Quadratic Mathematics

Prof. J.H. Coates

Michælmas Term 1996

These notes are maintained by Paul Metcalfe.
Comments and corrections to pdm23@cam.ac.uk.

```
Revision: 2.9
Date: 2004/07/26 07:35:41
```

The following people have maintained these notes.

– date Paul Metcalfe

Contents

Introduction

These notes are based on the course "Quadratic Mathematics" which was lectured by Prof. J. H. Coates in Cambridge in the Michælmas Term 1997. These typeset notes are totally unconnected with Prof. Coates.

Other sets of notes are available for different courses. At the time of typing these courses were:

Probability	Discrete Mathematics
Analysis	Further Analysis
Methods	Quantum Mechanics
Fluid Dynamics 1	Quadratic Mathematics
Geometry	Dynamics of D.E.'s
Foundations of QM	Electrodynamics
Methods of Math. Phys	Fluid Dynamics 2
Waves (etc.)	Statistical Physics
General Relativity	Dynamical Systems
Combinatorics	Bifurcations in Nonlinear Convection

They may be downloaded from

`http://www.istari.ucam.org/maths/`.

Chapter 1

Introduction to Bilinear Forms

This course is divided into two parts. The first part is about 2/3 of the course, and covers quadratic phenomena using the tools of linear algebra. The orthogonal, unitary and symplectic groups are introduced. The second part of the course looks at quadratic phenomena in number theory.

1.1 Definition of a field

A field $\mathbb{K}$ is a set with two binary operations written $+$ and $*$ satisfying these axioms :-

1. $\mathbb{K}$ is an Abelian group under $+$. The zero element of this group is written 0.
2. $\mathbb{K} \setminus \{0\}$ is Abelian under $*$. The identity element is written 1.
3. $a * (b + c) = a * b + a * c$ for all a, b, $c \in \mathbb{K}$.

Example. *Examples of fields include $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{Z}/p\mathbb{Z}$ (where p is prime).*

1.2 The characteristic of a field

Let $\mathbb{K}$ be a field. Then :-

Definition 1.1. *For $n \in \mathbb{Z}$, let*

$$n * 1 = \begin{cases} \overbrace{1 + \cdots + 1}^{n \text{ times}} & \text{if } n > 0; \\ -\overbrace{(1 + \cdots + 1)}^{|n| \text{ times}} & \text{if } n < 0; \\ 0 & \text{if } n = 0. \end{cases}$$

Definition 1.2. *We say that $\mathbb{K}$ has characteristic 0 if $n*1 = 0$ implies $n = 0$. Otherwise we say that $\mathbb{K}$ has characteristic n if n is the least (strictly) positive n such that $n * 1 = 0$.*

Example. *$\mathbb{Z}/p\mathbb{Z}$ has characteristic p.*

1.3 Some definitions

Definition 1.3. *A bilinear form ψ is a map $U \times V \mapsto \mathbb{K}$ satisfying :-*

1. *If $y = y_0$ is fixed, then $x \mapsto \psi(x, y_0)$ is linear in x.*
2. *If $x = x_0$ is fixed, then $y \mapsto \psi(x_0, y)$ is linear in y.*

Example. *If $U = V = \mathbb{R}^N$, $\psi(X, Y) = \sum_{i=1}^{N} x_i y_i$ is bilinear.*
If $V = C[a, b]$, take $\psi : V \times V \mapsto \mathbb{R}$ as $\psi(f, g) = \int_a^b f(x)g(x)dx$.

If U, V are finite-dimensional then a bilinear form has an attached matrix . Fix bases $\{d_1, \ldots, d_m\}$, $\{e_1, \ldots, e_n\}$ of U and V respectively.

Definition 1.4. *The matrix of ψ relative to the bases $\{d_1, \ldots, d_m\}$, $\{e_1, \ldots, e_n\}$ is the $m \times n$ matrix $A = (\psi(d_i, e_j))$.*

So if $x \in U$, $y \in V$ then $\psi(x, y) = x^T A y$.

1.4 Change Of Basis

Suppose U, V are finite dimensional vector spaces over a field $\mathbb{K}$. Then given a bilinear form $\psi : U \times V \mapsto \mathbb{K}$, and bases $\{d_1, \ldots, d_m\}$, $\{e_1, \ldots, e_n\}$ of U and V respectively there is an associated matrix $A = (\psi(d_i, e_j))$. If we take other bases $\{d'_1, \ldots, d'_m\}$, $\{e'_1, \ldots, e'_n\}$, then the matrix of ψ with respect to this basis is $A' = \big(\psi(d'_i, e'_j)\big)$.

Lemma 1.5. *There exist invertible matrices $\mathcal{M}$ ($m \times m$) and $\mathcal{N}$ ($n \times n$) such that :-*

$$A' = \mathcal{M}^T A \mathcal{N}$$

Proof. Since the d_j's and e_i's form bases for their respective vector spaces

$$d'_h = \sum_{i=1}^{m} \mathcal{M}_{ih} d_i \text{ and } e'_l = \sum_{j=1}^{n} \mathcal{N}_{jl} e_j$$

Now,

$$\begin{aligned} a'_{hl} = \psi(d'_h, e'_l) &= \psi\left(\sum_{i=1}^{m} \mathcal{M}_{ih} d_i, \sum_{j=1}^{n} \mathcal{N}_{jl} e_j\right) \\ &= \sum_{i=1}^{m} \mathcal{M}_{ih} \psi\left(d_i, \sum_{j=1}^{n} \mathcal{N}_{jl} e_j\right) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} \mathcal{M}_{ih} \mathcal{N}_{jl} a_{ij} \\ &= \left(\mathcal{M}^T A \mathcal{N}\right)_{hl} \end{aligned}$$

Both $\mathcal{M}$ and $\mathcal{N}$ are clearly invertible, since the d'_j's and e'_i's form bases for their respective vector spaces. □

Corollary 1.6. $A' = \mathcal{R}^{-1} A \mathcal{N}$

Definition 1.7. *The rank of ψ is the rank of its associated matrix for any choice of bases. This is well defined due to* (1.6) *and linear algebra results.*

Corollary 1.8. *Given ψ we can always find bases of U, V such that the matrix of ψ is of the form :-*

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

where $r = rank(\psi)$

If $U = V$, we choose the two new bases to be the same. We get :-

$$\begin{aligned} A &= (\psi(e_i, e_j)) \\ A' &= \left(\psi(e'_i, e'_j)\right) \\ &= \mathcal{M}^T A \mathcal{M} \end{aligned}$$

Definition 1.9. *If A and A' are related by $A' = \mathcal{M}^T A \mathcal{M}$ then A and A' are said to be congruent.*

1.5 Relation between bilinear forms and dual space

Definition 1.10. *Given a vector space V over a field $\mathbb{K}$, the* dual space V^* *is defined by*

$$V^* = \{\alpha : V \mapsto \mathbb{K}, \alpha \textit{ linear}\}.$$

Some definitions. For all of these, take $\psi : U \times V \mapsto \mathbb{K}$ to be bilinear.

