathrm{rad}(V))$ is called **rank of** Q. If $\mathrm{rad}(V) = \{\, 0 \,\}$, we say that (V, Q) is **nondegenerate** (we may leave out V or Q if it is clear from the context and just say that V is nondegenerate or Q is nondegenerate).

Proposition 4.12.

(i). The orthogonal complement $H^\perp$ of any set $H \subseteq V$ is a subspace of V.

(ii). $H \subseteq H^{\perp\perp}$.

(iii). Q is nondegenerate if and only if $\mathrm{disc}(Q) \neq 0$.

Proof.

(i). Let $H \subseteq V$, $\lambda \in \mathbb{F}$ and $x_1, x_2 \in H^\perp$. Now calculate for any $y \in H$ that

$$\begin{aligned} & (x_1 + x_2).y = x_1.y + x_2.y = 0 \\ \Rightarrow \quad & (\lambda x_1).y = \lambda\,(x_1.y) = 0, \end{aligned}$$

so $x_1 + x_2 \in H^\perp$ and $\lambda x_1 \in H^\perp$.

(ii). Let $x \in H$. To show that $x \in H^{\perp\perp}$, we have to verify that $\forall\, y \in H^\perp \colon x.y = 0$. So let $y \in H^\perp$ be arbitrary. By definition of $H^\perp$ we have that $\forall\, z \in H \colon y.z = 0$, especially for $z = x$.

(iii). Let A be the matrix of Q with respect to a basis $\{\, b_1, \ldots, b_n \,\}$. Now check

$$\begin{aligned} \mathrm{disc}(Q) = 0 &\Leftrightarrow \det(A) = 0 \\ &\Leftrightarrow \exists\, x \in V \setminus \{\, 0 \,\} \colon A\overline{x} = 0 \\ &\overset{*}{\Leftrightarrow} \exists\, x \in V \setminus \{\, 0 \,\} \colon \forall\, y \in V \colon \overline{y}^T A \overline{x} = 0 \\ &\Leftrightarrow Q \text{ is not nondegenerate} \end{aligned}$$

The implication "$\Leftarrow$" at $*$ can be seen like this:

$$\begin{aligned} & \exists\, x \in V \setminus \{\,0\,\} \colon \forall\, y \in V \colon \overline{y}^T A \overline{x} = 0 \\ \Rightarrow \quad & \exists\, x \in V \setminus \{\,0\,\} : \forall\, j \in [n] \colon \left(\overline{b}^T\right)_j A \overline{x} = 0 \end{aligned}$$

Since $\forall\, j \in [n]$ the j-th coordinate of $A\overline{x}$, namely $\left(\overline{b}^T\right)_j A\overline{x}$ is zero, $A\overline{x}$ is zero. □

Example 4.13. Being nondegenerate is not passed on to subspaces: Let for example $Q \colon \mathbb{R}^3 \to \mathbb{R}$ be given by

$$\begin{aligned} Q(X_1, X_2, X_3) &= X_1^2 + X_3^2 + 2X_2X_3 + 2X_1X_3 \\ &= (X_1, X_2, X_3) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \end{aligned}$$

Since $\operatorname{disc}(Q) = -1$ we get that the quadratic space $(\mathbb{R}^3, Q)$ is nondegenerate. But the subspace

$$U := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \in \mathbb{R}^3 \right\}$$

with the restriction $Q|_U$ is not nondegenerate since $Q(X_1, X_2, 0) = X_1^2$ and therefore

$$\forall \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \in U \colon \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0$$

which means that $(0, 1, 0)^T$ is orthogonal to every other element of U.

Definition 4.14. Let $U \subseteq V$ be a subspace and denote the **dual space** by

$$U^* := \{\, \phi : U \to \mathbb{F} \mid \phi \text{ is linear} \,\}.$$

Furthermore define

$$\begin{array}{rccl} q_U : V & \longrightarrow & U^* \\ x & \longmapsto & (U \ni y \mapsto x.y) \end{array}$$

Fact 4.15.

(i). $\ker(q_U) = U^\perp$

(ii). Q is nondegenerate if and only if $q_V : V \to V^*$ is an isomorphism.

Proof.

(i). For $x \in U$ with $q_U(x) = 0 \in V^*$ we have

$$\forall\, y \in V \colon 0 = (q_U(x))(y) = x.y$$

which exactly defines $U^\perp$.

(ii). $\ker(q_V) \overset{(i)}{=} V^\perp$ which is by definition $\{0\}$ if and only if Q is nondegenerate. Therefore q_V is injective. But since $V \cong V^*$ it is also surjective. □

Definition 4.16. Let $U_1, \ldots, U_m \subseteq V$ be subspaces. We say that V is the **orthogonal direct sum of the** U_i if they are pairwise orthogonal and if V is the direct sum of them. We then write:

$$V = U_1 \oplus \ldots \oplus U_m.$$

Remark 4.17. Let $V = U_1 \oplus \ldots \oplus U_m$ and decompose $x \in V$ into its components $x_i \in U_i$, then

$$Q(x) = Q_1(x_1) + \ldots + Q_m(x_m) \tag{4.2.1}$$

where $Q_i := Q|_{U_i}$ are the restrictions of Q to U_i. Conversely, if (U_i, Q_i) for $i \in [0:m]$ are quadratic modules, we can define a quadratic module (V, Q) where $V = \bigoplus_{i=1}^m U_i$ by Equation (4.2.1) above and have:

$$V = U_1 \oplus \ldots \oplus U_m.$$

Example 4.18. If $U \subseteq V$ is a supplementary subspace of $\operatorname{rad}(V)$ (i.e. $V = U \oplus \operatorname{rad}(V)$) then

$$V = U \oplus \operatorname{rad}(V).$$

Proposition 4.19. *Let (V, Q) be nondegenerate. Then the following statements hold:*

(i). All metric morphisms of V into a quadratic module (V', Q') are injective.

(ii). For all subspaces $U \subseteq V$, we have:

(a) $\dim(U) + \dim\left(U^\perp\right) = \dim(V)$

(b) $U^{\perp\perp} = U$

(c) $\operatorname{rad}(U) = \operatorname{rad}\left(U^\perp\right) = U \cap U^\perp$

The quadratic module $(U, Q|_U)$ is nondegenerate if and only if the quadratic module $\left(U^\perp, Q|_{U^\perp}\right)$ is nondegenerate in which case $V = U \oplus U^\perp$.

(iii). If V is the orthogonal direct sum of two subspaces, they are nondegenerate and each of them is orthogonal to the other.

Proof.

(i). If $f\colon V \to V'$ is a metric morphism and if $f(x) = 0$, we have

$$x.y = f(x).f(y) = 0 \quad \forall\, y \in V.$$

This implies $x = 0$ because (V, Q) is nondegenerate.

(ii). Let $U \subseteq V$ be a subspace. Note that $q_U = q_V \circ \pi_{U^*}$ where $\pi_{U^*} \colon V^* \to U^*$ is the dual of the inclusion $U \hookrightarrow V$. Since q_V is bijective (by Fact 4.15(i)), q_U is surjective, thus with the canonical injection $\iota : U^\perp \to V$ the following sequence is exact:

$$\{0\} \longrightarrow U^\perp \xhookrightarrow{\iota} V \xrightarrow{q_U} U^* \longrightarrow \{0\}$$

hence

$$\dim(V) = \dim(U^*) + \dim\left(U^\perp\right) = \dim(U) + \dim\left(U^\perp\right).$$

Taking $U^\perp$ as the subspace in this argument we also get

$$\dim(V) = \dim\left(U^\perp\right) + \dim\left(U^{\perp\perp}\right)$$

which implies that

$$\dim(U) + \dim\left(U^\perp\right) = \dim(V) = \dim\left(U^\perp\right) + \dim\left(U^{\perp\perp}\right)$$

giving that $\dim(U) = \dim\left(U^{\perp\perp}\right)$. Proposition 4.12(ii) now implies $U = U^{\perp\perp}$. By the definitions:

$$\begin{aligned} \operatorname{rad}(U) &:= \{\, x \in U \mid \forall\, y \in U \colon x.y = 0 \,\} \\ U^\perp &:= \{\, x \in V \mid \forall\, y \in U \colon x.y = 0 \,\} \end{aligned}$$

we clearly get $U \cap U^\perp = \operatorname{rad}(U)$. Applying this formula to $U^\perp$, we get $U^\perp \cap U^{\perp\perp} = \operatorname{rad}\left(U^\perp\right)$ and calculate

$$\operatorname{rad}\left(U^\perp\right) = U^\perp \cap U^{\perp\perp} \overset{(ii)b}{=} U^\perp \cap U = \operatorname{rad}(U).$$

(iii). This statement is trivial, because if $V = U \oplus W$ is nondegenerate, none of U and W can be not nondegenerate and orthogonality directly follows from the defintion of the orthogonal direct sum. □

Example 4.20. Example 4.13 does not yield a counterexample to Proposition 4.19(iii): Since even though with

$$U := \left\{ \begin{pmatrix} X_1 \\ X_2 \\ 0 \end{pmatrix} \in \mathbb{R}^3 \right\}, \quad W := \left\{ \begin{pmatrix} 0 \\ 0 \\ Y_3 \end{pmatrix} \in \mathbb{R}^3 \right\}$$

it follows that $\mathbb{R}^3 = U \oplus W$ we can also calculate

$$\begin{pmatrix} X_1 \\ X_2 \\ 0 \end{pmatrix} . \begin{pmatrix} 0 \\ 0 \\ Y_3 \end{pmatrix} = \begin{pmatrix} X_1 & X_2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ Y_3 \end{pmatrix} = (X_1 + X_2) Y_3.$$

This gives us that for example $(1,0,0)^T \in U$ and $(0,0,1)^T \in W$ are not orthogonal, which means that $\mathbb{R}^3$ is not the orthogonal sum of U and W.

4.3 Isotropic vectors

Definition 4.21. An element $x \in V$ is called **isotropic** if $Q(x) = 0$. A subspace $U \subseteq V$ is called **isotropic** if all its elements are isotropic.

Example 4.22. A nondegenerate space can contain isotropic vectors: Consider the quadratic form $Q(x, y) = 2xy$ over $\mathbb{R}^2$. The associated matrix is

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

One has $Q(1, 0) = 0$ which means that $(1, 0)^T$ is isotropic but $\det(A) = -1 \neq 0$ which means that $(\mathbb{R}^2, Q)$ is nondegenerate.

Fact 4.23.

$$U \text{ isotropic} \quad \Leftrightarrow \quad U \subseteq U^\perp \quad \Leftrightarrow \quad Q|_U = 0$$

Definition 4.24. A quadratic module having a basis formed of two isotropic elements $x, y \in V$ such that $x.y \neq 0$ is called **hyperbolic plane**.

Remark 4.25. Without loss of generality, we can assume that $x.y = 1$: Just multiply y by $\frac{1}{x.y}$. Then the matrix of the quadratic form with respect to the basis $\{ x, y \}$ is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The discriminant then is $\operatorname{disc}(Q) = -1$, in particular Q is nondegenerate.

Proposition 4.26. *Let $x \in V \setminus \{ 0 \}$ be isotropic and Q be nondegenerate. Then there exists a subspace $U \subseteq V$ which contains x and is a hyperbolic plane.*

Proof. Since V is nondegenerate, there exists $z \in V$ such that $x.z = 1$. The element $y = 2z - (z.z)x$ is isotropic and $x.y = 2$. The subspace $U = \langle x, y \rangle$ has the desired property. □

Corollary 4.27. *If (V, Q) is nondegenerate and contains a nonzero isotropic element, we have $Q(V) = \mathbb{F}$.*

Proof. We have to show that $\forall\, a \in \mathbb{F}\ \exists\, v \in V$ such that $Q(v) = a$. Without loss of generality, we may assume that V is a hyperbolic plane: Let $x \in V$ be the nonzero isotropic element, then with Proposition 4.26 we can get $y \in V$ such that $\langle x, y \rangle$ is a hyperbolic plane. Furthermore we can assume that x and y are isotropic and $x.y = 1$ (see Remark 4.25). Now for $a \in \mathbb{F}$ one calculates

$$\begin{pmatrix} 1 & \frac{a}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{a}{2} \end{pmatrix} = \begin{pmatrix} 1 & \frac{a}{2} \end{pmatrix} \begin{pmatrix} \frac{a}{2} \\ 1 \end{pmatrix} = a$$

and therefore get we get $a = Q\left(x + \frac{a}{2} y\right)$. □

4.4 Orthogonal bases

Definition 4.28. A basis $\{ b_1, \ldots, b_n \}$ is called **orthogonal** if its elements are pairwise orthogonal i.e.

$$V = \langle b_1 \rangle \oplus \ldots \oplus \langle b_n \rangle .$$

Remark 4.29. If V is nondegenerate, this is equivalent to saying that the matrix associated to Q with respect to the basis $B = \{ b_1, \ldots, b_n \}$ is a diagonal matrix with diagonal entries $a_1, \ldots, a_n \in \mathbb{F}^*$:

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

If $\overline{x} = (x_1, \ldots, x_n)^T$ is the coordinate vector of $x \in V$ with respect to the basis B, we have that $Q(x) = a_1 x_1^2 + \ldots + a_n x_n^2$.

Fact 4.30. *Let (V, Q) be a nondegenreate quadratic space with an orthogonal basis $\{ b_1, \ldots, b_n \}$, then $\forall\, i \in [1 : n]\colon (b_i . b_i) \neq 0$.*

Proof. Write

$$V = \langle b_1 \rangle \oplus \cdots \oplus \langle b_n \rangle .$$

Now Proposition 4.19(iii) gives that each $\langle b_i \rangle$ is nondegenerate which makes it impossible for b_i to be isotropic. □

Theorem 4.31. *Every quadratic module has an orthogonal basis.*

Proof. We prove this by induction on the dimension $n := \dim(V)$. The case $n = 0$ is trivial. Now let n be arbitrary. If V is isotropic, all bases of V are orthogonal. Otherwise, choose an element $b \in V$ such that $b.b \neq 0$. Now the orthogonal complement $U := \{ b \}^\perp$ is a hyperplane (i.e. has dimension $n - 1$) and since $b \notin U$, one has $V = \langle b \rangle \oplus U$. By induction hypothesis U has an orthogonal basis B and $\{ b \} \cup B$ is an orthogonal basis. □

Definition 4.32. Two orthogonal bases

$$B = \{ b_1, \ldots, b_n \} \quad \text{and} \quad C = \{ c_1, \ldots, c_n \}$$

of V are called **contiguous** if they have an element in common (i.e. if there exist i and j with $b_i = c_j$). A sequence of bases $(B_0, B_1, \ldots, B_m)$ is called a **chain contigouously relating B and C** if

- $B_i \subseteq V$ is an orthogonal basis for $1 \leq i \leq m$,
- $B_0 = B$ and $B_m = C$,
- B_i and B_{i+1} are contiguous for $0 \leq i < m$.

