The classical groups and their geometries

Notes from a seminar course

M. Chiara Tamburini Bellani

Lecce

Spring 2016

Index

Introduction

These notes are based on a 24 hours course given in the spring 2015 at the University of Milano Bicocca and the following year, in a revised and more complete version, at the University of Salento. In both cases it was part of the Dottorato di Ricerca programme. My aim here is to introduce students to the study of classical groups, an important instance of groups of Lie type, to their subgroup structure according to the famous classification Theorem of Aschbacher, and their matrix representations. My main references for such topics, which are absolutely central in abstract algebra and also reflect my personal tastes, have been [1], [2], [5], [6], [11], [13], [15] and [21].

These notes have no claim of completeness. For this reason each Chapter suggests more specific excellent textbooks, where a systematic treatment of the subject can be found. On the other hand a great deal of significant facts are presented, with proofs in several cases and a lot of examples.

As background I assume linear algebra and the basic notions of group theory, ring theory and Galois theory. As generale reference one may consult, for example, among many others: [9], [12], [14], [16], [17] and [19].

I am grateful to prof. Francesco Catino and the Università del Salento for the invitation and financial support. I appreciated a lot the warm hospitality of Maddalena and Francesco, which made so pleasant my short visits to the beautiful town of Lecce.

A special thank to my students of Milano and Lecce and also to prof. Salvatore Siciliano, dr. Paola Stefanelli and again to Maddalena and Francesco, for their stimulating and constructive attendance to my seminars.

Milano, September 2016.

Chapter I

Modules and matrices

Apart from the general reference given in the Introduction, for this Chapter we refer in particular to [8] and [20].

Let R be a ring with $1 \neq 0$. We assume most definitions and basic notions concerning left and right modules over R and recall just a few facts.

If M is a left R-module, then for every $m \in M$ the set $\mathrm{Ann}\,(m) := \{r \in R \mid rm = 0_M\}$ is a left ideal of R. Moreover $\mathrm{Ann}\,(M) = \bigcap_{m \in M} \mathrm{Ann}\,(m)$ is an ideal of R. The module M is *torsion free* if $\mathrm{Ann}\,(m) = \{0\}$ for all non-zero $m \in M$.

The regular module ${}_RR$ is the additive group $(R, +)$ considered as a left R-module with respect to the ring product. The submodules of ${}_RR$ are precisely the left ideals of R.

A finitely generated R-module is *free* if it is isomorphic to the direct sum of n copies of ${}_RR$, for some natural number n. Namely if it is isomorphic to the module

$$(0.1) \qquad ({}_RR)^n := \underbrace{{}_RR \oplus \cdots \oplus_R R}_{n \text{ times}}$$

in which the operations are performed component-wise. If R is commutative, then $({}_RR)^n \cong ({}_RR)^m$ only if $n = m$. So, in the commutative case, the invariant n is called the *rank* of $({}_RR)^n$. Note that $({}_RR)^n$ is torsion free if and only if R has no zero-divisors.

The aim of this Chapter is to determine the structure of finitely generated modules over a principal ideal domain (which are a generalization of finite dimensional vector spaces) and to describe some applications. But we start with an important result, valid for modules over any ring.

1 The Theorem of Krull-Schmidt

(1.1) Definition *An R-module M is said to be* indecomposable *if it cannot be written as the direct sum of two proper submodules.*

For example the regular module ${}_{\mathbb{Z}}\mathbb{Z}$ is indecomposable since any two proper ideals $n\mathbb{Z}$ and $m\mathbb{Z}$ intersect non-trivially. E.g. $0 \neq nm \in n\mathbb{Z} \cap m\mathbb{Z}$.

(1.2) Definition *Let M be an R-module.*

(1) *M is* noetherian *if, for every ascending chain of submodules*

$$M_1 < M_2 < M_3 < \dots$$

there exists $n \in \mathbb{N}$ such that $M_n = M_{n+r}$ for all $r \geq 0$;

(2) *M is* artinian *if, for every descending chain of submodules*

$$M_1 > M_2 > M_3 < \dots$$

there exists $n \in \mathbb{N}$ such that $M_n = M_{n+r}$for all $r \geq 0$.