Definition 1.11. *Suppose $A \subseteq U$. Then define*

$$A_R^\perp = \{v \in V : \psi(u, v) = 0, \forall u \in A\}$$

Now take $B \subseteq V$ and define

$$B_L^\perp = \{u \in U : \psi(u, v) = 0, \forall v \in B\}$$

Definition 1.12. *If we take $A = U$ or $B = V$, we get*

$$\begin{aligned} U_R^\perp &= \textit{ the } \text{right kernel } \textit{of } \psi \\ &= \{v \in V : \psi(u, v) = 0, \forall u \in U\} \end{aligned}$$

and

$$\begin{aligned} V_L^\perp &= \textit{ the } \text{left kernel } \textit{of } \psi \\ &= \{u \in U : \psi(u, v) = 0, \forall v \in V\} \end{aligned}$$

Given ψ, we get two canonical linear maps :-

$$\widehat{\psi}_L : U \mapsto V^*, \widehat{\psi}_L(u) = (v \mapsto \psi(u, v))$$

and

$$\widehat{\psi}_R : V \mapsto U^*, \widehat{\psi}_R(v) = (u \mapsto \psi(u, v))$$

Lemma 1.13. $\ker \widehat{\psi}_L = V_L^\perp$ *and* $\ker \widehat{\psi}_R = U_R^\perp$.

Proof. If $u \in \ker \widehat{\psi}_L$ then $\widehat{\psi}_L(u) = 0$ and so $u \in V_L^\perp$. Same for other three cases. □

Example. *Let* $U = \mathbb{R}^2$ *and* $V = \mathbb{R}^3$, $\psi : U \times V \mapsto \mathbb{R}$, $\psi(x, y) = x_1 y_2$.

The left kernel is $\left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle$ *and the right kernel is* $\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$.

Definition 1.14. *If* $V_L^\perp = \{0\}$ *and* $U_R^\perp = \{0\}$ *we say that* ψ *is* non-degenerate.

Theorem 1.15. *Assume* U, V *are finite dimensional vector spaces over* $\mathbb{K}$ *and that* ψ *is non-degenerate. Then*

1. $\dim U = \dim V$.
2. $\widehat{\psi}_L$ *is an isomorphism.*
3. $\widehat{\psi}_R$ *is an isomorphism.*

Proof. From Linear Maths, $\dim U = \dim U^*$ and $\dim V = \dim V^*$. Since ψ is non-degenerate, both $\widehat{\psi}_L$ and $\widehat{\psi}_R$ are injective. Now, $\widehat{\psi}_L$ injective implies

$$\begin{aligned} \dim U &= \dim \widehat{\psi}_L(U) \\ &\leq \dim V^* = \dim V \end{aligned}$$

And $\widehat{\psi}_R$ injective implies

$$\begin{aligned} \dim V &= \dim \widehat{\psi}_L(V) \\ &\leq \dim U^* = \dim U \end{aligned}$$

Therefore $\dim U = \dim V$. Now $\dim \widehat{\psi}_R(V) = \dim U$ and hence $\widehat{\psi}_R(V) = U$. □

Theorem 1.16. *Assume* $\dim U = \dim V < \infty$. *Then the following assertions about* ψ *are equivalent.*

1. ψ *is non-degenerate.*
2. *The left kernel of* ψ *is* $\{0\}$.
3. *The right kernel of* ψ *is* $\{0\}$.
4. *The matrix* A *representing* ψ *is non-singular relative to any bases of* U, V.

A lemma would be helpful. First of all, some notation.

Notation. *Let $\{d_1, \ldots, d_m\}$ be a basis for U and $\{e_1, \ldots, e_n\}$ be a basis for V. Then the dual bases are $\{d_1^*, \ldots, d_m^*\}$, $\{e_1^*, \ldots, e_n^*\}$ for U^* and V^* respectively, where d_i^* is defined by $d_i^*(d_j) = \delta_{ij}$ and $e_i^*(e_j) = \delta_{ij}$.*

Lemma 1.17. *The matrix of $\widehat{\psi}_R : V \mapsto U^*$ is $A = (\psi(d_i, e_j))$ relative to the bases $\{e_1, \ldots, e_n\}$, $\{d_1^*, \ldots, d_m^*\}$. The matrix of $\widehat{\psi}_L : U \mapsto V^*$ is A^T relative to the bases $\{e_1^*, \ldots, e_n^*\}$, $\{d_1, \ldots, d_m\}$.*

Proof. I'll only prove for $\widehat{\psi}_L$. Let R be the matrix for $\widehat{\psi}_L$.

$$\begin{aligned}\widehat{\psi}_L(d_j)(e_h) &= \psi(d_j, e_h) \\ &= a_{jh}\end{aligned}$$

$$\begin{aligned}\widehat{\psi}_L(d_j)(e_h) &= \sum_{i=1}^{n} r_{ij} e_i^*(e_h) \\ &= \sum_{i=1}^{n} r_{ij}\delta_{ih} \\ &= r_{hj}\end{aligned}$$

So $a_{jh} = r_{hj}$ giving $R = A^T$. □

Corollary 1.18. *Assume $\dim U = \dim V < \infty$. Then $\widehat{\psi}_R$ is an isomorphism if and only if $\widehat{\psi}_L$ is an isomorphism.*

Proof of Theorem 1.16. Immediate from Lemma 1.17. □

1.6 The adjoint map

Definition 1.19. *Given V a finite-dimensional vector space over $\mathbb{K}$, $\psi : V \times V \mapsto \mathbb{K}$ a non-degerate bilinear form and $\alpha : V \mapsto V$, a linear map, we define the* adjoint map *β of α with respect to ψ by*

$$\psi(\alpha(x), y) = \psi(x, \beta(y)) \forall x, y \in V$$

β is written as α_ψ^*.

Theorem 1.20. *Such a β always exists, and is unique.*

Proof. First prove uniqueness.

$$\begin{aligned}\psi(\alpha(x), y) = \psi(x, \beta_1(y)) = \psi(x, \beta_2(y)) \\ \Rightarrow \psi(x, (\beta_1 - \beta_2)(y)) = 0 \\ \Rightarrow \beta_1(y) - \beta_2(y) \in V_R^\perp \\ \Rightarrow \beta_1(y) = \beta_2(y) \\ \Rightarrow \beta_1 = \beta_2\end{aligned}$$

And for existence, look at the map $\phi : x \mapsto \psi(x, z)$. Firstly, $\phi \in V^*$. Now, I claim that every element of V^* is of form $x \mapsto \psi(x, z)$ for some z. Proof, either be subtle or blat it out in co-ordinates. Then pick β such that $\beta(y) = z$. Now $\beta : V \mapsto V$, and it is easy to see that β is linear. $\square$

Chapter 2

Special Bilinear Forms

In this section, we look at bilinear forms with some sort of additional structure.