Lemma 4.33. *Let (V, Q) be a nondegenerate quadratic module, $a, b \in V \setminus \{0\}$ and define P as the vector space spanned by a and b. Then*

$$(a.a)(b.b) \neq (a.b)^2 \Leftrightarrow (\dim(P) = 2 \text{ and } P \text{ is nondegenerate}).$$

Proof. We will prove the equivalent statement

$$(a.a)(b.b) = (a.b)^2 \Leftrightarrow (\dim(P) < 2 \text{ or } P \text{ is degenerate}).$$

"$\Leftarrow$": Asume that $\dim(P) < 2$, then there exists $\lambda \in \mathbb{F}$ with $a = \lambda b$ implying:

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

If P is degenerate. Then there exists $v = \lambda a + \mu b \in P \setminus \{0\}$ with the property that $\forall\, w \in P$: $(v.w) = 0$. Now calculate

$$0 = (v.a) = \lambda(a.a) + \mu(b.a) \quad \Leftrightarrow \quad -\mu(b.a) = \lambda(a.a) \tag{4.4.1}$$
$$0 = (v.b) = \lambda(a.b) + \mu(b.b) \quad \Leftrightarrow \quad -\lambda(a.b) = \mu(b.b). \tag{4.4.2}$$

Since $v \neq 0$ at least one of μ and λ is not zero. If $\lambda \neq 0$ we get by (4.4.2) that $(a.b) = -\frac{\mu}{\lambda}(b.b)$ which, substituting in (4.4.1), yields

$$\mu\frac{\mu}{\lambda}(b.b) = \lambda(a.a) \quad \Leftrightarrow \quad \frac{\mu^2}{\lambda^2}(b.b) = (a.a).$$

Putting this all together we get

$$(a.a)(b.b) - (a.b)^2 = \frac{\mu^2}{\lambda^2}(b.b)(b.b) - \left(-\frac{\mu}{\lambda}(b.b)\right)^2 = 0.$$

The same works for $\mu \neq 0$ because (4.4.1) and (4.4.2) are symmetric in a and b.

"$\Rightarrow$": Define the element

$$c := (b.b)a - (a.b)b \in P$$

and observe that

$$c.a = (b.b)(a.a) - (a.b)(b.a) = 0 \text{ by assumption}$$
$$c.b = (b.b)(a.b) - (a.b)(b.b) = 0$$

So $c \in \operatorname{rad}(P)$, which leads to either c is zero or P is degenerate. In the latter case, we are done, else we get that $(b.b)a - (a.b) = 0$, which is a linear combination of 0 in a and b (which generate P). Hence either $\dim(P) < 2$ (in which case we are done again) or $(b.b) = 0$ and $(a.b) = 0$, which implies that $b \in \operatorname{rad}(P)$. Now if $b = 0$, we analogously get that $\dim(P) < 2$ and if $b \neq 0$, P is degenerate. □

Lemma 4.34 (Gram-Schmidt). *Let (V, Q) be nondegenerate, $a, b \in V$ linearly independent and a nonisotropic. Then there exists $c \in V$ such that*

$$\langle a, b\rangle = \langle a\rangle \oplus \langle c\rangle$$

Proof. Set $p := \frac{a.b}{a.a}a$, $c := b - p$ and calculate:

$$a.c = a.(b-p) = a.b - a.p = a.b - a.a\frac{a.b}{a.a} = 0.$$

Let $\lambda, \mu \in \mathbb{F}$ and calculate

$$0 = \lambda a + \mu c = \lambda a + \mu b - \mu p = \lambda a + \mu b - \mu\frac{a.b}{a.a}a = \left(\lambda - \mu\frac{a.b}{a.a}\right)a + \mu b.$$

Since a and b are linearly independent, we get $\mu = 0$ and $\lambda - \mu\frac{a.b}{a.a} = 0$, i.e. $\lambda = 0$, meaning that a and c are linearly independent, too. □

Theorem 4.35. *Let (V, Q) be a nondegenerate quadratic module of dimension $\dim(V) \geq 3$ with two orthogonal bases B and C. Then there exists a chain contigously relating B and C.*

Proof. Let $B = \{ b_1, \ldots, b_n \}$ and $C = \{ c_1, \ldots, c_n \}$ and define

$$\forall\, i \in \{ 1, 2 \} : \mu_i := (b_1.b_1)(c_i.c_i) - (b_1.c_i)^2.$$

We now distinguish the cases where $\mu_1 = 0$, $\mu_2 = 0$ and $\mu_i \neq 0$ for $i \in \{ 1, 2 \}$:

Case 1 ($\mu_1 \neq 0$). By assumption and Lemma 4.33 (applied to b_1 and c_1) $P := \langle b_1, c_1 \rangle$ has dimension 2 and is nondegenerate. Since B and C are both orthogonal bases, by Fact 4.30 we know that b_1 and c_1 are both nonisotropic and Lemma 4.34 therefore yields $x, y \in V$ with

$$P = \langle b_1 \rangle \oplus\!\!\ominus \langle x \rangle \quad \text{and} \quad P = \langle c_1 \rangle \oplus\!\!\ominus \langle y \rangle$$

Additionally by Proposition 4.19 we get, that $P^\perp$ is nondegenerate, too, and ultimately $V = P \oplus\!\!\ominus P^\perp$. Now let $\{ d_3, \ldots d_n \}$ be an orthogonal basis of $P^\perp$ (which exists because of Theorem 4.31). Then the sequence

$$(\quad B \quad , \quad \{ b_1, x, d_3, \ldots, d_n \} \quad , \quad \{ c_1, y, d_3, \ldots, d_n \} \quad , \quad C \quad)$$

contigouously relates B and C.

Case 2 ($\mu_2 \neq 0$). This is analog to the case $\mu_1 \neq 0$ by switching c_1 and c_2.

Case 3 ($\mu_1 = \mu_2 = 0$). The condition on μ_i is equivalent to

$$\forall\, i \in \{ 1, 2 \} : (b_1.b_1)(c_i.c_i) = (b_1.c_i)^2.$$

With this, we prove the following claim:

Claim 4.35.1. *$\exists\, \lambda \in \mathbb{F}$: $e_\lambda := c_1 + \lambda c_2$ is nonisotropic and $\langle e_\lambda, b_1 \rangle$ is a nondegenerate plane.*

Proof. For e_λ being nonisotropic, we need to ensure that $0 \neq e_\lambda.e_\lambda$. So calculate

$$(e_\lambda.e_\lambda) = (c_1.c_1) + 2\lambda(c_1.c_2) + \lambda^2(c_2.c_2) = (c_1.c_1) + \lambda^2(c_2.c_2)$$

since C is orthogonal. We know that $(c_i, c_i) \neq 0$ (Fact 4.30) and therefore have that e_λ is nonisotropic if and only if $\lambda^2 \neq -\frac{c_1.c_1}{c_2.c_2}$.

Applying Lemma 4.33 to e_λ and b_1 yields that it is necessary and sufficient for them to generate a nondegenerate plane that

$$(b_1.b_1)(e_\lambda.e_\lambda) - (b_1.e_\lambda)^2 \neq 0$$

So calculate

$$\begin{aligned}(b_1.b_1)(e_\lambda.e_\lambda) &= (b_1.b_1)((c_1.c_1) + \lambda^2(c_2.c_2)) \\ &= (b_1.b_1)(c_1.c_1) + \lambda^2(b_1.b_1)(c_2.c_2) \\ &= (b_1.c_1)^2 + \lambda^2(b_1.c_2)^2 \quad \text{since } \mu_i = 0\end{aligned}$$

and

$$\begin{aligned}(b_1.e_\lambda)^2 &= (b_1.c_1 + \lambda(b_1.c_2))^2 \\ &= (b_1.c_1)^2 + 2\lambda(b_1.c_1)(b_1.c_2) + \lambda^2(b_1.c_2)^2\end{aligned}$$

leading to

$$0 \neq (b_1.b_1)(e_\lambda.e_\lambda) - (b_1.e_\lambda)^2 = -2\lambda(b_1.c_1)(b_1.c_2)$$

By Fact 4.30 and $\mu_i = 0$, we get that $(b_1.c_i)^2 \neq 0$ and therefore that $\lambda \neq 0$. Summarized, e_λ verifies the conditions of Claim 4.35.1 if and only if $\lambda^2 \neq -\frac{c_1.c_1}{c_2.c_2}$ and $\lambda \neq 0$. This rules out only 3 values for $\lambda \in \mathbb{F}$, so if $\mathbb{F}$ has at least 4 elements, we are done. We are left to show the statement for the case $\mathbb{F} = \mathbb{F}_3$ ($\mathbb{F} = \mathbb{F}_2$ is excluded, since $\operatorname{char}(\mathbb{F}) \neq 2$): In $\mathbb{F}_3$, all nonzero squares are 1, so $\mu_i = 0$ is equivalent to $(b_1.b_1)(c_i.c_i) = 1$. Now calculate

$$\lambda^2 \neq -\frac{c_1.c_1}{c_2.c_2} = -\frac{(b_1.b_1)(c_1.c_1)}{(b_1.b_1)(c_2.c_2)} = -1$$

and see that $\lambda = 1$ realizes the conditions $\lambda^2 \neq -1$ and $\lambda \neq 0$ which finishes the proof of Claim 4.35.1. □

Let $\lambda \in \mathbb{F}$ be as in Claim 4.35.1. Since $e_\lambda = c_1 + \lambda c_2$ is not isotropic there is $y \in V$ such that $\{ e_\lambda, y \}$ is an orthogonal basis of $\langle c_1, c_2 \rangle$. The set

$$D := \{ e_\lambda, y, c_3, \ldots, c_n \}$$

then is an orthogonal basis of V because

$$\forall\, i \in [3:n]\colon e_\lambda.c_i = (c_1 + \lambda c_2).c_i = c_1.c_i + \lambda c_2.c_i = 0.$$

So C and D are contigously related. But since $\langle b_1, e_\lambda \rangle$ is a nondegenerate plane, Case 1 of this proof contigously relates B and D, ultimately yielding that B and C are related contigously. □

4.5 Witt's Theorem

In this section let (V, Q) and (V', Q') be two nondegenerate quadratic spaces, $U \subseteq V$ a subspace of V and $s : U \to V'$ be an injective metric morphism. The goal is to extend s to a subspace larger than U, and, if possible, to all of V.

Proposition 4.36. *If U is degenerate, there exists $U_1 \subseteq V$ containing U with*

$$\dim(U_1) = \dim(U) + 1$$

extending s to an injective metric morphism $s_1 \colon U_1 \hookrightarrow V'$ with $s_1|_U = s$.

Proof. Let $x \in \operatorname{rad}(U) \setminus \{0\}$ and $g : U \to \mathbb{F}$ be linear such that $g(x) = 1$. Since V is nondegenerate, Fact 4.15(ii) implies that q_V is an isomorphism and therefore surjective i.e. that there exists $y \in V$ such that $q_V(y)|_U = g$ or in other words for all $u \in U \colon g(u) = u.y$. Since $x \in \operatorname{rad}(U)$ and $y.x = 1 \neq 0$ we get that $y \notin U$ and therefore that $U_1 := U \oplus \langle y \rangle$ contains U as a hyperplane.

Replacing y by $y - \lambda x$ with $\lambda = (y.y)/2$ does not change g since for any $u \in U$:

$$u.(y - \lambda x) = u.y - \lambda \underbrace{u.x}_{\substack{=0 \text{ since} \\ x \in \operatorname{rad}(U)}} = u.y.$$

But the replacement yields that $y.y = 0$ since

$$(y - \lambda x).(y - \lambda x) = (y.y) - 2\lambda \underbrace{(y.x)}_{\substack{=1 \text{ since} \\ g(x)=1}} + \lambda^2 \underbrace{(x.x)}_{\substack{=0 \text{ since} \\ x \in \operatorname{rad}(U)}} = (y.y) - 2\frac{(y.y)}{2} = 0.$$

The same construction works for $U' := s(U)$, $x' = s(x)$ and $g' = g \circ s^{-1}$ yielding $y' \in V'$ and $U_1' = U' \oplus \langle y' \rangle$. Define $s_1 \colon U_1 \to U_1'$ by

$$\begin{aligned} s_1 \colon U \oplus \langle y \rangle &\to U' \oplus \langle y' \rangle \\ (u, \alpha y) &\mapsto (s(u), \alpha y'). \end{aligned}$$

We claim that s_1 is a metric isomorphism. s_1 is indeed well-defined, linear, injective and surjective by definition and we check that it is metric. Let $u \in U$ and $\alpha y \in \langle y \rangle$, then

$$Q'(s_1(u, \alpha y)) = Q'(s(u), \alpha y') \overset{4.17}{=} Q'|_U(s(u)) + \underbrace{Q'|_{\langle y \rangle}(\alpha y')}_{=0 \text{ since } y'.y'=0} = Q'(s(u)).$$

And since s is metric, $Q' \circ s = Q$, implying that $Q' \circ s_1 = Q$ and finally that s_1 is metric. □

Theorem 4.37 (Witt). *If (V, Q) and (V', Q') are isomorphic and nondegenerate, every injective metric morphism*

$$s : U \hookrightarrow V'$$

from a subspace $U \subseteq V$ can be extended to a metric isomorphism of V onto V'.

Proof. Since V and V' are isomorphic, we can without loss of generality assume that $V = V'$. If V is degenerate, we can apply Proposition 4.36 to be finished or to be left with a nondegenerate subspace $U \subseteq V$. We argue by induction on $\dim(U)$.

If $\dim(U) = 1$, U is generated by a nonisotropic element $x \in U$. For $y := s(x)$, we have $y.y = s(x).s(x) = x.x$. One can choose $\epsilon = \pm 1$ such that $x + \epsilon y$ is nonisotropic too since otherwise we would have:

$$\begin{aligned} 0 &= (x+y).(x+y) = x.x + 2x.y + y.y = 2x.x + 2x.y \\ 0 &= (x-y).(x-y) = x.x - 2x.y + y.y = 2x.x - 2x.y \\ \Rightarrow 0 &= 4x.x \Leftrightarrow 0 = x.x \end{aligned}$$

Define $z = x + \epsilon y$ and let $H = \langle z \rangle^{\perp}$. We have $V = \langle z \rangle \oplus H$ by Proposition 4.19(iii) since (U, Q) is nondegenerate. Let $\sigma \colon V \to V$ be the unique automorphism defined by $\sigma|_H = \mathrm{id}_H$ and $\sigma(z) := -z$. We have

$$\begin{array}{lcll} \sigma(x - \epsilon y) & = & x - \epsilon y & \text{since } x - \epsilon y \in H \\ \sigma(x + \epsilon y) & = & -x - \epsilon y & \text{by definition} \end{array}$$

yielding

$$\sigma(2x) = \sigma(x - \epsilon y) + \sigma(x + \epsilon y) = x - \epsilon y - x - \epsilon y = -2\epsilon y$$

and ultimately $\sigma(x) = -\epsilon y$, hence the automorphism $-\epsilon\sigma$ extends s.