(1.3) Lemma *An R-module M is noetherian if and only if every submodule of M is finitely generated.*

(1.4) Examples

- *every finite dimensional vector space is artinian and noetherian;*
- *the regular $\mathbb{Z}$-modulo ${}_{\mathbb{Z}}\mathbb{Z}$ is noetherian, but it is not artinian;*
- *for every field $\mathbb{F}$, the polynomial ring $\mathbb{F}[x_1, \dots, x_n]$ is noetherian.*

(1.5) Theorem *(Krull-Schmidt) Let M be an artinian and noetherian R-module. Given two decompositions*

$$M = M_1 \oplus M_2 \oplus M_n = N_1 \oplus N_2 \oplus N_m$$

suppose that the M_i-s and the N_j-s are indecomposable submodules. Then $m = n$ and there exists a permutation of the N_i-s such that M_i is isomorphic to N_i for all $i \leq n$.

2 Finitely generated modules over a PID

We indicate by D a *principal ideal domain* (PID), namely a commutative ring, without zero-divisors, in which every ideal is of the form $Dd = \langle d \rangle$, for some $d \in D$.
Every euclidean domain is a PID. In particular we have the following

(2.1) Examples *of PID-s:*

- *the ring* $\mathbb{Z}$ *of integers;*
- *every field* $\mathbb{F}$*;*
- *the polynomial ring* $\mathbb{F}[x]$ *over a field.*

Let A be an $m \times n$ matrix with entries in D. Then there exist $P \in \mathrm{GL}_m(D)$ and $Q \in \mathrm{GL}_n(D)$ such that PAQ is a pseudodiagonal matrix in which the entry in position (i,i) divides the entry in position $(i+1, i+1)$ for all i-s. The matrix PAQ is called a normal form of A. A consequence of this fact is the following:

(2.2) Theorem *Let* V *be a free* D*-module of rank* n *and* W *be a submodule.*

(1) W *is free of rank* $t \leq n$*;*

(2) *there exist a basis* $\mathcal{B} = \{v_1, \cdots, v_n\}$ *of* V *and a sequence* $d_1, \cdots, d_t$ *of elements of* D *with the following properties:*

   i) d_i *divides* d_{i+1} *for* $1 \leq i \leq t-1$*,*

   ii) $\mathcal{C} = \{d_1 v_1, \cdots, d_t v_t\}$ *is a basis of* W*.*

We may now state the structure theorem of a finitely generated D-module M. To this purpose let us denote by $\mathrm{d}(M)$ the minimal number of generators of M as a D-module.

(2.3) Theorem *Let* M *be a finitely generated* D*-module, with* $\mathrm{d}(M) = n$*.*
There exists a descending sequence of ideals:

$$Dd_1 \geq \cdots \geq Dd_n \quad \text{(invariant factors of } M\text{)} \tag{2.4}$$

with $Dd_1 \neq D$*, such that:*

$$M \simeq \frac{D}{Dd_1} \oplus \cdots \oplus \frac{D}{Dd_n} \quad \text{(normal form of } M\text{)}. \tag{2.5}$$

Let $t \geq 0$ be such that $d_t \neq 0_D$ and $d_{t+1} = 0_D$. Then, setting:

$$T := \{0_M\} \quad \text{if } t = 0, \qquad T := \frac{D}{Dd_1} \oplus \cdots \oplus \frac{D}{Dd_t} \quad \text{if } t > 0, \tag{2.6}$$

we have that $\mathrm{Ann}\,(T) = Dd_t$ *and T is isomorphic to the torsion submodule of M.*

M is torsion free if and only if $t = n$, $M = T$. Indeed, by this Theorem:

$$M \simeq T \oplus D^{n-t}$$

where D^{n-t} is free, of rank $n - t$.