2.1 Symmetric Bilinear Forms

Notation. *In this subsection, V is a vector space over $\mathbb{K}$ and $\psi : V \times V \mapsto \mathbb{K}$ is always bilinear.*

Definition 2.1. *ψ is symmetric if and only if $\psi(x, y) = \psi(y, x) \forall x, y \in V$.*

If V is finite dimensional, it is clear that the matrix A representing ψ is symmetric, i.e. $A = A^T$.

Definition 2.2. *A quadratic form on V is a function $q : V \mapsto \mathbb{K}$ of the form $q(x) = \psi(x, x)$, where ψ is symmetric.*

Lemma 2.3. *If $1 + 1 \neq 0$ in $\mathbb{K}$, then $\psi(x, y)$ is determined by $q(x)$. Specifically,*

$$\psi(x, y) = \frac{q(x + y) - q(x) - q(y)}{2}.$$

Proof. Expand it. □

Now, an important theorem.

Theorem 2.4. *If the characteristic of $\mathbb{K}$ is not 2 and V is finite dimensional, then there exists a basis $\{v_1, \ldots, v_n\}$ such that $\psi(v_i, v_j) = 0$ if $i \neq j$.*

Proof. This is proved by induction on $n = \dim V$. It is true if $n = 1$ without too much effort. So assume true for all V' and $\psi' : V' \times V' \mapsto \mathbb{K}$, $\dim V' < \dim V$. Next, assume that ψ is not equivalently 0, since otherwise the result is trivial. So $\exists x,y$ such that $\psi(x, y) \neq 0 \Rightarrow \exists x_1$ such that $q(x_1) \neq 0$.

Let $V_1 = \{x \in V : \psi(x, x_1) = 0\}$. V_1 is clearly a subspace of V, and $V_1 \neq V$ (as $x_1 \notin V$). Define $\psi_1 : V_1 \times V_1 \mapsto \mathbb{K}$ by $\psi_1(x, y) = \psi(x, y)$. Now by the inductive hypothesis there exists a basis $\{e_1, \ldots, e_r\}$ of V_1 such that $\psi_1(e_i, e_j) = 0$ if $i \neq j$.

Now, must prove that $\{x_1, e_1, \ldots, e_r\}$ is a basis of V, as this gives the result immediately. Since $\{x_1, e_1, \ldots, e_r\}$ has at most n elements, it suffices to show that it spans V. Now take $y \in V$ and let $y' = y - \frac{\psi(y,x_1)}{\psi(x_1,x_1)} x_1$. Then $\psi(y', x_1) = 0$, so $y' = \sum_{i=2}^{r} a_i e_i$. □

Corollary 2.5. *A matrix interpretation of the theorem. If the characteristic of $\mathbb{K}$ is not 2, then for any symmetric matrix A, $\exists$ invertible N such that $N^T AN$ is diagonal.*

Proof. Obvious from theorem □

Corollary 2.6. *Let $\mathbb{K} = \mathbb{C}$, V be a finite dimensional vector space over $\mathbb{C}$ and ψ be a symmetric bilinear form $V \times V \mapsto \mathbb{C}$. Then $\exists$ a basis $\{e_1, \ldots, e_n\}$ of V such that if $x = x_1e_1 + \cdots + x_ne_n$ then $\psi(x, x) = x_1^2 + \cdots + x_r^2$, where r is the rank of ψ.*

Proof. Immediate from theorem. □

Corollary 2.7. *If A_1 and A_2 are two complex symmetric matrices, then they are congruent if and only if rank A_1 = rank A_2.*

Proof. Immediate from matrix interpretation. □

2.2 Real Quadratic Forms

Theorem 2.8 (Sylvester's Law Of Inertia).
Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $q\colon V \mapsto \mathbb{R}$ be any quadratic form. Then there exists a basis $\{e_1, \ldots e_n\}$ of V such that if $x = x_1e_1 + \cdots + x_ne_n$, $q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2$, where r is the rank of q, defined as the rank of ψ. Moreover, p is the same for all such bases.

Definition 2.9. *$2p - r = p - (r - p)$ is called the signature of q.*

Corollary 2.10 (Matrix interpretation). *Let $\mathcal{A}$ be any real symmetric matrix. Then there exists an invertible $\mathcal{N}$ such that*

$$\mathcal{N}^t\mathcal{A}\mathcal{N} = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Corollary 2.11. *Let $\mathcal{A}_1$,$\mathcal{A}_2$ be real symmetric matrices. Then $\mathcal{A}_1$ and $\mathcal{A}_2$ are congruent iff they have the same rank and signature.*

Definition 2.12. *We say that a quadratic form q is positive definite in a subspace W of V if $q(x) > 0 \forall x \neq 0 \in W$.*

Lemma 2.13. *p is the maximal dimension of any subspace of V on which q is positive definite.*

Proof. Let W be any subspace of V on which q is positive definite. Define $R = \langle e_{p+1}, \ldots e_n\rangle$. Now $q(x) \leq 0 \forall x \in R$, and so $R \cap W = \{0\}$. Now

$$\dim W + R = \dim W + \dim R \leq \dim V, \text{ so } \dim W \leq p.$$

□

Proof of Sylvester's Law of Inertia. Firstly, prove existence of basis. General result implies there exists a basis $\{v_1, \ldots v_n\}$ of V such that $\psi(v_i, v_j) = 0$ if $i \neq j$. Order basis such that $\psi(v_i, v_i) > 0$ for $i = 1, \ldots p$, $\psi(v_i, v_i) < 0$ for $i = p + 1, \ldots r$. Now, we can find $c_i \in \mathbb{R}$ st $c_i^2 = |\psi(v_i, v_i)|$. Define $e_i = \frac{v_i}{c_i}$ for $i = 1, \ldots r$, $e_i = v_i$ otherwise.

The uniqueness of p follows from Lemma 2.13. □

Theorem 2.14. *Let $\mathcal{A}$ be a real symmetric matrix. Then there exists a matrix $\mathcal{N}$ such that :-*

1. *$\mathcal{N}^t\mathcal{A}\mathcal{N}$ is diagonal.*
2. *$\mathcal{N}^t\mathcal{N} = I$.*

Proof. Proof later. □

2.3 Orthogonal Groups

Let V be a finite dimensional vector space over $\mathbb{K}$, with the characteristic of $\mathbb{K}$ not 2. Given $\psi : V \times V \mapsto \mathbb{K}$ bilinear, symmetric, non-degenerate.

Definition 2.15.

$$\mathcal{O}(V,\psi) = \{\alpha \colon V \mapsto V : (\alpha(x) = 0 \Leftrightarrow x = 0), q(x) = q(\alpha(x))\ \forall x \in V\}$$

is the orthogonal group of ψ, with the obvious group laws.