If $\dim(U) > 1$, we decompose U as $U_1 \oplus U_2$ with $U_1, U_2 \neq \{0\}$. By induction hypothesis, the restriction s_1 of s to U_1 extends to an automorphism σ_1 of V. After replacing s by $\sigma_1^{-1} \circ s$ one can thus suppose that s is the identity on U_1. Then the morphism s carries U_2 into the orthogonal complement V_1 of U_1. Again by induction hypothesis, the restriction of s to U_2 extends to an automorphism σ_2 of V_1. Define σ by $\sigma|_{U_1} = \mathrm{id}_{U_1}$ and $\sigma|_{V_1} = \sigma_2$ has the desired property. □

Corollary 4.38. *Two isomorphic subspaces of a nondegenerate quadratic module have isomorphic orthogonal complements.*

Proof. Let $U, W \subseteq V$ be two isomorphic subspaces. By Theorem 4.37 we can extend the isomorphism between them to an automorphism of the whole space and restrict it to the orthogonal complement $U^{\perp}$, yielding an isomorphism between $U^{\perp}$ and $W^{\perp}$. □

4.6 Application to quadratic form equivalence

Definition 4.39. Let $f \in \mathbb{F}[X_1, \ldots, X_n]$ be a quadratic form with

$$f(X_1, \ldots, X_n) = \sum_{i=1}^{n} a_{ii} X_i^2 + 2 \sum_{i<j} a_{ij} X_i X_j \quad \forall\, i \leq j \in [n] \colon a_{ij} \in \mathbb{F}$$

then $(\mathbb{F}^n, f)$ is the **quadratic module associated to** f.

Proposition 4.40. *Quadratic forms in the same number of variables are equivalent if and only if the associated quadratic modules are isomorphic.*

Proof. An isomorphism of quadratic modules $\phi : (\mathbb{F}^n, f) \xrightarrow{\cong} (\mathbb{F}^n, g)$ can be represented by a matrix $P \in \mathrm{Gl}_n(\mathbb{F})$ such that $f \circ P = g$ which is exactly the definition of form equivalence. □

Remark 4.41. Let $f, g \in \mathbb{F}[X_1, \ldots, X_n]$ be two quadratic forms with corresponding matrices A and B. Stating $f \sim g$ amounts to saying that there exists $X \in \mathrm{Gl}_n(\mathbb{F})$ with $B = XAX^T$ (by Remark 4.10).

Definition 4.42. Let $f \in \mathbb{F}[X_1, \ldots, X_n]$ and $g \in \mathbb{F}[x_1, \ldots, x_m]$ be two quadratic forms, then we define the **orthogonal sum** $f \oplus g \in \mathbb{F}[X_1, \ldots, X_{n+m}]$ by

$$(f \oplus g)(x_1, \ldots, x_{n+m}) := f(x_1, \ldots, x_n) + g(x_{n+1}, \ldots, x_{n+m}).$$

Correspondingly we write $f \hat{\ominus} g := f \oplus (-g)$.

Fact 4.43. *The orthogonal sum of forms corresponds to the orthogonal sum of quadratic spaces, i.e.* $\forall f \in \mathbb{F}[X_1, \ldots, X_n], g \in \mathbb{F}[x_1, \ldots, x_m]$*:*

$$\left(\mathbb{F}^{n+m}, f \oplus g\right) \cong (\mathbb{F}^n, f) \oplus (\mathbb{F}^m, g).$$

Proof. Define the map $\varphi : \mathbb{F}^n \oplus \mathbb{F}^m \to \mathbb{F}^{n+m}$ component-wise by

$$\mathbb{F}^n \ni (x_1, \ldots, x_n) \;\mapsto\; (x_1, \ldots, x_n, 0, \ldots, 0) \in \mathbb{F}^{n+m}$$

and

$$\mathbb{F}^m \ni (x_1, \ldots, x_m) \;\mapsto\; (0, \ldots, 0, x_1, \ldots, x_m) \in \mathbb{F}^{n+m}.$$

This map induces an isomorphism of vector spaces, which is also a metric morphism by definition of $f \oplus g$. □

Fact/Definition 4.44. A form $f \in \mathbb{F}[X, Y]$ is called **hyperbolic** if and only if:

$$f \sim xy \sim X^2 - Y^2.$$

This means that the quadratic space $(\mathbb{F}^2, f)$ is a hyperbolic plane.

Proof. First note that XY and $X^2 - Y^2$ are equivalent via

$$\tau\colon \begin{cases} X & \mapsto \frac{X+Y}{2} \\ Y & \mapsto \frac{X-Y}{2} \end{cases},$$

since

$$\begin{aligned} \tau(X)^2 - \tau(Y)^2 &= \left(\frac{X+Y}{2}\right)^2 - \left(\frac{X+Y}{2}\right)^2 \\ &= \frac{X^2 + 2XY + Y^2}{4} - \frac{X^2 - 2XY + Y^2}{4} = XY. \end{aligned}$$

Let now $f \in \mathbb{F}[X, Y]$ be hyperbolic. We need to find a basis $\{v, w\}$ of isotropic vectors such that $v.w \neq 0$. Since $f \sim XY$, there exists $A \in \mathrm{Gl}_n(\mathbb{F})$ such that

$$f\left(A\begin{pmatrix} X \\ Y \end{pmatrix}\right) = XY. \tag{4.6.1}$$

Now choose

$$v = A\begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad w = A\begin{pmatrix}0\\1\end{pmatrix},$$

which is, since A is bijective, a basis of $\mathbb{F}^2$. Now calculate with the help of (4.6.1):

$$\begin{aligned} f(v) &= f\left(A\begin{pmatrix}1\\0\end{pmatrix}\right) = 0 \\ f(w) &= f\left(A\begin{pmatrix}0\\1\end{pmatrix}\right) = 0 \\ f(v+w) &= f\left(A\begin{pmatrix}1\\1\end{pmatrix}\right) = 1 \end{aligned}$$

yielding that

$$\begin{aligned} v.v &= \frac{1}{2}\left(f(v+v) - f(v) - f(v)\right) = \frac{1}{2}\left(4f(v) - f(v) - f(v)\right) = 0 \\ w.w &= \frac{1}{2}\left(f(w+w) - f(w) - f(w)\right) = \frac{1}{2}\left(4f(w) - f(w) - f(w)\right) = 0 \\ v.w &= \frac{1}{2}\left(f(v+w) - f(v) - f(w)\right) = \frac{1}{2} \neq 0. \end{aligned}$$

This means that $\left(\mathbb{F}^2, f\right)$ is a hyperbolic plane. □

Definition 4.45. A form $f \in \mathbb{F}[X_1, \ldots, X_n]$ **represents** an element $a \in \mathbb{F}$ if there exists $x \in \mathbb{F}^n \setminus \{\mathbf{0}\}$ with $f(x) = a$.

Remark 4.46. A quadratic form represents zero if and only if the corresponding quadratic space contains a nonzero isotropic element.

Proposition 4.47. *If $f \in \mathbb{F}[X_1, \ldots, X_n]^{=2}$ represents zero and is nondegenerate, one has $f \sim h \oplus g$ where h is hyperbolic and g is a form of rank $n-2$. Moreover, f represents all elements of $\mathbb{F}$.*

Proof. Basically this is the translation of Proposition 4.26 and Corollary 4.27 into the language of quadratic form equivalence: Since f represents zero, there exists a nonzero isotropic element x. Then by Proposition 4.26 there exists a hyperbolic plane $U \subseteq V$ containing x, which by Remark 4.25 is nondegenreate. This implies with Proposition 4.19(iii) that $\mathbb{F} = U \oplus U^{\perp}$ where U is a hyperbolic plane. By Fact/Definition 4.44, we get that there is an h that is hyperbolic. Finally by Corollary 4.27 we have that $f(\mathbb{F}^n) = \mathbb{F}$ which means that every element of $\mathbb{F}$ is represented. □

Corollary 4.48. *Let $g \in \mathbb{F}[X_1, \ldots, X_{n-1}]^{=2}$ be nondegenerate and $a \in \mathbb{F}^*$. Then the following properties are equivalent:*

(i). g represents a.

(ii). $\exists\, h \in \mathbb{F}[X_1, \ldots, X_{n-2}]^{=2} : g \sim h \oplus aX_n^2$.

(iii). $g \mathbin{\hat{\ominus}} aX_n^2$ represents zero.

Proof.

Step 1 ((ii)⇒(i) and (ii)⇒(iii)). Let h be as in the statement. Set

$$t := (t_1, \ldots, t_{n-2}, t_n) \text{ with } \forall\, i \in [1 : n-2] \colon t_i = 0 \text{ and } t_n = 1$$

and calculate

$$\left(h \oplus a x_n^2\right)(t) = a.$$

Let τ be a form equivalence between g and $h \oplus a x_n^2$. Then we have

$$g\left(\tau\left(t\right)\right) = \left(h \oplus a x_n^2\right)(t) = a.$$

As τ is invertible and linear, from $t \neq 0$ it follows that $\tau\left(t\right) \neq 0$. Hence g represents a. This immediately gives

$$\left(g \mathbin{\hat{\ominus}} a x_n^2\right)\left(\tau\left(t\right), 1\right) = a - a = 0.$$

Step 2 ((i)⇒(ii)). Since g represents a, the quadratic space corresponding to g contains an element x such that $x.x = a$. By Proposition 4.19(ii) we then can write $\mathbb{F}^{n-1} = \{\, x \,\}^{\perp} \oplus \langle x \rangle$. Now let $\{\, b_1, \ldots, b_{n-2} \,\}$ be a basis of $\{\, x \,\}^{\perp}$ and define the quadratic form h by

$$h(X_1, \ldots, X_{n-2}) := \sum_{i=1}^{n-2} b_i.b_i X_i + 2 \sum_{i<j} b_i.b_j X_i X_j.$$

Fact 4.43 then implies the statement.

Step 3 ((iii)⇒(i)). If the form $f := g \mathbin{\hat{\ominus}} a x_n^2$ represents zero, there exists a nontrivial zero $(x_1, \ldots x_n)$ of f. Then we distinguish the following two cases:

Case 1 ($x_n = 0$). This implies that $(x_1, \ldots, x_{n-1})$ is a nontrivial zero of g and by Proposition 4.47 it follows that g represents a.

Case 2 ($x_n \neq 0$). In this case, we have

$$0 = \frac{f\left(x_1, \ldots, x_n\right)}{x_n^2} = f\left(\frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n}, 1\right) = g\left(\frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n}\right) - a. \square$$

Theorem 4.49 (Decomposition into Sums of squares). *It holds that*

$$\forall\, f \in \mathbb{F}\left[X_1, \ldots, X_n\right]^{=2} : \exists\, a_1, \ldots, a_n \in \mathbb{F} \colon f \sim \sum_{i=1}^{n} a_i X_i^2.$$

Proof. Theorem 4.31 together with Fact 4.43 gives the statement. □

Fact 4.50. *Let $f \in \mathbb{F}\left[X_1, \ldots, X_n\right]^{=2}$. By Theorem 4.49 there exist $a_1, \ldots, a_n \in \mathbb{F}$ such that $f \sim \sum_{i=1}^{n} a_i X_i^2$. The* $\operatorname{rank}\left(f\right)$ *defined in Definition 4.11 coincides with the number*

$$\left|\{\, i \in [n] \mid a_i \neq 0 \,\}\right|.$$

Two isomorphic quadratic modules $(\mathbb{F}^n, f) \cong (\mathbb{F}^n, g)$ are of the same rank.

Fact 4.51. *For any $f \in \mathbb{F}[X_1, \ldots, X_n]^{=2}$, we can write*

$$f(X_1, \ldots, X_n) = (X_1, \ldots, X_n) A (X_1, \ldots, X_n)^T$$

for a symmetric $A \in \mathbb{F}^{n \times n}$ with entries $a_{ij} \in \mathbb{F}$. Since A is a symmetric matrix, we can efficiently obtain $C \in \mathrm{Gl}_n(\mathbb{F})$ by the Gram-Schmidt process (without normalization) such that CAC^T is diagonal. Name the diagonal elements $b_i \in \mathbb{F}$. Then we have

$$f((X_1, \ldots, X_n)C) = (X_1, \ldots, X_n) CAC^T (X_1, \ldots, X_n)^T = \sum_{i=1}^{n} b_i X_i^2.$$

This process makes Theorem 4.49 explicit and takes $\mathcal{O}\left(b \cdot n^3\right)$ where

$$b := \log\left(\max_{1 \le i,j \le n} \{\, a_{ij} \,\}\right)$$

is the maximal bitsize of the entries of A.

Corollary 4.52. *Let g and h be two nondegenerate forms of rank ≥ 1 and $f := g \mathbin{\hat{\ominus}} h$. The following properties are equivalent:*

(i). f represents zero.

(ii). $\exists\, a \in \mathbb{F}^*$ which is represented by g and by h.

(iii). $\exists\, a \in \mathbb{F}^*$ such that $g \mathbin{\hat{\ominus}} aZ^2$ and $h \mathbin{\hat{\ominus}} aZ^2$ represent zero.

Proof.

Step 1 ((ii)⇔(iii)). This follows from Corollary 4.48.

Step 2 ((ii)⇒(i)). Since f is defined as the difference of g and h it represents zero.

Step 3 ((i)⇒(ii)). A nontrivial zero of f can be written as (x, y) with $f(x) = g(x)$ (by definition of $\hat{\ominus}$). If $a = g(x) = h(y)$ is $\neq 0$, we are done. So let $a = 0$ which means that at least one of the forms g and h represents zero, say g. By Proposition 4.47 g represents all elements of $\mathbb{F}$, and in particular all nonzero values taken by h. □

Theorem 4.53 (Witt's Cancellation Theorem). *Let $f = g \oplus h$ and $f' = g' \oplus h'$ be two nondegenerate quadratic forms. If $f \sim f'$ and $h \sim h'$, one has $g \sim g'$.*

Proof. Corollary 4.38 gives the statement. □

Corollary 4.54. *If f is nondegenerate, then there exist hyperbolic $g_1, \ldots, g_m$ and an h that does not represent zero with:*

$$f \sim g_1 \oplus \ldots \oplus g_m \oplus h$$

and this decomposition is unique up to equivalence.

Proof. Existence follows from Proposition 4.47 and uniqueness from Theorem 4.53.□

4.7 Quadratic forms over v-adic numbers

Let $v \in V$ and $p \in P$ in this section. Remember that for $v = \infty$, the v-adic numbers are $\mathbb{R}$. The goal of this chapter is to define invariants of quadratic forms over the v-adic numbers that fully classify their equivalence. Later, we will show that these invariants are computable in an efficient way under certain assumptions.