Proof (sketch) Let $m_1, \ldots, m_n$ be a set of generators of M as a D-module. Consider the epimorphism $\psi : D^n \to M$ such that

$$\begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^{n} x_i m_i.$$

By Theorem 2.2, there exist a basis $\{v_1, \cdots, v_n\}$ of D^n and a sequence $d_1, \cdots, d_t$ of elements of D with the property that d_i divides d_{i+1} for $1 \leq i \leq t-1$, such that $\{d_1 v_1, \cdots, d_t v_t\}$ is a basis of Ker ψ. It follows $\frac{D^n}{\mathrm{Ker}\psi} \cong M$, whence:

$$\begin{aligned} \tfrac{Dv_1 \oplus \cdots \oplus Dv_t}{Dd_1 v_1 \oplus \cdots \oplus Dd_t v_t} \oplus \tfrac{Dv_{t+1} \oplus \cdots \oplus Dv_n}{\{0\} \oplus \cdots \oplus \{0\}} &\cong M \\ \tfrac{D}{Dd_1} \oplus \cdots \oplus \tfrac{D}{Dd_t} \oplus D \oplus \cdots \oplus D &\cong M. \end{aligned}$$

■

(2.7) Corollary *Let V be a vector space over $\mathbb{F}$, with $\mathrm{d}(V) = n$. Then $V \simeq \mathbb{F}^n$.*

(2.8) Corollary *Let M be a f.g. abelian group, with $\mathrm{d}(M) = n$. Then either:*

(1) *$M \simeq \mathbb{Z}^n$, or*

(2) *$M \simeq \mathbb{Z}_{d_1} \oplus \cdots \mathbb{Z}_{d_t} \oplus \mathbb{Z}^{n-t}$, $t \leq n$,*

where $d_1, \cdots, d_t$ is a sequence of integers ≥ 2, each of which divides the next one.

It can be shown that the normal form (2.5) of a f.g. D-module M is unique. Thus:

(2.9) Theorem *Two finitely generated D-modules are isomorphic if and only if they have the same normal form* (2.5) *or, equivalently, the same invariant factors* (2.4).

In the notation of Theorem 2.3, certain authors prefer to call invariant factors the elements $d_1, \ldots, d_n$ instead of the ideals generated by them. In this case the invariant factors are determined up to unitary factors.

(2.10) Example *Every abelian group of order p^3, with p prime, is isomorphic to one and only one of the following:*

- $\mathbb{Z}_{p^3}$, $t = 1$, $d_1 = p^3$;
- $\mathbb{Z}_p \oplus \mathbb{Z}_{p^2}$, $t = 2$, $d_1 = p$, $d_2 = p^2$;
- $\mathbb{Z}_p \oplus \mathbb{Z}_p \oplus \mathbb{Z}_p$, $t = 3$, $d_1 = d_2 = d_3 = p$.

(2.11) Example *Every abelian group of order 20 is isomorphic to one and only one of the following:*

- $\mathbb{Z}_{20}$, $t = 1$, $d_1 = 20$;
- $\mathbb{Z}_2 \oplus \mathbb{Z}_{10}$, $t = 2$, $d_1 = 2$, $d_2 = 10$.

3 The primary decomposition

We recall that D is a PID. For any $a, b \in D$ we have $Da + Db = Dd$, whence $d =$ G.C.D.(a, b). It follows easily that D is a unique factorization domain.
The results of this Section are based on the previous facts and the well known Chinese remainder Theorem, namely:

(3.1) Theorem *Let $a, b \in D$ such that* M.C.D.$(a, b) = 1$. *For all $b_1, b_2 \in D$, there exists $c \in D$ such that*

$$\begin{cases} c \equiv b_1 \pmod{a} \\ c \equiv b_2 \pmod{b}. \end{cases} \tag{3.2}$$

Proof There exist $y, z \in D$ such that $ay + bz = 1$. Multiplying by b_1 and b_2:

$$\begin{matrix} ayb_1 + bzb_1 = b_1 \\ ayb_2 + bzb_2 = b_2 \end{matrix}.$$

It follows

$$\begin{matrix} bzb_1 \equiv b_1 \pmod{a} \\ ayb_2 \equiv b_2 \pmod{b} \end{matrix}.$$

We conclude that $c = bzb_1 + ayb_2$ satisfies (3.2). ■

(3.3) Theorem *Let $d = p_1^{m_1} \dots p_k^{m_k}$, where each p_i is an irreducible element of D and $p_i \neq p_j$ for $1 \leq i \neq j \leq k$. Then:*

$$\frac{D}{Dd} \simeq \frac{D}{Dp_1^{m_1}} \oplus \cdots \oplus \frac{D}{Dp_k^{m_k}} \quad \text{(primary decomposition)}. \tag{3.4}$$