Definition 2.16 (Orthogonal direct sum). *If U, W are subspaces of V, then $V = U \perp W$ iff $V = U \oplus W$ and $\psi(u,w) = 0\ \forall (u,w) \in U \times W$.*

Definition 2.17. *Let $V = U \perp W$. A reflexion with respect to (U,W) is a map $r : V \mapsto V$ such that $r(u+w) = u - w\ \forall (u,w) \in U \times W$.*

Lemma 2.18. *Let $r \in \mathcal{O}(V,\psi)$ such that $r^2 = \iota$. Then r is a reflexion wrt subspace U, W of V with $V = U \perp W$.*

Proof. Define U, W as

$$U = \{x \in V : r(x) = x\}, W = \{x \in V : r(x) = -x\}.$$

These work! □

Theorem 2.19 (Main Theorem). *Every element of $\mathcal{O}(V,\psi)$ can be written as a product of n reflexions, where $n = \dim V$.*

Lemma 2.20. *Let x, y be any elements of V with $\psi(x,x) = \psi(y,y) \neq 0$. Then there exists a reflexion $r \in \mathcal{O}(V,\psi)$ with $y = r(x)$.*

Proof. Define $u = \frac{x+y}{2}$ and $v = \frac{x-y}{2}$. Firstly, $\psi(u,v) = 0$. Secondly $\psi(u,u) + \psi(v,v) = \psi(x,x) \neq 0$ and so one of $\psi(u,u)$ and $\psi(v,v)$ is non-zero, say $\psi(u,u) \neq 0$.

Define $U = \{\lambda u : \lambda \in \mathbb{K}\}$ and $W = \{w \in V : \psi(u,w) = 0\}$. Claim : $V = U \oplus W$. $U \cap W = \{0\}$ trivially, and given $v \in V$, define $v_1 = b - \frac{\psi(v,u)}{\psi(u,u)}u$. Now $\psi(u,v_1) = 0$ and so $V = U \perp W$. Let r be the reflexion wrt (U,W). $r(x) = r(u+v) = y$. □

Proof of theorem. Induction on $\dim V = n$. Trivial when $n = 1$. Now assume $n > 1$ and the theorem is true for all V', ψ' st $\dim V < n$ and $\psi' : V' \times V' \mapsto \mathbb{K}$ non-degenerate, symmetric and bilinear. In V, and given $\alpha \in \mathcal{O}(V,\psi)$, choose a basis $\{e_1, \ldots e_n\}$ of V such that $\psi(e_i, e_j) = 0$ if $i \neq j$. Note that $\psi(e_i, e_i) \neq 0\ \forall i$ since ψ is non-degenerate.

Define $U = \langle e_1 \rangle$ and $W = \langle e_2, \dots e_n \rangle$. Note that $V = U \perp W$ and $W = U^{\perp}$. Define $\psi' : W \times W \mapsto \mathbb{K}$ by $\psi'(x, y) = \psi(x, y)\ \forall x, y \in W$. ψ' is a non-degenerate, symmetric bilinear form.

Return to α. By lemma, there exists a reflexion $r_1 \in \mathcal{O}(V, \psi)$ with $r(\alpha(e_1)) = e_1$. Now consider $\beta = r_1\alpha \in \mathcal{O}(V, \psi)$. By construction $\beta(e_1) = e_1$. Given $w \in W$, $\psi(\beta(w), e_1) = \psi(\beta(w), \beta(e_1)) = \psi(w, e_1) = 0$ and so $\beta(W) \subseteq W$. Let β' be the restriction of β to W. Now $\beta' \in \mathcal{O}(W, \psi')$ and so $\beta' = s_2 \dots s_n$, where the s_i are reflexions. Extend s_i to V by $s_i(e_1) = e_1$, let r_i be this extension. So $\alpha = r_1 \dots r_n$. $\square$

Chapter 3

Hermitian Forms

3.1 Introduction

Let V be a vector space over $\mathbb{C}$.

Definition 3.1. *A Hermitian form on V is a function $\psi : V \times V \mapsto \mathbb{C}$ such that :-*

1. *$\psi(x, y)$ is linear in x if y fixed.*
2. *$\psi(x, y) = \overline{\psi(y, x)}$.*

If $\psi : V \times V \mapsto \mathbb{C}$ is Hermitian, then define $q(x) = \psi(x, x)$. $q(x) \in \mathbb{R} \forall x \in V$. Possibly useful (?) to know

$$\psi(x, y) = \frac{q(x + y) - q(x - y) + iq(x + iy) - iq(x - iy)}{4}$$

3.2 Hermitian Matrices and Change of Basis

Given $A \in M_n(\mathbb{C})$ ...

Definition 3.2. *$A^h = \overline{A}^t$.*

Definition 3.3. *A is Hermitian if $A^h = A$.*

If V is finite dimensional, we can define the matrix of ψ relative to some basis $\{v_1, \ldots, v_n\}$ of V by $A = (\psi(v_i, v_j))$. A is Hermitian iff ψ is Hermitian.

Theorem 3.4 (Change of Basis). *Take bases $\{v_1, \ldots, v_n\}$, $\{v'_1, \ldots, v'_n\}$ of V such that $v'_j = \sum_{i=1}^{n} \mathcal{M}_{ij} v_i$, then*

$$A' = \overline{\mathcal{M}}^h A \overline{\mathcal{M}}$$

Proof. DIY! □

3.3 Sylvester's Law?

Theorem 3.5 (Analogue of Sylvester's Law of Inertia). *Assume that V is a finite dimensional vector space over $\mathbb{C}$ and that $\psi : V \times V \mapsto \mathbb{C}$ is Hermitian. Then there exists a basis $\{e_1, \ldots, e_n\}$ of V such that if $x = x_1 e_1 + \cdots + x_n e_n$, then*

$$\psi(x, x) = |x_1|^2 + \cdots + |x_p|^2 - |x_{p+1}|^2 - \cdots - |x_r|^2$$

where r is the rank of ψ and p is the same for all such bases.

Proof. See previous. □

3.4 The Unitary Group

Definition 3.6 (The Unitary Group). *Define the unitary group $\mathcal{U}(V, \psi)$ just like the orthogonal group.*

Chapter 4

Inner Product Spaces

4.1 Euclidean Space

Let V be a vector space over $\mathbb{R}$. An inner product on V is a symmetric bilinear form ψ such that $\psi(x,x) > 0$ if $x \neq 0$. We thus get the Euclidean space (V, ψ).

Definition 4.1. *1.* $||x|| = \sqrt{\psi(x,x)}$,

2. x *is orthogonal to* y *if* $\psi(x,y) = 0$.

We also get Cauchy-Schwarz (a transplantable proof will be given for unitary space) and thus the triangle inequality.

4.2 Unitary Space

Definition 4.2. *An inner product on V is a Hermitian form* ψ *is* $\psi(x,x) > 0 \; \forall x \neq 0$. *This gives rise to unitary space* (V, ψ).

Theorem 4.3 (Cauchy-Schwarz).

$$|\psi(x,y)| \leq ||x||\,||y||$$

Proof. Given $\lambda \in \mathbb{C}$,

$$\psi(x - \lambda y, x - \lambda y) \geq 0$$

$$\psi(x,x) - \lambda\psi(y,x) - \overline{\lambda\psi(y,x)} + |\lambda|^2 \psi(y,y) \geq 0$$

Assume $y \neq 0$ and put

$$\lambda = \frac{\psi(x,y)}{\psi(y,y)}.$$

This gives result. □

4.3 Orthogonal Projection

(V, ψ) is either orthogonal or unitary space. Let $W \neq V$ be a subspace of V, and let $\alpha \in V$. How do we define the "foot of the perpendicular" from α to W?