Definition 4.55 (The Hasse-Minkowski invariant). With the help of the Hilbert symbol[1], we now define the following map for every $v \in V$:

$$\begin{array}{rcl} \varepsilon\colon (\mathbb{Q}_v^*)^n & \longrightarrow & \{\pm 1\} \\ (a_1, \ldots, a_n) & \longmapsto & \prod_{i<j} (a_i, a_j)_v \,. \end{array}$$

If (V, Q) is a nondegenreate quadratic module of rank n with orthogonal basis $B = \{ b_1, \ldots, b_n \}$ we define furthermore

$$\varepsilon(B) := \varepsilon(b_1.b_1, \ldots, b_n.b_n)\,.$$

Theorem 4.56. *ε does not depend on the choice of the orthogonal basis.*

Proof. Let $v \in V$ and $B = \{ b_1, \ldots, b_n \}$ be an orthogonal basis of a nondegenerate module (V, Q) of rank n. We will use induction on n:

Step 1 ($n = 1$). We have $\varepsilon(B) = 1$.

Step 2 ($n = 2$). We have that $\varepsilon(B) = (b_1.b_1, b_2.b_2)_v$. By definition we have $(b_1.b_1, b_2.b_2)_v = 1$ if and only if $z^2 = b_1.b_1x^2 + b_2.b_2y^2$ has a nontrivial solution or, in other words, the form

$$Z^2 - (b_1.b_1)\,X^2 - (b_2.b_2)\,Y^2 \tag{4.7.1}$$

represents zero. Note that Z^2 and $(b_1.b_1)\,X^2 + (b_2.b_2)\,Y^2$ are nondegenerate forms. So by Corollary 4.52 we have that (4.7.1) represents zero if and only if there exists $a \in \mathbb{F}^*$ and $z \in \mathbb{F}^*, (x, y) \in \mathbb{F}^2 \setminus \{ \mathbf{0} \}$:

$$z^2 = a = (b_1.b_1)\,x^2 + (b_2.b_2)\,y^2.$$

But $(b_1.b_1)\,X^2 + (b_2.b_2)\,Y^2$ represents $a \in \mathbb{F}^*$ if and only if $\exists\, v \in V\colon Q(v) = a$, which does not depend on the choice of the basis B.

Step 3 ($n - 1 \to n$). Let $n \geq 3$ and the statement be proven for $n - 1$. By Theorem 4.35 every pair of orthogonal bases is contigously related. So if we prove the statement for two contiguous bases, it follows for any two bases (since there exists a chain contigously relating them). So let $B = \{ b_1, \ldots, b_n \}$ and $B' = \{ b'_1, \ldots, b'_n \}$ be two contiguous orthogonal bases of V. Since the Hilbert symbol is symmetric (see Proposition 3.9(ii)), $\varepsilon(B)$ does not change if we permute the elements of B. So without loss of generality assume that $b_1 = b'_1$. For shortage of notation, define

$$\forall\, i \in [n]\colon a_i := b_i.b_i,\ a'_i := b'_i.b'_i.$$

So the assumption above translates into

$$a_1 = a'_1 \tag{4.7.2}$$

[1] See Definition 3.2

First note that

$$(a_1, a_2 \cdot \dots \cdot a_n) \overset{3.4}{=} \left(a_1, a_1^2 \cdot a_2 \cdot \dots \cdot a_n\right)_v = \left(a_1, a_1 \mathrm{disc}\,(Q)\right)_v \tag{4.7.3}$$

and calculate

$$\begin{aligned}
\varepsilon(B) &= \prod_{i<j} (a_i, a_j)_v = \prod_{i=1}^{n} \prod_{\substack{j=1 \\ j\neq i}}^{n} (a_i, a_j)_v \\
&= \left(\prod_{j=2}^{n} (a_1, a_j)_v\right) \cdot \left(\prod_{i=2}^{n} \prod_{\substack{j=2 \\ j\neq i}}^{n} (a_i, a_j)_v\right) \\
&\overset{3.13}{=} (a_1, a_2 \cdot \dots \cdot a_n) \cdot \left(\prod_{i=2}^{n} \prod_{\substack{j=2 \\ j\neq i}}^{n} (a_i, a_j)_v\right) \\
&\overset{(4.7.3)}{=} \left(a_1, a_1 \mathrm{disc}\,(Q)\right) \cdot \left(\prod_{i=2}^{n} \prod_{\substack{j=2 \\ j\neq i}}^{n} (a_i, a_j)_v\right)
\end{aligned}$$

Similarly we have that

$$\begin{aligned}
\varepsilon(B') &= \left(a_1', a_1' \mathrm{disc}\,(Q)\right) \cdot \left(\prod_{i=2}^{n} \prod_{\substack{j=2 \\ j\neq i}}^{n} \left(a_i', a_j'\right)_v\right) \\
&\overset{(4.7.2)}{=} \left(a_1, a_1 \mathrm{disc}\,(Q)\right) \cdot \left(\prod_{i=2}^{n} \prod_{\substack{j=2 \\ j\neq i}}^{n} \left(a_i', a_j'\right)_v\right).
\end{aligned}$$

The induction hypothesis, applied to the orthogonal complement of b_1, yields

$$\left(\prod_{i=2}^{n} \prod_{\substack{j=2 \\ j\neq i}}^{n} (a_i, a_j)_v\right) = \left(\prod_{i=2}^{n} \prod_{\substack{j=2 \\ j\neq i}}^{n} \left(a_i', a_j'\right)_v\right)$$

from which the desired result follows. □

Notation 4.57. From now on, we write $\varepsilon(Q)$ instead of $\varepsilon(B)$.

Corollary 4.58. *For $v \in V$ and $f \in \mathbb{Q}_v[X_1, \dots, X_n]^{=2}$ with*

$$f \sim a_1 X^1 + \dots + a_n X^n$$

we have that the two elements

$$\begin{aligned}
\mathrm{disc}\,(f) &= a_1 \cdot \dots \cdot a_n \in \mathbb{Q}_v^* \Big/ \left[\mathbb{Q}_v^*\right]^2 \\
\varepsilon(f) &= \prod_{i<j} (a_i, a_j)_v \in \{\pm 1\}
\end{aligned}$$

are invariants of the equivalence class of f.

4.7.1 Representing p-adic numbers

Lemma 4.59. *Define*

$$r := \begin{cases} 2 & \text{if } p \neq 2 \\ 3 & \text{else.} \end{cases}$$

Furthermore, define for $a \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ *and* $\epsilon \in \{\pm 1\}$ *the set*

$$H_a^\epsilon := \left\{ x \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2 \;\middle|\; (x,a)_p = \epsilon \right\}.$$

Then the following statements hold:

(i). $\left| \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2 \right| = 2^r$

(ii). We have that for $a \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ *and* $\epsilon \in \{\pm 1\}$*:*

- *If* $a = 1$: $\left|H_a^1\right| = 2^r$ *and* $\left|H_a^{-1}\right| = 0$.
- *If* $a \neq 1$: $|H_a^\epsilon| = 2^{r-1}$.

(iii). Let $a, a' \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ *and* $\epsilon, \epsilon' \in \{\pm 1\}$*. If* H_a^ϵ *and* $H_{a'}^{\epsilon'}$ *are nonempty it holds that*

$$H_a^\epsilon \cap H_{a'}^{\epsilon'} = \varnothing \quad \Leftrightarrow \quad a = a' \text{ and } \epsilon = -\epsilon'.$$

Proof.

(i). This summarizes Corollaries 2.112 and 2.114.

(ii). The case $a = 1$ directly follows from Fact 3.7(i). Now let $a \neq 1$ and then by Theorem 3.13, we have that the homomorphism

$$\begin{array}{rccc} \varphi\colon & \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2 & \longrightarrow & \{\pm 1\} \\ & b & \longmapsto & (a,b)_p \end{array}$$

is onto. Thus $\ker(\varphi)$ is a hyperplane in $\mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ and has 2^{r-1} elements. Its complement H_a^{-1} has 2^{r-1} elements since it is an affine hyperplane parallel to H_a^1.

(iii). If both H_a^ϵ and $H_{a'}^{\epsilon'}$ are nonempty and disjoint, they have necessarily 2^{r-1} elements each and are complementary to one another, which implies $H_a^\epsilon = H_{a'}^{\epsilon'}$. Hence

$$\forall\, x \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2 : (x,a)_p = (x,a')_p .$$

This, again by Theorem 3.13, implies that $a = a'$ and $\epsilon = \epsilon'$. The converse is trivial. □

Now we will give equivalent conditions for a quadratic form f to represent zero. First note that by Theorem 4.49, we can assume without loss of generality that the rank is full and:

$$f = \sum_{i=1}^{n} a_i X_i^2 \quad \forall\, i \in [1:n] \colon a_i \neq 0.$$

We will handle the different numbers of variables separately. Each of the different cases for n will give a corollary — not only for $p \in P$ but also for $v \in V$: For $a \in \mathbb{Q}_v^* \Big/ \left[\mathbb{Q}_v^*\right]^2$ define $f_a := f \mathbin{\hat{\ominus}} aZ^2$ for a fresh variable Z. We know by Corollary 4.48 that f_a represents zero if and only if f represents a. The discriminant and ε_v evaluate to:

$$\operatorname{disc}(f_a) = -a \cdot \operatorname{disc}(f) \text{ and } \varepsilon_v(f_a) = (-a, \operatorname{disc}(f))_p \, \varepsilon_v(f). \tag{4.7.4}$$

Applying the following Lemmas 4.60, 4.62, 4.64 and 4.66 to f_a, we get the corresponding Corollaries 4.61, 4.63, 4.65 and 4.67.

Lemma 4.60. *$f \in \mathbb{Q}_p[X_1, X_2]^{=2}$ represents zero if and only if*

$$\operatorname{disc}(f) = -1.$$

Corollary 4.61. *For $a \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$, $f \in \mathbb{Q}_p[X_1]^{=2}$ represents a if and only if*

$$\operatorname{disc}(f) = a.$$

Proof. We are looking for $(x_1, x_2) \in \mathbb{Q}_p^2 \setminus \{0\}$ with

$$0 = a_1 x_1^2 + a_2 x_2^2 \quad \Leftrightarrow \quad -\frac{a_1}{a_2} x_1^2 = x_2^2$$

Since none of the x_i can be zero, such x_i can be found if and only if $-\frac{a_1}{a_2}$ is a square. But in $\mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ we can calculate

$$-\frac{a_1}{a_2} = -a_1 a_2 = -\operatorname{disc}(f).$$

Together with the fact that if $-\frac{a_1}{a_2}$ is a square, it is equal to 1 in $\mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$, we get $\operatorname{disc}(f) = -1$. □

Lemma 4.62. *$f \in \mathbb{Q}_p[X_1, X_2, X_3]^{=2}$ represents zero if and only if*

$$(-1, -\operatorname{disc}(f))_p = \varepsilon_p(f).$$

Corollary 4.63. *For $a \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$, $f \in \mathbb{Q}_p[X_1, X_2]^{=2}$ represents a if and only if*

$$(a, -\operatorname{disc}(f))_p = \varepsilon_p(f).$$

Proof. f represents zero if and only if

$$-a_3 f = -a_3a_1X_1^2 - a_3a_2X_2^2 - a_3^2X_3^2 \sim -a_3a_1X_1^2 - a_3a_2X_2^2 - X_3^2 =: g$$

represents zero. Simply by Definition 3.2, we have that g represents zero if and only if $(-a_3a_1, -a_3a_2)_p = 1$. Using the bilinearity of the Hilbert symbol, we get

$$\begin{aligned}
1 &= (-a_3a_1, -a_3a_2)_p \\
&= (-1,-1)_p\,(-1,a_3)_p\,(-1,a_2)_p\,(a_3,-1)_p \\
&\quad (a_3,a_3)_p\,(a_3,a_2)_p\,(a_1,-1)_p\,(a_1,a_3)_p\,(a_1,a_2)_p \\
&= (-1,-1)_p\,(-1,a_1)_p\,(-1,a_3)_p\,(-1,a_2)_p\,(-1,a_3)_p \\
&\quad (a_1,a_2)_p\,(a_1,a_3)_p\,(a_2,a_3)_p\,(a_3,a_3)_p \\
&= (-1,-1)_p\,(-1,a_1)_p\,(-1,a_2)_p\,(a_1,a_2)_p\,(a_1,a_3)_p\,(a_2,a_3)_p\,(a_3,a_3)_p \\
&\overset{3.9(\text{ii})}{=} (-1,-1)_p\,(-1,a_1)_p\,(-1,a_2)_p\,(a_1,a_2)_p\,(a_1,a_3)_p\,(a_2,a_3)_p\,(-1,a_3)_p \\
&= (-1,-a_1a_2a_3)_p\,(a_1,a_2)_p\,(a_1,a_3)_p\,(a_2,a_3)_p \\
&= (-1,-\operatorname{disc}(f))_p\,\varepsilon(f)
\end{aligned}$$

which is equivalent to saying that $(-1,-\operatorname{disc}(f))_p = \varepsilon(f)$. □

Lemma 4.64. *$f \in \mathbb{Q}_p[X_1,X_2,X_3,X_4]^{=2}$ represents zero if and only if*

$$\operatorname{disc}(f) \neq 1 \text{ or } \Big(\operatorname{disc}(f) = 1 \text{ and } \varepsilon_p(f) = (-1,-1)_p\Big).$$

Corollary 4.65. *For $a \in \mathbb{Q}_p^*\Big/\left[\mathbb{Q}_p^*\right]^2$, $f \in \mathbb{Q}_p[X_1,X_2,X_3]^{=2}$ represents a if and only if*

$$\operatorname{disc}(f) \neq -a \text{ or } \Big(\operatorname{disc}(f) = -a \text{ and } \varepsilon_p(f) = (-1,-\operatorname{disc}(f))_p\Big).$$

Proof. Define

$$g := a_1X_1^2 + a_2X_2^2 \quad \text{and} \quad h := -a_3X_3^2 - a_4X_4^2.$$

Since $f = g \mathbin{\hat{\ominus}} h$, Corollary 4.52 implies that f represents zero if there exists an element $x \in \mathbb{Q}_p^*\Big/\left[\mathbb{Q}_p^*\right]^2$ that is represented by g and by h. By Corollary 4.63 such an element is characterized by the conditions

$$\begin{aligned}
(x,-\operatorname{disc}(g))_p = \varepsilon(g) &\quad\Leftrightarrow\quad (x,-a_1a_2)_p = (a_1,a_2)_p \\
(x,-\operatorname{disc}(h))_p = \varepsilon(h) &\quad\Leftrightarrow\quad (x,-a_3a_4)_p = (-a_3,-a_4)_p
\end{aligned}$$

Now define the subsets of $\mathbb{Q}_p^*\Big/\left[\mathbb{Q}_p^*\right]^2$ that satisfy these two conditions:

$$\begin{aligned}
A &:= \left\{ x \in \mathbb{Q}_p^*\Big/\left[\mathbb{Q}_p^*\right]^2 \;\middle|\; (x,-a_1a_2)_p = (a_1,a_2)_p \right\} \\
B &:= \left\{ x \in \mathbb{Q}_p^*\Big/\left[\mathbb{Q}_p^*\right]^2 \;\middle|\; (x,-a_3a_4)_p = (-a_3,-a_4)_p \right\}
\end{aligned}$$