$Dp_1^{m_1}, \cdots, Dp_k^{m_k}$ (or simply $p_1^{m_1}, \cdots, p_k^{m_k}$) are the elementary divisors *of $\frac{D}{Dd}$.*

Proof Setting $a = p_1^{m_1}$, $b = p_2^{m_2} \dots p_k^{m_k}$, we have $d = ab$ with G.C.D.$(a, b) = 1$. The map

$$f : D \to \frac{D}{Da} \oplus \frac{D}{Db} \quad \text{such that} \quad x \mapsto \begin{pmatrix} Da + x \\ Db + x \end{pmatrix}$$

is a D-homomorphism. Moreover it is surjective by theorem 3.1. Finally Ker $f = Da \cap Db = Dd$. We conclude that

$$\frac{D}{Dd} \simeq \frac{D}{Da} \oplus \frac{D}{Db} = \frac{D}{Dp_1^{m_1}} \oplus \frac{D}{D\left(p_2^{m_2} \dots p_k^{m_k}\right)}$$

and our claim follows by induction on k. ■

(3.5) Examples

- $\mathbb{Z}_6 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_3$, *elementary divisors* $2, 3$;
- $\mathbb{Z}_6 \oplus \mathbb{Z}_6 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_3$, *elementary divisors* $2, 2, 3, 3$;
- $\mathbb{Z}_{40} \cong \mathbb{Z}_8 \oplus \mathbb{Z}_5$, *elementary divisors* $8, 5$;
- $\frac{\mathbb{C}[x]}{\langle x^3-1\rangle} \cong \frac{\mathbb{C}[x]}{\langle x-1\rangle} \oplus \frac{\mathbb{C}[x]}{\langle x-\omega\rangle} \oplus \frac{\mathbb{C}[x]}{\langle x-\overline{\omega}\rangle}$, *el. div.* $x-1,\ x-\omega,\ x-\overline{\omega}$ *where* $\omega = e^{\frac{i2\pi}{3}}$.

4 Modules over $\mathbb{F}[x]$ defined by matrices

Let $\mathbb{F}$ be a field. We recall that two matrices $A, B \in \mathrm{Mat}_n(\mathbb{F})$ are *conjugate* if there exist $P \in \mathrm{GL}_n(\mathbb{F})$ such that $P^{-1}AP = B$. The conjugacy among matrices is an equivalence relation in $\mathrm{Mat}_n(\mathbb{F})$, whose classes are called *conjugacy classes*. Our goal here is to find representatives for these classes.

The additive group $(\mathbb{F}^n, +)$ of column vectors is a left module over the ring $\mathrm{Mat}_n(\mathbb{F})$, with respect to the usual product of matrices. For a fixed matrix $A \in \mathrm{Mat}_n(\mathbb{F})$, the map: $\varphi_A : \mathbb{F}[x] \to \mathrm{Mat}_n(\mathbb{F})$ such that

$$f(x) \mapsto f(A)$$

is a ring homomorphism. It follows that $\mathbb{F}^n$ is an $\mathbb{F}[x]$-module with respect to the product:

$$f(x)\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} := f(A)\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}. \tag{4.1}$$

The $\mathbb{F}[x]$-module defined by (4.1) will be denoted by ${}_A\mathbb{F}^n$. Identifying $\mathbb{F}$ with the subring $\mathbb{F}x^0$ of $\mathbb{F}[x]$, the module ${}_A\mathbb{F}^n$ is a vector space over $\mathbb{F}$ in the usual way. Indeed, for all $\alpha \in \mathbb{F}$ and all $v \in {}_A\mathbb{F}^n$, we have: $(\alpha x^0)v = (\alpha A^0)v = \alpha v$.
Clearly, if V is any $\mathbb{F}[x]$-module, the map $\mu_x : V \to V$ such that

$$v \mapsto xv, \quad \forall\, v \in V \tag{4.2}$$

is an $\mathbb{F}[x]$-homomorphism. In particular μ_x is $\mathbb{F}$-linear.