We want $\mu \in W$ such that $\psi(\alpha - \mu, w) = 0 \forall w \in W$. Or alternatively, $\alpha - \mu \in W^{\perp}$. If such an $\alpha - \mu$ exists for all α, we can write $V = W + W^{\perp}$. This is not always possible, but...

Theorem 4.4. *Assume V is finite dimensional. Then for any subspace W of V, we have $V = W \oplus W^{\perp}$.*

Proof in Euclidean case. Define $\theta : V \mapsto W^*$ as $\theta(v)(w) = \psi(w, v)$. Now

$$\begin{aligned} \dim V &= \dim(\ker \theta) + \dim(\operatorname{im} \theta) \\ &= \dim W^{\perp} + \dim(\operatorname{im}\ \theta) \end{aligned}$$

Now $\dim W^* = \dim W$, and so θ surjective gives theorem.

Since ψ is non-degenerate, every element of V^* is of the form $x \mapsto \psi(x, v)$ for some $v \in V$. So given $\phi \in W^*$, extend ϕ to $\rho : V \mapsto \mathbb{R}$. Then $\rho(x) = \psi(x, v) \forall x \in V$, the restriction of which gives ϕ. □

Now, assume V finite-dimensional over $\mathbb{C}$ or $\mathbb{R}$ and W any subspace of $V \Rightarrow V = W \oplus W^{\perp}$. Define the orthogonal projection $\Pi_W : V \mapsto W$ by $\Pi_W(w + v) = w \, \forall w \in W, v \in W^{\perp}$.

Assume $W = \langle \eta \rangle, \eta \neq 0$. Now $\Pi_W(v) = \lambda_v \eta$.

Lemma 4.5.

$$\lambda_v = \frac{\psi(v, \eta)}{\psi(\eta, \eta)}$$

Proof. $v - \lambda_v \eta \in W^{\perp}$ gives result. □

Definition 4.6. *Given a subset $S = \{e_1, e_2, \dots\} \subset V$, we say S is orthonormal iff $\psi(e_i, e_j) = \delta_{ij}$.*

4.4 Gram-Schmidt Process

Let V be an Euclidean or unitary space, with ψ the inner product.

Theorem 4.7. *Let $\{v_1, v_2, \dots\}$ be a linearly independent set in V. Then there exists an orthonormal set $\{e_1, e_2, \dots\}$ such that $\forall n \geq 1$*

$$\langle e_1, \dots, e_n \rangle = \langle v_1, \dots, v_n \rangle.$$

Proof. By induction on n. For $n = 1$, put $e_1 = \frac{v_1}{||v_1||}$.

Now assume $n > 1$ and have already constructed $\{e_1, \dots, e_{n-1}\}$ as required. Put

$$e'_n = v_n - \sum_{i=1}^{n-1} \psi(v_n, e_i) e_i.$$

Now $\langle e_1, \dots, e'_n \rangle = \langle v_1, \dots, v_n \rangle$ so $e'_n \neq 0$. Put $e_n = \frac{e'_n}{||e'_n||}$. □

4.5 Spectral Theory for $\mathbb{C}$

Let V be a finite dimensional vector space over $\mathbb{C}$. A linear map $\alpha : V \mapsto V$ is called Hermitian or self-adjoint if $\alpha = \alpha^*$ wrt a Hermitian inner product ψ.

Lemma 4.8. *The eigenvalues of α are real and $\psi(\xi_1, \xi_2) = 0$ if ξ_1 and ξ_2 are eigenvectors belonging to different eigenvalues.*

Proof.

$$\lambda\psi(\xi,\xi) = \psi(\alpha(\xi),\xi) = \overline{\lambda}\psi(\xi,\xi)$$

Now

$$\psi(\alpha(\xi_1),\xi_2) = \psi(\xi_1,\alpha(\xi_2))$$

and so

$$\lambda_1\psi(\xi_1,\xi_2) - \lambda_2\psi(\xi_1,\xi_2) = 0$$

Since $\lambda_1 \neq \lambda_2$, result follows. □

Theorem 4.9 (Self-Adjoint Case). *Let V be a finite dimensional vector space over $\mathbb{C}$ endowed with an inner product ψ. Let $\alpha : V \mapsto V$ be a linear map such that $\alpha = \alpha^*$. Then there exists an orthonormal basis of V consisting of eigenvectors of α.*

Proof. By induction on $\dim V$. Trivial when $\dim V = 1$. So assume $\dim V > 1$, and the theorem true for all subspaces of V.

α has one eigenvalue λ_1, with corresponding eigenvector $\xi_1 \neq 0$. Let $V_1 = \langle\xi_1\rangle$, and $W = V_1^{\perp}$. Now $V = V_1 \oplus W$, so $\dim W = \dim V - 1$. Let ψ_W be the restriction of ψ. Now, does α take W to W? But if $w \in W$, then

$$\psi(\alpha(w),\xi_1) = \psi(w,\alpha(\xi_1)) = \overline{\lambda_1}\psi(w,\xi_1) = 0$$

Now define $\beta : W \mapsto W$ by $\beta(w) = \alpha(w)\ \forall w \in W$. Now $\beta = \beta^*_{\psi_W}$ and so by inductive hypothesis W has an orthonormal basis $\{e_2,\ldots,e_n\}$ of eigenvectors of β. Put $e_1 = \frac{\xi_1}{||\xi_1||}$ to get $\{e_1,e_2,\ldots,e_n\}$, the desired orthonormal basis of V. □

Theorem 4.10 (Unitary Case). *Let V be a finite dimensional vector space over $\mathbb{C}$ endowed with an inner product ψ. Let $\alpha : V \mapsto V$ be a linear map such that $\alpha^* = \alpha^{-1}$. Then there exists an orthonormal basis of V consisting of eigenvectors of α.*

Proof. There exists one eigenvalue $\lambda_1 \neq 0$ with eigenvector ξ_1. Let $V_1 = \langle\xi_1\rangle$ and $W = V_1^{\perp}$. Given $w \in W$,

$$\psi(\alpha(w),\xi_1) = \psi(x,\alpha^{-1}(\xi_1) = \psi(x,\frac{\xi_1}{\lambda_1}) = 0$$

and so $\alpha(w) \in W$. Fill in the blanks. □

4.6 Spectral Theory for $\mathbb{R}$

So V is a finite dimensional vector space over $\mathbb{R}$, with $\psi : V \times V \mapsto \mathbb{R}$ an inner product. α is self-adjoint wrt ψ if $\alpha = \alpha^*$.