In order that f does not represent zero, it is necessary and sufficient that $A \cap B = \varnothing$. Note that A and B are nonempty (we have $a_1 \in A$ and $-a_3 \in B$ for example) and Lemma 4.59(iii) yields that $A \cap B = \varnothing$ is equivalent to

$$a_1a_2 = a_3a_4 \quad \text{and} \quad (a_1,a_2)_p = -(-a_3,-a_4)_p. \tag{4.7.5}$$

The first condition in (4.7.5) is equivalent to $\operatorname{disc}(f) = 1$. So if $\operatorname{disc}(f) \neq 1$, we have that $A \cap B \neq \varnothing$ and f represents zero. So assume $\operatorname{disc}(f) = 1$ and calculate

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

which is

$$-(-a_3, -a_4)_p (-1, -1)_p (-a_3, -a_4)_p = -(-1, -1)_p$$

if and only if $(a_1, a_2)_p = -(-a_3, -a_4)_p$, the second condition in (4.7.5). □

Lemma 4.66. *For every $n \in \mathbb{N}_{\geq 5}$ we have that $f \in \mathbb{Q}_p[X_1, \ldots, X_n]$ represents zero.*

Corollary 4.67. *For every $a \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ and $n \in \mathbb{N}_{\geq 4}$ we have that $f \in \mathbb{Q}_p[X_1, \ldots, X_n]$ represents a.*

Proof. We will prove the case $n = 5$ which then implies the statement for $n \geq 5$: If every quadratic form of rank 5 represents zero, every form of higher rank does so too. Let r and H_a^ϵ be as in Lemma 4.59 and observe that for every quadratic form $g = a_1X_1 + a_2X_2$ over $\mathbb{Q}_p$:

$$\begin{aligned}
H^{\varepsilon(g)}_{-\operatorname{disc}(g)} &= \left\{ x \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2 \;\middle|\; (x, -\operatorname{disc}(g))_p = \varepsilon(g) \right\} \\
&\overset{4.63}{=} \left\{ x \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2 \;\middle|\; g \text{ represents } x \right\}.
\end{aligned}$$

Now if $a_1a_2 = \operatorname{disc}(g) = -1$ we can calculate

$$\begin{aligned}
1 &\overset{3.9(\text{iv})}{=} (a_1, -a_1)_p = (a_1, -1)_p (a_1, a_1)_p = (a_1, a_1a_2)_p (a_1, a_1)_p \\
&= (a_1, a_1)_p (a_1, a_2)_p (a_1, a_1)_p = (a_1, a_2)_p = \varepsilon(g),
\end{aligned}$$

which means that there cannot be the case of H_1^{-1} in Lemma 4.59. Statement (ii) of this Lemma implies that g represents at least 2^{r-1} elements. Since this is true for forms of rank 2, it is also true for forms with greater rank e.g. f. Since $2^{r-1} \geq 2$, f represents at least one element $a \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ distinct from $\operatorname{disc}(f)$. We then have

$$f \sim aX^2 \oplus h$$

where h is a form of rank 4. The discriminant of h is $\frac{\operatorname{disc}(f)}{a}$ and since $a \neq \operatorname{disc}(f)$ we have $\operatorname{disc}(h) \neq 1$, which by 4.67 implies that h represents zero. The same is then true for f. □

Lemmas 4.60, 4.62, 4.64 and 4.66 can be summarized in the following theorem.

Theorem 4.68. *Let $f \in \mathbb{Q}_p[X_1, \ldots, X_n]^{=2}$ be of full rank. It is necessary and sufficient for f to represent zero that one of the following statements hold*

(i). $n = 2$ *and* $\operatorname{disc}(f) = -1$.

(ii). $n = 3$ *and* $\varepsilon(f) = (-1, -\operatorname{disc}(f))_p$.

(iii). $n = 4$ *and* $(\operatorname{disc}(f) \neq 1$ *or* $(\operatorname{disc}(f) = 1$ *and* $\varepsilon(f) = (-1,-1)_p))$.

(iv). $n \geq 5$.

Corollaries 4.61, 4.63, 4.65 and 4.67 can be summarized as:

Corollary 4.69. *Let $p \in V$, $a \in \mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$ and $f \in \mathbb{Q}_p[X_1, \ldots, X_n]^{=2}$ of full rank. In order that f represents a it is necessary and sufficent that one of the following holds:*

(i). $n = 1$ *and* $a = \operatorname{disc}(f)$.

(ii). $n = 2$ *and* $(a, -\operatorname{disc}(f))_p = \varepsilon(f)$.

(iii). $n = 3$ *and* $(a \neq -\operatorname{disc}(f)$ *or* $(a = -\operatorname{disc}(f)$ *and* $(-1, -\operatorname{disc}(f))_p = \varepsilon(f)))$.

(iv). $n \geq 4$.

Remark 4.70. As always, a and $\operatorname{disc}(f)$ are elements of $\mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$. So the inequality $a \neq -\operatorname{disc}(f)$ for example means that a is not equal to a product of $-\operatorname{disc}(f)$ with a square.

4.7.2 Classification of p-adic quadratic forms

This section will give a complete classification of the form equivalence for the quadratic p-adic case via the rank, discriminant and Hasse-Minkowski invariant of a quadratic form.

Theorem 4.71. *Two quadratic forms over $\mathbb{Q}_p$ are equivalent if and only if they have the same rank, same discriminant and the same invariant ε.*

Proof. From the definitions and Corollary 4.58 it follows that two forms have the same invariants follows. We will prove the converse by induction on the rank. So let $f, g \in \mathbb{Q}_p[X_1, \ldots, X_n]^{=2}$ with $f \sim g$ and $n := \operatorname{rank}(f) = \operatorname{rank}(g)$.

Step 1 ($n = 0$). The set $\mathbb{Q}_p^{=2}$ is empty, so the statement is trivially true.

Step 2 ($n-1 \to n$). Let the statement be true for $n-1$. By Corollary 4.69, we know that f and g represent the same — nonempty — set of elements of $\mathbb{Q}_p^* \Big/ \left[\mathbb{Q}_p^*\right]^2$. Let a be such an element. By Corollary 4.48, there exist $f', g' \in \mathbb{Q}_p\left[X_1, \ldots, X_{n-1}\right]$ such that

$$f \sim f' \oplus aZ^2 \quad \text{and} \quad g \sim g' \oplus aZ^2.$$

We see that

$$\begin{aligned} \operatorname{disc}(f') &= a \cdot \operatorname{disc}(f) = a \cdot \operatorname{disc}(g) = \operatorname{disc}(g'), \\ \varepsilon(f') &= \varepsilon(f)\,(a, \operatorname{disc}(f'))_p = \varepsilon(g)\,(a, \operatorname{disc}(g'))_p = \varepsilon(g') \end{aligned}$$

which shows that f' and g' have the same invariants. The statement follows by induction. □

4.7.3 Classification of real quadratic forms

For $v = \infty$ respectively $\mathbb{Q}_v = \mathbb{R}$, we can prove a statement similar to Theorem 4.68 and Corollary 4.69. For this, we first prove some formulas for the invariants in the real case.

Fact/Definition 4.72 (Signature of a quadratic form over $\mathbb{R}$). For a real quadratic form f of rank n, there exist $r, s \in \mathbb{N}_0$ with $r + s = n$ such that

$$f \sim X_1^2 + \cdots + X_r^2 - Y_1^2 - \cdots - Y_s^2.$$

The pair (r, s) is called **signature of** f. We say that f is **definite** if $r = 0$ or $s = 0$, i.e. if f does not change signs. Otherwise f is called **indefinite**, i.e. if f represents zero.

Proof. From Theorem 4.49 we know that

$$\exists\, a_1, \ldots, a_n \in \mathbb{R}\colon f \sim \sum_{i=1}^{n} a_i X_i^2.$$

First note that f is of rank n, so for all $i \in [1:n]$ we have that $a_i \neq 0$. Since $\mathbb{R}$ contains all square roots of positive numbers, we can use the following linear transformation of the variables to obtain the desired equivalent form:

$$\forall\, i \in [1:n]\colon \tau(X_i) := \frac{X_i}{\sqrt{|a_i|}}. \qquad \square$$

Lemma 4.73. *For $f \in \mathbb{R}\left[X_1, \ldots, X_n\right]^{=2}$ with signature (r, s) we have*

$$\begin{aligned} \varepsilon(f) = (-1)^{\frac{s(s-1)}{2}} &= \begin{cases} 1 & \text{if } s \equiv 0, 1 \pmod 4 \\ -1 & \text{if } s \equiv 2, 3 \pmod 4 \end{cases} \\ \operatorname{disc}(f) = (-1)^s &= \begin{cases} 1 & \text{if } s \equiv 0 \pmod 2 \\ -1 & \text{if } s \equiv 1 \pmod 2. \end{cases} \end{aligned}$$

Proof. By Theorem 3.11, we have that $(-1,-1)_\infty = -1$ and $(a,b)_\infty = 1$ if $a = 1$ or $b = 1$. So the formula for $\varepsilon(f)$ are evident. The formula for $\operatorname{disc}(f)$ is immediate by the definition of discriminant and signature. □

Theorem 4.74. *For $n \in [2:4]$ and $f \in \mathbb{R}[X_1,\dots,X_n]^{=2}$ to represent zero it is necessary and sufficient that one of the following statements hold*

(i). $n = 2$ *and* $\operatorname{disc}(f) = -1$.

(ii). $n = 3$ *and* $\varepsilon(f) = (-1, -\operatorname{disc}(f))_\infty$

(iii). $n = 4$ *and* $(\operatorname{disc}(f) = -1$ *or* $(\operatorname{disc}(f) = 1$ *and* $\varepsilon(f) = -1))$.

Proof.

Case 1 ($n = 2$). Let $f = a_1X_1 + a_2X_2$ where $a_1, a_2 \in \{\pm 1\}$. $\operatorname{disc}(f) = a_1a_2 = -1$ means that $a_1 \neq a_2$. Note that $0 = X_1^2 + X_2^2$ and $0 = -X_1^2 - X_2^2$ have $(1,1)$ as a nontrivial solution and that $0 = X_1^2 - X_2^2$ and $0 = -X_1^2 + X_2^2$ both do not have a solution over reals.

Case 2 ($n = 3$). Let $a = \varepsilon(f)$ and $b = (-1, -\operatorname{disc}(f))_\infty$. With Lemma 4.73 we get

$$a = (-1)^{\frac{s(s-1)}{2}}$$
$$b = (-1, -\operatorname{disc}(f))_\infty = (-1, -(-1)^s)_\infty = \left(-1, (-1)^{s-1}\right)_\infty$$

and it remains to show that f represents zero if and only if $a = b$:

s	f	Repr.	a	b	
0	$X_1^2 + X_2^2 + X_3^2$	none	1	$(-1,-1)_\infty = -1$	$a \neq b$
1	$X_1^2 + X_2^2 - X_3^2$	$(0,1,1)$	1	$(-1,1)_\infty = 1$	$a = b$
2	$X_1^2 - X_2^2 - X_3^2$	$(1,0,1)$	-1	$(-1,-1)_\infty = -1$	$a = b$
3	$-X_1^2 - X_2^2 - X_3^2$	none	-1	$(-1,1)_\infty = 1$	$a \neq b$

Comparing the "Repr." and the last column yields the statement.

Case 3 ($n = 4$). Lemma 4.73 translates

$$\operatorname{disc}(f) = -1 \text{ or } (\operatorname{disc}(f) = 1 \text{ and } \varepsilon(f) = -1)$$

into

$$(-1)^s = -1 \text{ or } ((-1)^s = 1 \text{ and } (-1)^{\frac{s(s-1)}{2}} = -1)$$

which can be reformulated as

$$s \text{ is odd or } (s \text{ is even and } \frac{s(s-1)}{2} \text{ is odd })$$

The following table then finishes the proof

s	f	Representation	$\frac{s(s-1)}{2}$
0	$X_1^2 + X_2^2 + X_3^2 + X_4^2$	none	0
1	$X_1^2 + X_2^2 + X_3^2 - X_4^2$	$(0,0,1,1)$	0
2	$X_1^2 + X_2^2 - X_3^2 - X_4^2$	$(0,1,0,1)$	1
3	$X_1^2 - X_2^2 - X_3^2 - X_4^2$	$(1,0,0,1)$	3
4	$-X_1^2 - X_2^2 - X_3^2 - X_4^2$	none	6

□

Example 4.75. The last part of Theorem 4.68 does not translate to $\mathbb{R}$, since for example the form

$$X_1^2 + X_2^2 + X_3^2 + X_4^2$$

does not represent zero over $\mathbb{R}$.

Corollary 4.76. *Let $a \in \mathbb{R}^* \Big/ \left[\mathbb{R}^*\right]^2$. In order that $f \in \mathbb{R}\left[X_1, \ldots, X_n\right]^{=2}$ of full rank represents a it is necessary and sufficient that one of the following holds:*

(i). $n = 1$ *and* $a = \operatorname{disc}(f)$,

(ii). $n = 2$ *and* $(a, -\operatorname{disc}(f))_\infty = \varepsilon(f)$,

(iii). $n = 3$ *and either* $a \neq -\operatorname{disc}(f)$ *or* $(a = -\operatorname{disc}(f)$ *and* $(-1, -\operatorname{disc}(f))_\infty = \varepsilon(f))$,

Proof. Use the argument from (4.7.4) and apply Theorem 4.74 to f_a. □

4.8 Quadratic forms over rational numbers

In this section, we will give the theoretical results that are the core of Algorithm 6 that will decide quadratic form equivalence over rational numbers in polynomial time (using an oracle for IntFact). The algorithm heavily uses the fact that two such forms are equivalent if and only if they have the same invariants sig, disc and ε, which will also be proven in this section. The complexity of the computation of these invariants is discussed in Section 5.1.

Definition 4.77 (Local Invariants). Let $f \in \mathbb{F}\left[X_1, \ldots, X_n\right]^{=2}$ be a quadratic form of full rank and write

$$f \sim \sum_{i=1}^{n} a_i X_i^2.$$

We now associate the following invariants

- The discriminant $\operatorname{disc}(f) \in \mathbb{Q}^* \Big/ \left[\mathbb{Q}^*\right]^2$ which is $\prod_{i=1}^n a_i$.
- For $v \in V$ we use the injection $\mathbb{Q} \mapsto \mathbb{Q}_v$ to view f as a form over $\mathbb{Q}_v$ (which we will denote by f_v). We denote the invariants of f_v by $\operatorname{disc}_v(f)$ and $\varepsilon_v(f)$. It is clear that $\operatorname{disc}_v(f)$ is the image of $\operatorname{disc}(f)$ under

  $$\mathbb{Q}^* \Big/ \left[\mathbb{Q}^*\right]^2 \to \mathbb{Q}_v^* \Big/ \left[\mathbb{Q}_v^*\right]^2.$$

  We have

  $$\varepsilon_v(f) = \prod_{i<j} (a_i, a_j)_v.$$

- When we view f as a real quadratic form, $\operatorname{sig}(f)$ is another invariant.