(4.3) Theorem *Let V be an $\mathbb{F}[x]$-module, $dim_{\mathbb{F}}(V) = n$, and let $A, B \in \mathrm{Mat}_n(\mathbb{F})$.*

(1) *$V \simeq {}_A\mathbb{F}^n$ if and only if μ_x has matrix A with respect to a basis $\mathcal{B}$ of V;*

(2) *${}_A\mathbb{F}^n \simeq {}_B\mathbb{F}^n$ if and only if B is conjugate to A.*

Proof
(1) Suppose that μ_x has matrix A with respect to a basis $\mathcal{B}$ and call η the map which assigns to each $v \in V$ its coordinate vector $v_{\mathcal{B}}$ with respect to $\mathcal{B}$. We have:

$$Av_{\mathcal{B}} = (\mu_x(v))_{\mathcal{B}} = (xv)_{\mathcal{B}}, \quad \forall\, v \in V.$$

Clearly $\eta : V \to {}_A\mathbb{F}^n$ is an isomorphism of $\mathbb{F}$-modules. Moreover:

$$\eta(xv) = (xv)_{\mathcal{B}} = Av_{\mathcal{B}} = x\, v_{\mathcal{B}} = x\, \eta(v).$$

It follows easily that η is an isomorphism of $\mathbb{F}[x]$-modules. Thus $V \simeq {}_A\mathbb{F}^n$.
Vice versa, suppose that there exists an $\mathbb{F}[x]$-isomorphism $\gamma : V \to {}_A\mathbb{F}^n$. Set $\mathcal{B} = \{\gamma^{-1}(e_1),\ \dots,\ \gamma^{-1}(e_n)\}$, where $\{e_1, \dots, e_n\}$ is the canonical basis of $\mathbb{F}^n$. Then

$$\gamma(v) = \gamma\left(\sum_{i=1}^{n} k_i\, \gamma^{-1}(e_i)\right) = \sum_{i=1}^{n} k_i e_i = v_{\mathcal{B}}, \quad \forall\, v \in V.$$

Now $\gamma(xv) = x\gamma(v)$ gives $(\mu_x(v))_{\mathcal{B}} = Av_{\mathcal{B}}$. So μ_x has matrix A with respect to $\mathcal{B}$.

(2) Take $V = {}_A\mathbb{F}^n$, the $\mathbb{F}[x]$-module for which $\mu_x = \mu_A$. By the previous point ${}_A\mathbb{F}^n \simeq {}_B\mathbb{F}^n$ if and only if the linear map μ_A, induced by A with respect to the canonical basis, has matrix B with respect to an appropriate basis $\mathcal{B}$ of V. By elementary linear algebra this happens if and only if B is conjugate to A. ■

5 The rational canonical form of matrices

(5.1) Theorem *Let $A \in \mathrm{Mat}_n(\mathbb{F})$. The $\mathbb{F}[x]$-module ${}_A\mathbb{F}^n$ defined in* (4.1) *is finitely generated and torsion free.*

Proof $\mathbb{F}^n$ is finitely generated as a $\mathbb{F}$-module. Hence, a fortiori, as a $\mathbb{F}[x]$-module. In order to show that it is torsion free we must show that, for all $v \in \mathbb{F}^n$, there exists a non-zero polynomial $f(x) \in \mathbb{F}[x]$ such that $f(x)v = f(A)v = 0_{\mathbb{F}^n}$. This is clear if $A^i v = A^j v$ for some non-negative $i \neq j$. Because, in this case, we may take $f(x) = x^i - x^j$. Otherwise the subset $\{v,\ Av,\ \cdots,\ A^n v\}$ of $\mathbb{F}^n$ has cardinality $n+1$. It follows that there exist $k_0, \cdots, k_n$ in $\mathbb{F}$, not all zero, such that $k_0 v + k_1 Av + \cdots k_n A^n v = 0_{\mathbb{F}^n}$. So we may take $f(x) = k_0 + k_1 x + \cdots + k_n x^n$. ■

By Theorem 2.3 there exists a chain of ideals $\langle d_1(x)\rangle \geq \cdots \geq \langle d_t(x)\rangle \neq \{0\}$ such that

$$ {}_A\mathbb{F}^n \quad \simeq \quad \frac{\mathbb{F}[x]}{\langle d_1(x)\rangle} \oplus \ \cdots \ \oplus \frac{\mathbb{F}[x]}{\langle d_t(x)\rangle}. \tag{5.2}$$

Clearly $\langle d_t(x)\rangle = \mathrm{Ann}({}_A\mathbb{F}^n) = \mathrm{Ker}\ \varphi_A$. Moreover each $d_i(x)$ can be taken monic.