Lemma 4.11. *Let α be self-adjoint. Then all the eigenvalues of α are real and if ξ_1 and ξ_2 are eigenvectors belonging to distinct eigenvalues, then they are automatically orthogonal.*

Proof. For the first part, choose a basis of V, then α corresponds to a matrix $A = A^t$. The map $X \mapsto AX, \mathbb{C}^n \mapsto \mathbb{C}^n$ is Hermitian, so has real eigenvalues.

$$\psi(\alpha(\xi_1),\xi_2) = \psi(\xi_1,\alpha(\xi_2))$$

and so

$$\lambda_1\psi(\xi_1,\xi_2) - \lambda_2\psi(\xi_1,\xi_2) = 0$$

Since $\lambda_1 \neq \lambda_2$, result follows. □

Theorem 4.12 (Real, self-adjoint case). *Let V be a finite dimensional vector space over $\mathbb{R}$ endowed with an inner product ψ. Let $\alpha : V \mapsto V$ be a linear map such that $\alpha^* = \alpha$. Then there exists an orthonormal basis of V consisting of eigenvectors of α.*

Proof. As before, noting that α has real eigenvalues. □

Chapter 5

Alternating Forms

Let V be a vector space over any field $\mathbb{K}$ and $\psi : V \times V \mapsto \mathbb{K}$ bilinear.

Definition 5.1. *ψ is alternating if $\psi(x,x) = 0\ \forall x \in V$. ψ is anti-symmetric if $\psi(x,y) = -\psi(y,x)$.*

Alternating implies anti-symmetric (consider $\psi(x+y, x+y)$), and if the characteristic of $\mathbb{K}$ is not 2, anti-symmetric implies alternating.

5.1 Nice matrices

Theorem 5.2. *Assume V is finite dimensional and ψ is alternating. Then ψ has even rank $2m$ and there exists a basis $\{e_1, \ldots, e_n\}$ such that the matrix $(\psi(e_i, e_j))$ is of the form*

$$\begin{pmatrix} 0 & I_m & 0 \\ -I_m & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Proof. By induction on $\dim V$. Obvious when $\dim V = 1$, then $\psi \equiv 0$. Assume $\dim V > 1$ and result proven for all (V', ψ') where $\psi' : V' \times V' \mapsto \mathbb{K}$ is alternating and $\dim V' < \dim V$.

We want a basis $\{u_1, \ldots, u_m, v_1, \ldots, v_m, w_1, \ldots, w_s\}$ where $2m + s = n$ such that $\psi(u_i, v_i) = -\psi(v_i, u_i) = 1$ and $\psi(\text{anything else}) = 0$. If $\psi \equiv 0$, there is nothing to prove, so assume $\exists x, y$ such that $\psi(x,y) \neq 0$. Put $u_1 = \frac{x}{\psi(x,y)}$ and $v_1 = y$. Then $\psi(u_1, v_1) = 0$. Let $V_1 = \langle u_1, v_1 \rangle$. Note that $\dim V_1 = 2$. Let $W = V_1^{\perp}$. Claim that $V = V_1 \oplus W$.

Firstly note that $V_1 \cap W = \{0\}$ (by putting $\zeta = \lambda u_1 + \mu v_1 \in V_1 \cap W$). Now, given $z \in V$, define $z_1 = \psi(z, v_1)u_1 + \psi(u_1, z)v_1$ and $z - z_1 \in W$. So $V = V_1 \perp W$. Now given (W, ψ_W) apply inductive hypothesis. □

5.2 Symplectic Group

Let V be a finite dimensional vector space over $\mathbb{K}$, $\psi : V \times V \mapsto \mathbb{K}$ be bilinear, alternating and non-degenerate (implies $\dim V = 2m$).

Definition 5.3. *The symplectic group $Sp(V, \psi)$ is the set of linear maps $\alpha : V \mapsto V$ satisfying*

1. *α is an isomorphism of vector spaces*
2. *$\psi(\alpha(x), \alpha(y)) = \psi(x, y)$.*

Or alternatively, for matrix definition, choose a basis of V st ψ has a matrix

$$J_{2m} = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}$$

Then $Sp_{2m}(\mathbb{K}) = \{P \in M_{2m}(\mathbb{K}) \,|P$ is invertible and $P^t J_{2m} P = J_{2m}\}$.

Chapter 6

Number Theory

6.1 Introduction

Lemma 6.1. *Let p be any prime. Then $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is a field with p elements and $\mathbb{F}_p^\times$ the set of non-zero elements is a multiplicative group of order $p-1$.*

Proof. DIY! □

Corollary 6.2. *Let a be any integer with $(a,p)=1$. Then $a^{p-1} \equiv 1 \pmod{p}$.*

Proof. $a + p\mathbb{Z} \in \mathbb{F}_p^\times$, so $(a+p\mathbb{Z})^{p-1} = 1 + p\mathbb{Z}$. □

Definition 6.3. *Let $p > 2$ and take a with $(a,p)=1$. We say a is a quadratic residue modulo p if $a + p\mathbb{Z}$ is a square in $\mathbb{F}_p^\times$. Or equivalently, the congruence*

$$x^2 \equiv a \mod p$$

is soluble.

Lemma 6.4. *Let $a \in \mathbb{Z}$ have $(a,p)=1$, $p>2$. Then the congruence $x^2 \equiv a \pmod{p}$ has either no solutions or two solutions modulo p.*

Proof. If x_0 is a solution then $-x_0$ is a solution. $x_0 \not\equiv -x_0 \pmod{p}$ since $p \neq 2$. Now suppose x_0 and x_1 are both solns of $x^2 \equiv a \pmod{p}$. Then $x_0^2 \equiv x_1^2 \pmod{p}$, and so $p | x_0^2 - x_1^2 = (x_0 - x_1)(x_0 + x_1)$. So either $x_0 \equiv x_1 \pmod{p}$ or $x_0 \equiv -x_1 \pmod{p}$. □

Lemma 6.5. *Let p be an odd prime. Then there are precisely $\frac{p-1}{2}$ quadratic residues modulo p.*

Proof. Define $\theta : \{1, \ldots, p-1\} \mapsto \{1, \ldots, p-1\}$ by $\theta(x)$ is the least positive residue of x^2 modulo p. Now by above, θ is 2 to 1, so $\#Im(\theta) = \frac{p-1}{2}$. □

6.2 Quadratic Reciprocity

Definition 6.6. *For p odd, $(a,p)=1$, we define the Legendre symbol $\left(\frac{a}{p}\right)$ by*

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p; \\ -1 & \text{otherwise.} \end{cases}$$

Lemma 6.7.

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$$

Proof. Follows from Euler's Criterion. □

Lemma 6.8 (Euler's Criterion).