Theorem 4.78 (Local-Global-Principle / Hasse-Minkowski Theorem). *In order for $f \in \mathbb{Q}[X_1, \ldots, X_n]^{=2}$ to represent zero it is necessary and sufficient that, for all $v \in V$, the form f_v represents zero. Or in other words:*

$$f \text{ represents zero} \quad \Longleftrightarrow \quad \forall\, v \in V\colon f_v \text{ represents zero} \ .$$

Proof. The necessity is trivial. To see the sufficiency, write f as

$$f = \sum_{i=1}^{n} a_i X_i^2, \quad \text{with } a_i \in \mathbb{Q}^* \ \forall\, i[1:n].$$

Note that f represents zero if and only if $a_1 f$ represents zero. But the form

$$a_1 f = a_1^2 X_1^2 + \sum_{i=2}^{n} a_1 a_i X_i^2$$

represents zero if and only if

$$g = X_1^2 + \sum_{i=2}^{n} b_i X_i^2 \quad \text{with } b_i := a_1 a_i \ \forall\, i[2:n]$$

represents zero — a nontrivial zero $(x_1, \ldots, x_n) \in \mathbb{Q}^n$ of $a_1 f$ yields a nontrivial zero $(a_1 x_1, \ldots, x_n) \in \mathbb{Q}^n$ of g and vice versa. So without loss of generality, we can assume that $a_1 = 1$. Consider the following 4 cases:

Case 1 ($n = 2$). We have $f = X_1^2 - aX_2^2$ since f_∞ represents zero we have that $a > 0$. Write a as

$$a = \prod_{p \in P} p^{\operatorname{ord}_p(a)}. \tag{4.8.1}$$

Since for every $p \in P$ the form f_p represents zero, we know that there exists $(x_1, x_2) \in \mathbb{Q}_p^2 \setminus \{\, \mathbf{0} \,\}$ such that $0 = x_1^2 - ax_2^2$. If $x_2 = 0$, we would have that $x_1 = 0$ which contradicts the fact that $(x_1, x_2) \neq (0, 0)$. So a is a square in $\mathbb{Q}_p$ since

$$0 = x_1^2 - ax_2^2 \quad \Leftrightarrow \quad ax_2^2 = x_1^2 \quad \Leftrightarrow \quad a = \left(\frac{x_1}{x_2}\right)^2.$$

This implies that $\operatorname{ord}_p(a)$ is even for every $p \in P$ which means by (4.8.1) that a is a square in $\mathbb{Q}$. Let $c \in \mathbb{Q}$ be such that $c^2 = a$ and one sees that $(c, 1)$ is a nontrivial zero of f.

Case 2 ($n = 3$ — "Legendre case"). This case is very similar to Section 3.4 but much less explicit: Write $f = X_1^2 - aX_2^2 - bX_3^2$. Utilizing Notation 3.23 every nontrivial zero (x_1, x_2, x_3) of

$$g = X_1^2 - \overline{a}X_2^2 - \overline{b}X_3^2$$

yields the nontrivial $(x_1, \widetilde{a}x_2, \widetilde{b}x_3)$ zero of f and vice versa. So without loss of generality assume that a and b are squarefree i.e. that for all $p \in P$ we have that $\operatorname{ord}_p(a), \operatorname{ord}_p(b) \in \{\, 0, 1 \,\}$. By symmetry additionally assume that $|a| \leq |b|$. Set $m := |a| + |b|$ and use induction on m:

Step 1 ($m = 2$). We have that

$$f = X_1^2 \pm X_2^2 \pm X_3^2$$

where the case that $f = X_1^2 + X_2^2 + X_3^2$ is excluded because f_∞ represents zero. In all other cases, f represents zero:

$X_1^2 + X_2^2 - X_3^2$	has $(0, 1, 1)$ as a solution.
$X_1^2 - X_2^2 + X_3^2$	has $(1, 1, 0)$ as a solution.
$X_1^2 - X_2^2 - X_3^2$	has $(1, 0, 1)$ as a solution.

Step 2 ($m > 2$). For some $k \in \mathbb{N}$ we can write:

$$b = \pm \prod_{i=1}^{k} p_i \quad \text{where } p_i \in P$$

and the p_i are pairwise different since b is squarefree. Let p be one of the p_i. We will prove that a is a square modulo p. This is obvious if $a \equiv 0 \pmod{p}$. Otherwise, a defines a p-adic unit. By the assumption of this theorem, there exists a solution $(x, y, z) \in \mathbb{Q}_p^3$ such that $z^2 - ax^2 - by^2 = 0$ and by Proposition 2.98 we can assume that it is in $\mathbb{Z}_p^3$ and primitive. By the definition of p we have that $z^2 - ax^2 \equiv 0 \pmod{p}$. Now if $x \equiv 0 \pmod{p}$, this implies that $z^2 \equiv 0 \pmod{p}$ which means that $z \equiv 0 \pmod{p}$. Reduction of $z^2 - ax^2 - by^2 = 0$ by p^2 therefore yields that $by^2 \equiv 0 \pmod{p^2}$. But since $\operatorname{ord}_p(b) = 1$ this means that $y^2 \equiv 0 \pmod{p}$ and $y \equiv 0 \pmod{p}$ contrary to the fact that (x, y, z) is primitive. So we know that $x \not\equiv 0 \pmod{p}$, which together with $z^2 - ax^2 \equiv 0 \pmod{p}$ implies that a is a square modulo p. By the Chinese Remainder Theorem we have that $\mathbb{Z}/b\mathbb{Z} \cong \prod_{i=1}^{k} \mathbb{Z}/p_i\mathbb{Z}$ which now implies that a is a square modulo b, i.e.

$$\exists\, t, b' \in \mathbb{Z} \colon t^2 = a + bb'$$

and we can choose t such that $|t| \leq \frac{|b|}{2}$. For $\mathbb{F} = \mathbb{Q}$ or $\mathbb{F} = \mathbb{Q}_v$ for some $v \in V$, the formula $bb' = t^2 - a$ shows that bb' is the norm of $t + \sqrt{a}$ in the extension $\mathbb{F}[\sqrt{a}]/\mathbb{F}$ which implies

$$\begin{aligned}
& g := X_1^2 - aX_2^2 - b'X_3^2 \text{ represents zero} \\
\overset{3.8}{\iff} \quad & b' \text{ is a norm in } \mathbb{F}[\sqrt{a}] \\
\overset{(*)}{\iff} \quad & b \text{ is a norm in } \mathbb{F}[\sqrt{a}] \\
\overset{3.8}{\iff} \quad & f = X_1^2 - aX_2^2 + bX_3^2 \text{ represents zero}
\end{aligned}$$

Where $(*)$ holds because bb' is a norm in $\mathbb{F}[\sqrt{a}]$. In particular, g represents zero in each of the $\mathbb{Q}_v$. But we have

$$\begin{aligned}
|b'| &= \left|\tfrac{t^2 - a}{b}\right| \leq \tfrac{\left|t^2\right| + |a|}{|b|} \overset{|t| \leq \frac{|b|}{2}}{\leq} \tfrac{\frac{|b|^2}{4} + |a|}{|b|} \\
&\leq \tfrac{|b|}{4} + \tfrac{|a|}{|b|} \overset{|a| \leq |b|}{\leq} \tfrac{|b|}{4} + 1 \overset{|b| \geq 2}{<} |b|
\end{aligned}$$

Now write $b' = \overline{b'}\widetilde{b'}^2$ where $\widetilde{b'}, \overline{b'} \in \mathbb{Z}$ and $\overline{b'}$ is squarefree. We have that $\left|\overline{b'}\right| \leq |b'| < |b|$. Now apply the induction hypothesis to the form

$$h := X_1^2 - aX_2^2 - \overline{b'}X_3^2.$$

We have that $h \sim g$, hence h represents zero in $\mathbb{Q}$ implying that the same is true for f.

Case 3 ($n = 4$). Write

$$f = aX_1^2 + bX_2^2 - (cX_3^2 + dX_4^2)$$

and let $v \in V$. Since f_v represents zero, Corollary 4.52 shows that there exists $x_v \in \mathbb{Q}_v^*$ which is represented both by $aX_1^2 + bX_2^2$ and by $cX_3^2 + dX_4^2$. By Corollary 4.69(ii) and Corollary 4.76(ii) this is equivalent to saying that

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

Applying Theorem 3.20 it follows that

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

By Theorem 3.22 this implies the existence of $x \in \mathbb{Q}^*$ such that

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

The form $aX_1^2 + bX_2^2 - xZ^2$ represents zero in each of the $\mathbb{Q}_v$ hence by Case 2 it represents zero in $\mathbb{Q}$. So again by Corollary 4.52 we have that $aX_1^2 + bX_2^2$ represents x. The same argument applies to $cX_3^2 + dX_4^2$ implying that f represents zero.

Case 4 ($n \geq 5$). We use induction on n. Write f as

$$f = h \mathbin{\hat{\ominus}} g \quad \text{with } h = a_1X_1^2 + a_aX_2^2 \text{ and } g = -\sum_{i=3}^{n} a_iX_i^2.$$

Define $S \subseteq V$ as

$$S := \{\, p \in P \mid \exists\, i \in [3:n]\colon \operatorname{ord}_p(a_i) \neq 0 \,\} \cup \{\, 2, \infty \,\}\,.$$

This is a finite set. Let $v \in S$. Since f_v represents zero, by Corollary 4.52 there exists $a_v \in \mathbb{Q}_v^*$ which is represented in $\mathbb{Q}_v$ by h and by g i.e. there exists for every $i \in [1:n]$ an $x_{i,v} \in \mathbb{Q}_v$ such that

$$h(x_{1,v}, x_{2,v}) = a_v = g(x_{3,v}, \ldots, x_{n,v})\,.$$

But the squares of $\mathbb{Q}_v^*$ form an open set by Remark 2.115. By the Approximation Theorem (Lemma 3.21), there exists $x_1, x_2 \in \mathbb{Q}$ such that for $a = h(x_1, x_2)$ one has

$$\forall\, v \in S\colon \frac{a}{a_v} \in [\mathbb{Q}_v^*]^2\,.$$

Now define $f_1 := aZ^2 \mathbin{\hat{\ominus}} g$.

Case 4.1 ($v \in S$). If $v \in S$, we have that g represents a_v over $\mathbb{Q}_v$. But then g also represents a because $\frac{a}{a_v}$ is a square in $\mathbb{Q}_v$.

Case 4.2 ($v \notin S$). The coefficients $-a_3, \ldots, -a_n$ of g are v-adic units, and so is $\operatorname{disc}_v(g)$. Because $v \neq 2$, we then have $\varepsilon_v(g) = 1$.

In both cases, f_1 represents zero in $\mathbb{Q}_v$. Since $\operatorname{rank}(f_1) = n-1$, the induction hypothesis yields that f_1 represents zero in $\mathbb{Q}$. This means that g represents a in $\mathbb{Q}$ and since h represents a, f represents zero, which finishes the proof. □

Corollary 4.79. *Let $a \in \mathbb{Q}^*$ and $f \in \mathbb{Q}[X_1, \ldots, X_n]^{=2}$. Then it follows that f represents a if and only if it does in each of the $\mathbb{Q}_v$*

Proof. Apply the Local-Global-Principle / Hasse-Minkowski Theorem (Theorem 4.78) to $aZ^2 \mathbin{\hat{\ominus}} f$. □

Corollary 4.80 (Meyer). *For all $f \in \mathbb{Q}[X_1, \ldots, X_n]^{=2}$ with $\operatorname{rank}(f) \geq 5$ it holds that f represents zero if and only if it is indefinite (i.e. it represents zero over $\mathbb{R}$).*

Proof. By Theorem 4.68 (or Lemma 4.66 to be more precise) quadratic forms with rank ≥ 5 represent zero over every $\mathbb{Q}_p$ for $p \in P$. Together with the fact that f is indefinite, the Local-Global-Principle / Hasse-Minkowski Theorem (Theorem 4.78) implies the statement. □

Remark 4.81. The name of Theorem 4.78 (Local-Global-Principle) comes from the following reformulation: f has a "global" zero if and only if f has a "local" zero everywhere.

Theorem 4.82. *Two quadratic forms over $\mathbb{Q}$ are equivalent if and only if they are equivalent over $\mathbb{Q}_v$ for every $v \in V$.*

Proof. The necessity is obvious since for any $v \in V$ we have that $\mathbb{Q} \subseteq \mathbb{Q}_v$. So it is left to show that

$$\forall\, n \in \mathbb{N}_0 \colon \forall\, f, g \in \mathbb{Q}[X_1, \ldots, X_n]^{=2} \colon \quad (f \sim_{\mathbb{Q}} g \Rightarrow \forall\, v \in V \colon f \sim_{\mathbb{Q}_v} g)\,.$$

Since f and g are equivalent over $\mathbb{Q}$, they share the same rank. We now use induction on the rank $n := \operatorname{rank}(f) = \operatorname{rank}(g)$:

Step 1 ($n = 0$). The set $\mathbb{Q}^{=2}$ is empty and therefore the statement is trivially true.

Step 2 ($n - 1 \to n$ for $n \in \mathbb{N}$). Let $f, g \in \mathbb{Q}[X_1, \ldots, X_n]^{=2}$ with $f \sim_{\mathbb{Q}} g$. Then there exists $a \in \mathbb{Q}^*$ that is represented by f and by g. By Corollary 4.79 we have that for every $v \in V$ the forms f and g represent a over $\mathbb{Q}_v$ too. Then by Corollary 4.48 there exist $f', g' \in \mathbb{Q}_v[X_1, \ldots, X_{n-1}]^{=2}$ such that

$$f \sim_{\mathbb{Q}_v} aX_n^2 \mathbin{\hat{\ominus}} f' \quad \text{and} \quad g \sim_{\mathbb{Q}_v} aX_n \mathbin{\hat{\ominus}} g'.$$

By Witt's Cancellation Theorem (Theorem 4.53) it follows that $f \sim_{\mathbb{Q}_v} g$ if and only if $f' \sim_{\mathbb{Q}_v} g'$. Since $\operatorname{rank}(f') = \operatorname{rank}(g') = n - 1$ the induction hypothesis finishes the proof. □

Theorem 4.83. *Let $f, f' \in \mathbb{Q}[X_1, \ldots, X_n]^{=2}$ with signatures (r, s) and (r', s'). For f and f' to be equivalent it is necessary and sufficient that*

(i). $\operatorname{disc}(f) = \operatorname{disc}(f')$,

(ii). $(r, s) = (r', s')$ *and*

(iii). $\forall\, v \in V : \varepsilon_v(f) = \varepsilon_v(f')$.

Proof. By Theorem 4.71 the conditions just mean that f and f' are equivalent over every $\mathbb{Q}_v$, which by Theorem 4.82 implies the statement. □

Chapter 5

Results

Abstract

In this chapter, we will present the details about finding rational quadratic form equivalence and the proof that IntFact can be done in randomized polynomial time if $\text{FuncQuadFormEquiv}_{\mathbb{Q}} \in \mathbf{FP}$. This lower bound

$$\text{IntFact} \leq_{\mathbf{R}} \text{FuncQuadFormEquiv}_{\mathbb{Q}}$$

seriously improves the result by [Har08] where it is only shown that one can calculate $\sqrt{-1} \pmod{n}$ using an oracle for IntFact and that

$$\text{FuncQuadFormEquiv}_{\mathbb{Q}} \leq \text{IntFact}.$$

5.1 Decide rational quadratic form equivalence

In this section, we will see an algorithm that decides rational quadratic form equivalence in polynomial time only using an oracle for $\text{HilbertSymbol}_{\mathbb{Q}}$ (which can be replaced by an oracle for IntFact by Theorem 3.34).

Algorithm 6 Decide-Quadratic-Form-Equivalence

Input: $f, g \in \mathbb{Q}[X_1, \ldots, X_n]^{=2}$.
Output: **true** if $f \sim g$, **false** else.

1: Assume $f = \sum_{i=1}^{n} a_i X_i^2$ and $g = \sum_{i=1}^{n} b_i X_i^2$ with $a_i, b_i \in \mathbb{Q}$.
2: **assert** $\operatorname{rank}(f) = \operatorname{rank}(g)$
3: Without loss of generality assume $n = \operatorname{rank}(f)$ i.e. $a_i, b_i \in \mathbb{Q}^*$.
4: **assert** $\operatorname{sig}(f) = \operatorname{sig}(g)$
5: **assert** $\operatorname{disc}(f) = \operatorname{disc}(g)$
6: **calculate** $\forall\, i, j \in [1:n] \colon \mathcal{H}_{a_i,a_j}, \mathcal{H}_{b_i,b_j}$
7: **for** $v \in \mathcal{H}_{a_i,a_j} \cup \mathcal{H}_{b_i,b_j}$ **do**
8: **assert** $\prod_{i<j} (a_i, a_j)_v = \prod_{i<j} (b_i, b_j)_v$
9: **end for**
10: **return true**

Theorem 5.1. *Algorithm 6 is correct.*

Proof. This follows directly from Theorem 4.83. □

Theorem 5.2. *Given an oracle for* HilbertSymbol$_{\mathbb{Q}}$ *Algorithm 6 runs in polynomial time in the bitsize of the input n. If the coefficients of f and g are bound by some constant, it runs in $\mathcal{O}\left(n^3\right)$ even without an oracle for* HilbertSymbol$_{\mathbb{Q}}$.

Proof. We simply list the argumentation of efficiency for every step:

Case 1 (Step 1). This is polynomial in the bitsize of the input and cubic for constant coefficients by Fact 4.51.

Case 2 (Step 2). We only have to count the a_i and b_i that are nonzero. If and only if this number is that same for both f and g, the ranks coincide. This is linear.

Case 3 (Step 3). Permute the variables such that

$$\exists\, k \in [1:n] \colon \forall\, i \in [1:k] \colon a_i, b_i \in \mathbb{Q}^*$$

and this k is the same for f and g by the previous step. Now we can consider f and g as forms in $\mathbb{Q}[X_1, \ldots, X_k]$ and assume that $k = n$ in the first place. Permutation of the variables is linear too.

Case 4 (Step 4). Simply count the signs of a_i and b_i. If and only the number of positive coefficients coincides for f and g, the signatures are the same. This is linear.

Case 5 (Step 5). First calculate the products

$$\pm\frac{\alpha_1}{\alpha_2} := \prod_{i=1}^{n} a_i = \operatorname{disc}(f) \qquad \text{and} \qquad \pm\frac{\beta_1}{\beta_2} := \prod_{i=1}^{n} b_i = \operatorname{disc}(g)$$

for positive $\alpha_1, \alpha_2, \beta_1, \beta_2$. We now have to check whether $\pm\frac{\alpha_1}{\alpha_2} = \pm\frac{\beta_1}{\beta_2}$ modulo squares in $\mathbb{Q}^*$ (see Remark 4.10). This is equivalent to checking whether $\pm\frac{\alpha_1\beta_2}{\alpha_2\beta_1}$ is a square in $\mathbb{Q}^*$ or not. For this, we only have to check that all of the following conditions hold:

- The sign is $+$,
- $\alpha_1\beta_2$ is a square in $\mathbb{Z}$,
- $\alpha_2\beta_1$ is a square in $\mathbb{Z}$.

This is polynomial in the bitsizes by Algorithm 3. For bound coefficients, the whole process is linear.

Case 6 (Step 6). This is quadratic given an oracle for IntFact by Theorem 3.17. By the fact that IntFact takes constant time for a constantly bounded input and the same theorem for bound coefficients, this step takes constant time.

Case 7 (Step 8). The products can easily be calculated since a factor $(a_i, a_j)_v$ is -1 if and only if $v \in \mathcal{H}_{a_i,a_j}$. So this boils down to comparing the parity of $\left|\left\{\, v \in \mathcal{H}_{a_i,a_j} \;\middle|\; 1 \le i < j \le n \,\right\}\right|$ and $\left|\left\{\, v \in \mathcal{H}_{b_i,b_j} \;\middle|\; 1 \le i < j \le n \,\right\}\right|$, which is polynomial in the input and constant for bound coefficients. □

5.2 Find rational quadratic form equivalence

In this section, we will present an algorithm to find a rational quadratic form equivalence very explicitly. The algorithm reduces the problem of finding an equivalence to the problem of finding a solution to a rational equation of the form

$$\sum_{i=1}^{n} a_i X_i = b \quad \text{where } a_i, b \in \mathbb{Q}.$$

Since by [Sim05] this can be done in polynomial time using an oracle for IntFact, the whole algorithm can be implemented in polynomial time (with such an oracle). Additionally many results from Chapter 4 will then prove the correctness of the algorithm.

Fact 5.3. *For $f, g \in \mathbb{F}[X]^{=2}$ with $f(x) = ax^2$ and $g(x) = bx^2$ we have that*

$$f \sim g \Leftrightarrow \frac{a}{b} \in \mathbb{F}^{*2}.$$

This criterium can be checked in polynomial time for $\mathbb{F} = \mathbb{Q}$.

Proof. Being equivalent for these two forms means that there exists a number $\lambda \in \mathbb{F}^*$ such that $g(\lambda x) = f(x)$:

$$g(\lambda x) = f(x) \Leftrightarrow b(\lambda x)^2 = ax^2 \Leftrightarrow b\lambda^2 x^2 = ax^2 \Leftrightarrow b\lambda^2 = a \Leftrightarrow \lambda^2 = \frac{a}{b}.$$

The last equation is solvable if and only if $\frac{a}{b}$ is a square in $\mathbb{F}$ i.e. $\frac{a}{b} \in \mathbb{F}^{*2}$. For $\mathbb{F} = \mathbb{Q}$ this is the case if and only if the nominator and denominator are squares in $\mathbb{Z}$. To check this one can simply perform a binary search which can be done in linear time in the number of digits by Algorithm 3. □

Theorem 5.4. *Algorithm 7 is correct and, given an oracle for solving diagonal quadratic equations, runs in polynomial time.*

Proof. Correctness is obvious. All steps are just linear algebra except 8, where we need the oracle to solve a diagonal quadratic equation, so the polynomial running time follows instantly. □

Algorithm 7 FIND-QUADRATIC-FORM-EQUIVALENCE

Input: $f, g \in \mathbb{Q}[X_1, \ldots, X_n]^{=2}$.
Output: **true** if $f \sim g$, **false** else.

1: By Fact 4.51 assume $f = \sum_{i=1}^{n} a_i X_i^2$ and $g = \sum_{i=1}^{n} b_i X_i^2$ with $a_i, b_i \in \mathbb{Q}$.
2: Without loss of generality **set** $n = \operatorname{rank}(f)$ and permute the variables such that $a_i, b_i \in \mathbb{Q}^* \ \forall\, i \in [1:n]$.
3: **assert** $\operatorname{rank}(f) = \operatorname{rank}(g)$
/* Fact 4.50 */
4: **if** $\operatorname{rank}(f) = 1$ **then**
5: Write $f(x) = ax^2$ and $g(x) = bx^2$
6: **return** truth value of $\frac{a}{b} \in \mathbb{Q}^{*2}$
/* Fact 5.3 */
7: **end if**
8: Let $\alpha \in \mathbb{Q}^n$ be a solution for the diagonal quadratic equation $f(X_1, \ldots, X_n) = b_n$.
9: The subspace $U := \langle \alpha \rangle^{\perp}$ is nondegenerate since $b_n \neq 0$, from which it follows by Proposition 4.19(ii) that we have

$$V = \langle \alpha \rangle \oplus U$$

So every $v \in V$ can be written as $v = \lambda\alpha + u$ with $\lambda \in \mathbb{Q}$ and $u \in U$. Thus

$$\begin{aligned} f(v) = v.v &= (\lambda\alpha + u).(\lambda\alpha + u) = \lambda^2 \alpha.\alpha + u.u \\ &= \lambda^2 f(\alpha) + f(u) = \lambda^2 b_n + f(u). \end{aligned}$$

This simply means that $f \sim b_n X_n^2 \oplus f_1(X_1, \ldots, X_{n-1})$ for some quadratic form $f_1 \in \mathbb{Q}[X_1, \ldots, X_{n-1}]$.
10: Now we have

$$\begin{aligned} f &\sim b_n X_n^2 \oplus f_1(X_1, \ldots, X_{n-1}) \\ g(x_1, \ldots, x_n) &= b_n X_n^2 \oplus \sum_{i=1}^{n-1} b_i X_i^2 \end{aligned}$$

Witt's Cancellation Theorem allows us to calculate:

$$\begin{aligned} b_n X_n^2 \oplus f_1(X_1, \ldots, X_{n-1}) &\sim b_n X_n^2 \oplus \sum_{i=1}^{n-1} b_i X_i^2 \\ \iff \quad f_1(X_1, \ldots, X_{n-1}) &\sim \sum_{i=1}^{n-1} b_i X_i^2 \end{aligned}$$

11: **set** $f := f_1, g := \sum_{i=1}^{n-1} b_i X_i^2$ and **goto** 1.

5.3 IntFact $\in$ ZPP$^{\text{FuncQuadFormEquiv}_{\mathbb{Q}}}$

In this section we will show that, should the oracle for IntFact be unnecessary in the polynomialtime algorithm for rational quadratic form equivalence, we can solve IntFact in randomized polynomial time. For this, we will first consider an algorithm to calculate square roots of invertible elements modulo n. This implies by standard methods that we can factor integers in randomized polynomial time (see for example [CP05, p. 309, ex. 6.5]). All this can be summarized in the following theorem that is a corollary of Theorems 5.8 and 5.10:

Theorem 5.5.

$$\text{FuncQuadFormEquiv}_{\mathbb{Q}} \in \mathbf{FP} \Rightarrow \text{IntFact} \in \mathbf{ZPP}.$$

5.3.1 SqrtMod* using FuncQuadFormEquiv$_{\mathbb{Q}}$

Lemma 5.6. *For $n \in \mathbb{N}_{>1}$ the inverse of a number $a \in \mathbb{Z}/n\mathbb{Z}$ or the fact that it does not exist, can be found in polynomial time.*

Proof. First calculate $g := \gcd(a, n)$. If $g \neq 1$ we have that a is not invertible. Otherwise by definition, the modular multiplicative inverse x of a is the solution of

$$ax \equiv 1 \pmod{n}.$$

This is equivalent to $n \mid ax - 1$, i.e. that n is a divisor of $ax - 1$, which means

$$\begin{aligned} & ax - 1 = qn \\ \Leftrightarrow \quad & ax - nq = 1 = g, \end{aligned}$$

With the extended Euclidean algorithm, one can find the numbers x and q for given a and n. We discard the latter and get the inverse of x. Since we only used the extended Euclidean algorithm, this process is polynomial in the bitsize of the input. □

Algorithm 8 Square-Root-Modulo-n-With-QuadFormEquiv-Oracle

Input: $n \in \mathbb{Z}, a \in (\mathbb{Z}/n\mathbb{Z})^*$.

Output: $\begin{cases} \textbf{false} & \text{if } a \text{ is not a square} \\ \sqrt{a} \pmod{n} & \text{else.} \end{cases}$

1: $b := a^{-1}$ by Lemma 5.6
2: $f := bX^2 + nY^2$
   $g := X^2 + bnY^2$
3: **assert** $f \sim g$
4: Calculate an equivalence τ for f and g
5: **return** $(\tau(X))(1, 0)$.

Theorem 5.7. *Algorithm 8 is correct and runs in polynomial time in the bitsize of the input with an oracle for* FuncQuadFormEquiv$_{\mathbb{Q}}$.

Proof. To prove correctness, let τ be given by

$$\tau(X) = \alpha_1 X + \alpha_2 Y$$
$$\tau(Y) = \beta_1 X + \beta_2 Y$$

and explicitly write down $f(\tau(X), \tau(Y)) = g(X, Y)$:

$$b(\alpha_1 X + \alpha_2 Y)^2 + n(\beta_1 X + \beta_2 Y)^2 = X^2 + bnY^2.$$

Now set $X = 1$ and $Y = 0$:

$$1 = b\alpha_1^2 + n\beta_1^2$$

Reduce the equation modulo n to get

$$b\alpha_1^2 \equiv 1 \pmod{n} \quad \Leftrightarrow \quad \alpha_1^2 \equiv b^{-1} \equiv a \pmod{n}.$$

So $\alpha_1 = (\tau(X))(1, 0)$ is a root of $a \pmod{n}$. To finish the proof of correctness, we have to observe, that $a\mathrm{R}n$ is equivalent to $f \sim g$. So let α_1 be a root of a. First observe that α_1 is invertible too, since it has $\alpha_1 a^{-1}$ as an inverse. Now define an equivalence by

$$\tau(X) = \alpha_1 X \quad \text{and} \quad \tau(Y) = \alpha_1^{-1} Y.$$

The statement about the running time is obvious since the hardest step is to find the equivalence τ, for which we are allowed to use the oracle. □

We summarize this result in the following theorem. Recall that SqrtMod* is the special case of invertible instances of SqrtMod.

Theorem 5.8. *It holds that:*

$$\text{SqrtMod}^* \in \mathbf{FP}^{\text{FuncQuadFormEquiv}_{\mathbb{Q}}}.$$

5.3.2 Factoring integers using SqrtMod*

In this chapter we will give a randomized polynomial time algorithm for integer factoring with the help of an oracle for taking square roots modulo n. Given a composite number n, the basic idea is to take a random element $r \neq 1$ in $(\mathbb{Z}/n\mathbb{Z})^*$ and square it. The resulting number has 4 different square roots and the oracle has some chance not to choose $\pm r$, but one of the other two roots. It will turn out that we can then determine a factor of n using Euclidean's algorithm.