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \mod p$$

Proof. Trivial if $\left(\frac{a}{p}\right) = 1$. So take $\left(\frac{a}{p}\right) = -1$. Now take $y \in \{1, \ldots, p-1\}$, then there exists unique $z \in \{1, \ldots, p-1\}$ such that $zy \equiv a \pmod p$, with $z \neq y$. So can break up $\{1, \ldots, p-1\}$ into $\frac{p-1}{2}$ distinct pairs whose product $\equiv a \pmod p$. So

$$\begin{aligned}(p-1)! &\equiv a^{\frac{p-1}{2}} \mod p \\ &\equiv -1 \qquad \text{by Wilson's Theorem}\end{aligned}$$

□

Theorem 6.9 (The Law of Quadratic Reciprocity). *If p and q are odd primes then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} = \begin{cases} 1 & \text{one of } p \text{ or } q \equiv 1 \pmod 4; \\ -1 & p, q \equiv -1 \pmod 4. \end{cases}$$

Lemma 6.10. *If p an odd prime, then*

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1 & p \equiv \pm 1 \pmod 8; \\ -1 & p \equiv \pm 3 \pmod 8. \end{cases}$$

Proof. None given - take it on trust... □

Example. *Compute* $\left(\frac{34}{97}\right)$.

$$\begin{aligned}\left(\frac{34}{97}\right) &= \left(\frac{2}{97}\right)\left(\frac{17}{97}\right) = +1\left(\frac{17}{97}\right) \\ &= \left(\frac{97}{17}\right) = \left(\frac{12}{17}\right) \\ &= \left(\frac{3}{17}\right)\left(\frac{4}{17}\right) = \left(\frac{3}{17}\right) \\ &= \left(\frac{17}{3}\right) = \left(\frac{2}{3}\right) \\ &= -1\end{aligned}$$

Example. *Is* $x^2 \equiv 20964 \pmod{1987}$ *soluble?* 1987 *is known to be prime.*

$$\begin{aligned}\left(\frac{20964}{1987}\right) &= \left(\frac{1094}{1987}\right)\\ &= \left(\frac{2}{1987}\right)\left(\frac{547}{1987}\right) = -\left(\frac{547}{1987}\right)\\ &= \left(\frac{1987}{547}\right) = \left(\frac{346}{547}\right)\\ &= \left(\frac{2}{547}\right)\left(\frac{173}{547}\right) = -\left(\frac{173}{547}\right)\\ &= -\left(\frac{547}{173}\right) = -\left(\frac{28}{173}\right) = -\left(\frac{7}{173}\right)\\ &= -\left(\frac{173}{7}\right) = -\left(\frac{5}{7}\right) = -\left(\frac{2}{5}\right)\\ &= 1\end{aligned}$$

So the congruence is soluble.

Example. *Compute* $\left(\frac{5}{p}\right)$, $p \neq 5$. *Let* $p = 5a + r, r = 1, 2, 3, 4$.

$$\begin{aligned}\left(\frac{5}{p}\right) &= \left(\frac{p}{5}\right)\\ &= \left(\frac{r}{5}\right)\\ &= \begin{cases}+1 & \text{if } r = 1, 4;\\ -1 & \text{if } r = 2, 3.\end{cases}\end{aligned}$$

Example. *Compute* $\left(\frac{3}{p}\right)$, $p \neq 3$.

$$\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right)(-1)^{\frac{p-1}{2}}$$

Let $p = 12a + r, r = 1, 5, 7, 11$

$$\begin{aligned}&= \left(\frac{r}{3}\right)\\ &= \begin{cases}+1 & r = 1, 11;\\ -1 & r = 5, 7.\end{cases}\end{aligned}$$

6.3 Introduction to Binary Quadratic Forms

Something of the form

$$f(x, y) = ax^2 + bxy + cy^2 \qquad a, b, c \in \mathbb{Z}$$

is called a binary quadratic form. We want to look at the problem of representation, i.e., given a fixed $f(x, y)$ and $m \in \mathbb{Z}$, find $x_0, y_0 \in \mathbb{Z}$ such that $f(x_0, y_0) = m$.

Definition 6.11.

$$SL_2(\mathbb{Z}) = \{\begin{pmatrix} p & q \\ r & s \end{pmatrix} \mid p, q, r, s \in \mathbb{Z}, ps - rq = 1\}$$

Definition 6.12.

$$\mathcal{B} = \{f(x, y) = ax^2 + bxy + cy^2 \mid a, b, c \in \mathbb{Z}\}$$

For the action of $SL_2(\mathbb{Z})$ on $\mathcal{B}$, take $\sigma \in SL_2(\mathbb{Z})$ and $f \in \mathcal{B}$. Define

$$\begin{aligned} \sigma \circ f &= f(px + qy, rx + sy) \\ &= a'x^2 + b'xy + c'y^2 \end{aligned}$$

where

$$\begin{aligned} a' &= f(p, r) \\ b' &= 2apq + 2crs + b(ps + qr) \\ c' &= f(q, s). \end{aligned}$$

You can check (if sufficiently bored), that $\sigma_1 \circ (\sigma_2 \circ f) = (\sigma_1\sigma_2) \circ f$.

Definition 6.13. *Two binary quadratic forms $f_1, f_2 \in \mathcal{B}$ are said to be equivalent if there exists $\sigma \in SL_2(\mathbb{Z})$ such that $f_2 = \sigma \circ f_1$.*

Definition 6.14. *The discriminant $\Delta(f) = b^2 - 4ac$.*

Lemma 6.15. $\Delta(\sigma \circ f) = \Delta(f)$

Proof. DIY! □

Note that inequivalent forms can have the same discriminant, for instance, $x^2 + 6y^2$ and $2x^2 + 3y^2$ both have discriminant -24, but are not equivalent.

$\Delta(f) = b^2 - 4ac$, so $\Delta(f) \equiv 0, 1 \pmod 4$.

Lemma 6.16. *For each $d \in \mathbb{Z}$ with $d \equiv 0, 1 \pmod 4$, there exists a binary quadratic form with d as discriminant.*

Proof. Given d, seek $a, b, c \in \mathbb{Z}$ with $b^2 - 4ac = d$. Take $a = 1$ and $b = 0, 1$ according as $d \equiv 0, 1 \pmod 4$. Take $c = \frac{-d}{4}$ if $d \equiv 0 \pmod 4$ and $c = \frac{1-d}{4}$ otherwise. These work! □

$$4af(x, y) = (2ax + by)^2 - \Delta(f)y^2$$

if $a \neq 0$, then f is positive definite when

$$a > 0 \text{ and } \Delta(f) < 0$$

negative definite when

$$a < 0 \text{ and } \Delta(f) < 0$$

and indeterminate when

$$a \neq 0 \text{ and } \Delta(f) > 0$$

6.4 Problem of Representation

Definition 6.17. *Let $m \in \mathbb{Z}$. We say m is properly represented by $f \in \mathcal{B}$ if $\exists p, r \in \mathbb{Z}$ with $(p, r) = 1$ such that $f(p, r) = m$.*

Lemma 6.18. *$m \in \mathbb{Z}$ is properly represented by $f \in \mathcal{B} \Leftrightarrow \exists f'$ equivalent to f such that m is the coefficient of x^2 in f'.*

Proof. If $f' = f(px + qy, rx + sy)$ with $ps - qr = 1$, and $f' = mx^2 + \ldots$, then $f' = f(p, r)x^2 + \ldots$, so $m = f(p, r)$ and $(p, r) = 1$.

Now assume $f(p, r) = m$ with $p, r \in \mathbb{Z}$ such that $(p, r) = 1$. Choose $q, s \in \mathbb{Z}$ such that $ps - qr = 1$. Form $\sigma = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$, then $\sigma \circ f = mx^2 + \ldots$. □

Corollary 6.19. *Assume $m \neq 0$ is properly represented by f. Then the congruence*

$$z^2 \equiv \Delta(f) \mod 4|m|$$

is soluble.

Proof. $f \sim f' = mx^2 + b'xy + c'y^2$ if m is properly represented by f. Now

$$\begin{aligned} \Delta(f) &= \Delta(f') \\ &= b'^2 - 4mc' \end{aligned}$$

So b' is a solution of the congruence. □

Lemma 6.20. *Assume f given, and $0 \neq m \in \mathbb{Z}$. Then if the congruence $z^2 \equiv \Delta(f) \pmod{4|m|}$ is soluble m is properly represented by some form with discriminant $\Delta(f)$.*

Proof. $z = b'$ is a solution of the congruence. Now $b'^2 - \Delta(f) = 4mc', c' \in \mathbb{Z}$ and define $f'(x, y) = mx^2 + b'xy + c'y^2$, which has discriminant $\Delta(f)$ and properly represents m. □

Example. *The primes represented by $x^2 + y^2$ are 2 and all p with $p \equiv 1 \pmod 4$. Trivial for p=2, so take $p > 2'$. Now all forms with discriminant -4 are equivalent to $x^2 + y^2$ (proof later), so*

$$\begin{aligned} p \text{ represented by } f &\Leftrightarrow z^2 \equiv -4 \mod 4p \text{ is soluble} \\ &\Leftrightarrow z = 2z_1 \text{ and } z_1^2 \equiv -1 \mod p \text{ is soluble} \\ &\Leftrightarrow z = 2z_1 \text{ and } \left(\frac{-1}{p}\right) = 1 \\ &\Leftrightarrow z = 2z_1 \text{ and } p \equiv 1 \mod 4 \end{aligned}$$

Example. *$f(x, y) = x^2 + xy + 2y^2$. The primes represented by f are 2 and all odd primes congruent to 1,2 or 4 modulo 7. 2 is trivial, so take $p > 2$. All forms of discriminant -7 are equivalent to f (proof later). So*

$$\begin{aligned} p \text{ represented by } f &\Leftrightarrow z^2 \equiv -7 \mod 4p \text{ is soluble} \\ &\Leftrightarrow z_1^2 \equiv -7 \mod 4 \text{ and } z_2^2 \equiv -7 \mod p \text{ are both soluble} \\ &\Leftrightarrow \left(\frac{-7}{p}\right) = 1 \\ &p \equiv 1, 2 \text{ or } 4 \mod 7 \end{aligned}$$

Step 2 is made using the Chinese remainder theorem.

6.5 Reduction Theory

Definition 6.21. $\mathcal{P} = \{f \in \mathcal{B} \mid a > 0$ *and* $\Delta(f) < 0\}$ *is the set of positive definite binary quadratic forms.*

$SL_2(\mathbb{Z})$ acts on $\mathcal{P}$.

Notation. *Write* (a, b, c) *for* $f(x, y) = ax^2 + bxy + cy^2$.

We now produce two members of $SL_2(\mathbb{Z})$ which make a, $|b|$ as small as possible.

If $c < a$, replace (a, b, c) by $(c, -b, a)$ using $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. If $|b| > a$, replace (a, b, c) by the equivalent form (a, b_1, c_1) where $b_1 = b + 2\mu a$, μ chosen such than $|b_1| < a$ and c_1 given by $\Delta(f) = b_1^2 - 4ac_1$, using $\left(\begin{smallmatrix} 1 & \mu \\ 0 & 1 \end{smallmatrix}\right)$

Now start with any form and apply these successively. At each stage either a or $|b|$ is reduced, so algorithm must terminate with a form which has $c \geq a$ and $|b| \leq a$. If $b = -a$ we can apply the second operation with $\mu = 1$ to change b to $+a$. If $c = a$, apply operation 1 to get $b \geq 0$. We have thus proved the following theorem.

Theorem 6.22. *Any element of* $\mathcal{P}$ *is equivalent to a form* $f(x, y) = ax^2 + bxy + cy^2$ *satisfying either* $c > a$ *and* $-a < b \leq a$ *or* $c = a$ *and* $0 \leq b \leq a$. *An element of* $\mathcal{P}$ *satisfying these conditions is said to be reduced. Additionally no two reduced forms are equivalent.*

Corollary 6.23. *If* $\Delta < 0$ *fixed, there are only finitely many positive definite reduced forms* (a, b, c) *of discriminant* Δ.

Proof. Put $D = -\Delta$. Now $4ac - b^2 = D$. If (a, b, c) reduced then $b^2 \leq a^2 \leq ac \Rightarrow 3ac \leq D$. There are only a finite number of possibilities for (a, c), each with only two choices of b. □

Definition 6.24. *If* $\Delta < 0$, *then* $h(\Delta)$ *is the number of equivalence classes of positive definite* (a, b, c) *with discriminant* Δ.

The above proof gives an algorithm to find $h(\Delta)$.

Example. $D = 4 \Rightarrow |b| \leq \sqrt{\frac{4}{3}}$ *and* b *even* $\Rightarrow b = 0$. *Now factor* 1*! to get* $a = c = 1$. *Thus there is a unique reduced form* $x^2 + y^2$.

Example. $D = 7 \Rightarrow |b| \leq \sqrt{\frac{7}{3}}$ *and* b *odd* $\Rightarrow b = \pm 1$. $b = -1$ *ruled out, since we want reduced form, so now factor* 2 *to get* $a = 1$, $c = 2$. *Thus there is a unique reduced form* $x^2 + xy + 2y^2$.

And so on. For an example with $h(\Delta) > 1$, put $\Delta = -20$ or $\Delta = -15$.

Example. *When* $\Delta = -15$, *get the two reduced forms* $x^2 + xy + 4y^2$ *and* $2x^2 + xy + 2y^2$. *Question: which primes are represented by at least one of these?*

Get $p \equiv 1, 2, 4,$ *or* $8 \pmod{15}$ *eventually. Now, can we decide which one?*

If $p = x^2 + xy + 4y^2$, *then* $4p = (2x + y)^2 + 15y^2$, *and* $4p \equiv (2x + y)^2 \pmod{15}$. *This implies that* p *is a square modulo 15, so* $p \equiv 1$ *or* $4 \pmod{15}$. *Similarly, by considering* $8p$, $p \equiv 2$ *or* $8 \pmod{15}$ *to be represented by* $2x^2 + xy + 2y^2$.

This is not always possible. No congruence condition on p can decide between $x^2 + 55y^2$ and $5x^2 + 11y^2$.

References

- P.M. Cohn, *Algebra Vol. 1*, Second ed., Wiley, 1993.

  This is the best book I found for the first part of the course. It has more in than is needed and is also quite good for Linear Maths.

- H. Davenport, *The Higher Arithmetic*, Sixth ed., CUP, 1992.

  A *very* good book for the last part of the course. It's also worth a read just for interest's sake.