Algorithm 9 IntFact-With-Sqrt-Mod-n-Oracle

Input: $n = pq \in \mathbb{Z}$ with p, q.
Output: p or q or **DoNotKnow**.

1: **if** $2 \mid n$ **then**
2: **return** 2
3: **end if**
4: **if** $\sqrt{n} \in \mathbb{Z}$ **then**
5: **return** $\sqrt{n}$
6: **end if**
7: $r \in_R (\mathbb{Z}/n\mathbb{Z})^* \setminus \{1\}$
8: **if** $r \mid n$ **then**
9: **return** r
10: **end if**
11: $s := r^2 \pmod{n}$
12: Query the oracle for $t := \sqrt{s} \pmod{n}$
13: **if** $t \equiv \pm r \pmod{n}$ **then**
14: **return DoNotKnow**
15: **end if**
16: **return** $\gcd(t - r, n)$

Theorem 5.9. *Algorithm 9 is correct and, given an oracle for* SqrtMod**, runs in zero-error probabilistic polynomial time in the bitsize of the input.*

Proof. Steps 1 to 6 ensure that the requirements of Lemma 3.38 are met. If $2 \mid n$, then $p = 2$ or $q = 2$ and the algorithm returns the right value. If $\sqrt{n} \in \mathbb{Z}$, we know that $p = q$, and finding $\sqrt{n} = \sqrt{p^2} = p$ can be done efficiently by Algorithm 3.

Lemma 3.38 now implies that $r \mathrm{R} n$ if and only if $r \mathrm{R} p$ and $r \mathrm{R} q$. This is important since $r \mathrm{R} p$ means that r is a root of $\left(X^2 - p\right)$ implying that $-r$ is a root, too, yielding that $(-r) \mathrm{R} p$. The same holds for q. So s has four different roots. Say the oracle for $\sqrt{\cdot} \pmod{n}$ chooses one of them, then since r was chosen at random, half of them are not $\pm r$. So Step 14 is only reached in less than half of the cases. Otherwise define $g := \gcd(t - r, n)$. If $g = n$ we would have that n is a factor of $(t - r)$ which directly implies $t - r \equiv 0 \pmod{n}$. But this is not allowed by Step 13. Additionally, we know that

$$t^2 \equiv r^2 \pmod{n}$$

or, in other words,

$$\exists\, k \in \mathbb{Z}\colon kpq = kn = t^2 - r^2 = (t - r)(t + r)$$

If n is a factor of either $(t - r)$ or $(t + r)$ we would have that $t \equiv \pm r \pmod{n}$, which again is not possible by Step 13. This means that exactly one of p and q divides $(t - r)$ implying that $g \in \{p, q\}$. □

We summarize these results in the following theorem.

Theorem 5.10. *We have:*

$$\text{IntFact} \in \mathbf{ZPP}^{\text{SqrtMod}^*}.$$

Chapter 6

Outlook and Open Questions

Abstract

In this chapter, we will give an outlook over possible development in the future and state some open questions that might be of interest.

6.1 Hilbert-Symbol

In this work, we established the following lower and upper bounds for the Hilbert symbol:

$$\textsc{QuadResidue}^* \leq \textsc{HilbertSymbol}_{\mathbb{Q}} \leq \textsc{IntFact}$$

where $\textsc{QuadResidue}^*$ is a special case of the problem to decide whether a given number is a quadratic residue modulo a composite number n or not. The differences between $\textsc{QuadResidue}^*$ and QuadResidue seem rather technical. So a natrual question is whether one can weaken the restrictions to obtain a better lower bound.

IntFact is believed to be a hard problem. Since $\textsc{HilbertSymbol}_{\mathbb{Q}}$ helps to decide $\textsc{QuadFormEquiv}_{\mathbb{Q}}$ and its functional version $\textsc{FuncQuadFormEquiv}_{\mathbb{Q}}$ can decide IntFact (in randomized polynomial time) the upper bound IntFact for $\textsc{HilbertSymbol}_{\mathbb{Q}}$ means that at least one of the following statements is the case

- $\textsc{FuncQuadFormEquiv}_{\mathbb{Q}}$ is much harder than $\textsc{QuadFormEquiv}_{\mathbb{Q}}$.
- The randomization to reduce SqrtMod to IntFact helps a lot.
- IntFact is not as hard as we think it is. At the time of this work, the fastest sub-exponential algorithm — the general number fields sieve (GNFS) — runs in

$$\mathcal{O}\left(\exp\left(\left(\frac{64}{9}\log\left(\log\left(n\right)\right)\right)^{\frac{1}{3}}\left(\log\left(\log\left(n\right)\right)\right)^{\frac{2}{3}}\right)\right)$$

steps.

6.2 Connected complexity classes

In this chapter we want to present some results about the complexity of the general polynomial equivalence problem and its special cases. The algebra isomorphism problem defined in Section 2.4 plays a major role, so we will first give upper bounds for some interesting fields.

Theorem 6.1.

(i). $\textsc{CommAlgIso}_{\mathbb{F}_q} \in \mathbf{NP} \cap \mathbf{coAM}$ *for a prime power* q.

(ii). $\textsc{CommAlgIso}_{\mathbb{R}} \in \mathbf{EEXP}$.

(iii). $\textsc{CommAlgIso}_{\mathbb{F}} \in \mathbf{PSPACE}$ *if* $\mathbb{F} = \overline{\mathbb{F}}$.

Proof. A proof can be found in

(i). [KS05, Theorem 3.1.].

(ii). [DH88].

(iii). [Bro06]. □

The complexity of $\textsc{CommAlgIso}_{\mathbb{Q}}$ is unknown. It is not even known if the problem is decidable at all.

Theorem 6.2. *For every field* $\mathbb{F}$ *one has:*

(i). $\textsc{GraphIso} \leq_T^p \textsc{CommAlgIso}_{\mathbb{F}}$.

(ii). $\textsc{GraphIso} \leq_T^p \textsc{CubicFormEquiv}_{\mathbb{F}}$.

Proof. A proof can be found in

(i). [KS05, Theorem 3.2.] or [AS05, Theorem 2] or [AS06b, Lemma 6.13].

(ii). [AS05, Theorem 4]. □

Theorem 6.3.

(i). $\textsc{PolyEquiv}_{d,\mathbb{F}_q} \in \mathbf{NP} \cap \mathbf{coAM}$ *for a prime power* q.

(ii). $\textsc{PolyEquiv}_{d,\mathbb{R}} \in \mathbf{EEXP}$.

(iii). $\textsc{PolyEquiv}_{d,\mathbb{F}} \in \mathbf{PSPACE}$ *if* $\mathbb{F} = \overline{\mathbb{F}}$.

Proof. The proof is given in [AS06b, Theorem 2.1]. □

Theorem 6.4.

(i). $\textsc{CommAlgIso}_{\mathbb{F}} \leq_T^p \textsc{CubicFormEquiv}_{\mathbb{F}}$.

(ii). $\textsc{CommAlgIso}_{\mathbb{F}} \leq_T^p \textsc{CubicPolyEquiv}_{\mathbb{F}}$.

(iii). $\textsc{FormEquiv}_{d,\mathbb{F}} \leq_T^p \textsc{CommAlgIso}_{\mathbb{F}}$ *(if* $\mathbb{F}$ *contains* d*-th roots)* .

Proof. A proof can be found in

(i). [AS06a, Theorem 4.1] or [AS06b, Theorem 3.10].

(ii). [AS06b, Theorem 2.7].

(iii). [AS06b, Theorem 2.3]. □

The preceding theorems have some interesting consequences:

- A subexponential algorithm for CubicFormEquiv$_{\mathbb{F}}$, where $\mathbb{F}$ is an arbitrary field, will result in a subexponential algorithm for GraphIso. Can for example the theory of cubic forms over $\mathbb{C}$ help find a subexponential algorithm for GraphIso?
- The decidability of CubicFormEquiv$_{\mathbb{Q}}$ will imply the decidability of CommAlgIso$_{\mathbb{Q}}$. Due to the rich structure of cubic forms, this might be easier to show.

6.3 Cubic form equivalence

Quadratic forms are well understood due to works of Minkowski, Hasse and Witt — some of the results can be found in Chapter 4. But the "slightly" more general case of cubic forms turns out to seem much more complicated. It may even be the hardest case: Theorem 6.3 shows that for a field $\mathbb{F}$ that contains d-th roots, we have that

$$\text{FormEquiv}_{d,\mathbb{F}} \leq \text{CommAlgIso}_{\mathbb{F}} \leq \text{CubicFormEquiv}_{\mathbb{F}}.$$

So in this case, solving CubicFormEquiv$_{\mathbb{F}}$ enables us to solve the general FormEquiv$_{d,\mathbb{F}}$ problem for any d.

But cubic forms have a lot of structure that may enable us to gain insight into the isomorphism problems of commutative $\mathbb{F}$-algebras and graphs. [AS06b] states some questions that may be of great interest in this context:

- Are there invariants for cubic forms like the ones for quadratic forms? The invariants for quadratic forms are defined in Chapter 4.
- Is there something like a limited Local-Global-Principle / Hasse-Minkowski Theorem (Theorem 4.78) for cubic forms? In general it is false, since the form

  $$3X^3 + 4Y^3 + 5Z^3$$

  does not represent zero over $\mathbb{Q}$ but it does over every completion $\mathbb{Q}_v$ for $v \in V$. A proof can for example be found in [Con].
- Even if this is not possible: Is there an algorithm that decides rational cubic form equivalence? The problem QuadFormEquiv$_{\mathbb{Q}}$ basically reduces to the question of finding $\mathbb{Q}$-roots of diagonal equations — see for example Algorithm 7. Chapter 4 shows that two quadratic forms that are equivalent over $\mathbb{R}$ and that represent the same set of points over $\mathbb{Q}$ are also equivalent over $\mathbb{Q}$. Such a result does not hold for cubic forms, see for example [AS06b, Section 6. Discussion]. So finding $\mathbb{Q}$-roots for cubic forms may not be related to the equivalence problem.

- Can one generalize the reduction $\textsc{CommAlgIso}_{\mathbb{F}} \leq \textsc{CubicFormEquiv}_{\mathbb{F}}$ to general fields, not only the ones that contain d-th roots? Or even further to the case where $\mathbb{F}$ is allowed to be a ring?

Bibliography

[AM] Leonard M. Adleman and Kevin S. Mccurley. Open problems in number theoretic complexity, ii.

[AM69] M.F. Atiyah and I.I.G. Macdonald. *Introduction to Commutative Algebra (on Demand)*. Addison-Wesley series in mathematics. Addison-Wesley Publishing Company, 1969.

[AS05] Manindra Agrawal and Nitin Saxena. Automorphisms of finite rings and applications to complexity of problems. *STACS'05, Springer LNCS 3404*, pages 1–17, 2005.

[AS06a] Manindra Agrawal and Nitin Saxena. Equivalence of $\mathbb{F}$-algebras and cubic forms. *STACS'06, Springer LNCS*, pages 115–126, 2006.

[AS06b] Manindra Agrawal and Nitin Saxena. On the complexity of cubic forms. to be submitted, 2006.

[Bro06] W. D. Brownawell. Bounds for the degrees in the nullstellensatz. *Annals of Maths 126*, pages 577–591, 2006.

[BZ10] Richard P. Brent and Paul Zimmermann. An o(m(n) log n) algorithm for the jacobi symbol. *CoRR*, abs/1004.2091, 2010.

[Con] Keith Conrad. Selmer's example.

[CP05] R. Crandall and C.B. Pomerance. *Prime Numbers: A Computational Perspective*. Lecture notes in statistics. Springer Science+Business Media, Incorporated, 2005.

[DH88] J. Davenport and J. Heintz. Real quantifier elimination is doubly exponential. *Journal of Symbolic Computation, 5*, pages 29–35, 1988.

[Har08] RupertJ. Hartung. Cryptography based on quadratic forms: Complexity considerations. In Stefan Lucks, Ahmad-Reza Sadeghi, and Christopher Wolf, editors, *Research in Cryptology*, volume 4945 of *Lecture Notes in Computer Science*, pages 52–64. Springer Berlin Heidelberg, 2008.

[JS08] Tibor Jager and Jörg Schwenk. The generic hardness of subset membership problems under the factoring assumption. Cryptology ePrint Archive, Report 2008/482, 2008. `http://eprint.iacr.org/`.

[Kob77] Neal Koblitz. *P-adic numbers, p-adic analysis, and zeta-functions / Neal Koblitz*. Springer-Verlag, New York :, 1977.

[KS05] Neeraj Kayal and Nitin Saxena. On the ring isomorphism and automorphism problems. *Proceedings of the 20th Annual IEEE Conference on Computational Complexity*, pages 2–12, 2005.

[Ser73] Jean-Pierre Serre. *A Course in Arithmetic*. Springer, 1973.

[Sim05] Denis Simon. Solving quadratic equations using reduced unimodular quadratic forms. *Math. of Comp*, 74:1531–1543, 2005.

Index

Nomenclature

assert	in algorithms: terminates with false if argument is false, page 17
$\textsc{CommAlgIso}_{\mathbb{F}}$	commutative algebra isomorphism decision problem, page 20
coNP	complement of **NP**, page 16
$\textsc{CubicFormEquiv}_{\mathbb{F}}$	problem of cubic form equivalence, page 19
$\textsc{CubicPolyEquiv}_{\mathbb{F}}$	problem of cubic polynomial equivalence, page 19
deg	degree of a polynomial, page 14
disc_v	disc relative to v, page 96
disc	discriminant of a quadratic vector space, page 71
$a \mid b$	a divides b, page 15
U^*	the dual of a vector space U, page 73
ε	Hasse-Minkowski invariant for quadratic forms, page 87
ε_v	Hasse-Minkowski invariant relative to v, page 96
$f \sim g$	f and g are equivalent, page 17
EXP	problems decidable in expoential time, page 16
$\mathbb{F}$	field, page 13
$\textsc{QuadPolyEquiv}_{\mathbb{F}}$	problem of form equivalence, page 19
$\mathbb{F}_q$	field with q elements, page 13
FuncIntFact	functional integer factoring problem, page 21
$\mathbb{F}[X_1, \ldots, X_n]$	polynomial ring in n variables, page 13
$\mathbb{F}[X_1, \ldots, X_n]^{=d}$	forms of degree d, page 18
$\mathbb{F}[X_1, \ldots, X_n]^{\leq d}$	forms of degree at most d, page 18
$[\mathbb{F}^*]^n$	$\{ x^n \mid x \in \mathbb{F}^* \}$, page 14
$\mathcal{H}_{a,b}$	set of places $v \in V$ where $(a,b)_v \neq 1$, page 52

Algorithms