(5.3) Definition

(1) $d_1(x), \cdots,\ d_t(x)$ *are called the* similarity invariants *of A;*

(2) $d_t(x)$ *is called the* minimal polynomial *of A.*

(5.4) Definition *For a given monic polynomial of degree s*

$$d(x) = k_0 + k_1 x + k_2 x^2 \cdots + k_{s-1} x^{s-1} + x^s \in \mathbb{F}[x]$$

its companion matrix $C_{d(x)}$ *is defined as the matrix of* $\mathrm{Mat}_s(\mathbb{F})$ *whose columns are respectively $e_2, \ldots, e_s, [-k_0, \ldots, -k_{s-1}]^T$, namely the matrix:*

$$C_{d(x)} \ := \ \begin{pmatrix} 0 & 0 & \cdots & -k_0 \\ 1 & 0 & \cdots & -k_1 \\ 0 & 1 & \cdots & -k_2 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & -k_{s-1} \end{pmatrix}. \tag{5.5}$$

(5.6) Lemma *The companion matrix $C_{d(x)}$ has $d(x)$ as characteristic polynomial and as minimal polynomial.*

The first claim can be shown by induction on s, the second noting that

$$C_{d(x)}e_i = e_{i+1}, \quad i \leq s-1.$$

(5.7) Theorem *Consider the $\mathbb{F}[x]$-module $V = \frac{\mathbb{F}[x]}{\langle d(x)\rangle}$ and the map $\mu_x : V \to V$.*

(1) $\mathcal{B} := \left\{\langle d(x)\rangle + x^0,\ \langle d(x)\rangle + x,\ \cdots,\ \langle d(x)\rangle + x^{s-1}\right\}$ *is a basis of V over $\mathbb{F}$;*

(2) μ_x *has matrix $C_{d(x)}$ with respect to $\mathcal{B}$.*

Proof Routine calculation, noting that $\mu_x\left(\langle d(x)\rangle + f(x)\right) = \langle d(x)\rangle + xf(x)$. ■

We may now consider the general case. Let

$$V = \frac{\mathbb{F}[x]}{\langle d_1(x)\rangle} \oplus \cdots \oplus \frac{\mathbb{F}[x]}{\langle d_t(x)\rangle} = V_1 \oplus \cdots \oplus V_t$$

where each $d_i(x)$ is a monic, non-constant polynomial, and

$$d_i(x) \text{ divides } d_{i+1}(x),\ 1 \leq i \leq t-1. \tag{5.8}$$

With respect to the basis $\mathcal{B}_1 \times \{0_{V_2\oplus\cdots\oplus V_t}\}\ \dot\cup \ldots\ \dot\cup\ \mathcal{B}_t \times \left\{0_{V_1\oplus\cdots\oplus V_{t-1}}\right\}$, where each $\mathcal{B}_i$ is the basis of $\frac{\mathbb{F}[x]}{\langle d_i(x)\rangle}$ defined in Theorem 5.7, the map μ_x has matrix:

$$C = \begin{pmatrix} C_{d_1(x)} & & \\ & \cdots & \\ & & C_{d_t(x)} \end{pmatrix}. \tag{5.9}$$

(5.10) Definition *Every matrix C as in* (5.9)*, with $d_1(x),\ \ldots,\ d_t(x)$ satisfying* (5.8)*, is called a* rational canonical form.

(5.11) Lemma *The rational canonical form C in* (5.9) *has characteristic polynomial $\prod_1^t d_i(x)$ and minimal polynomial $d_t(x)$.*

From the above results we may conclude the following

(5.12) Theorem *For any field $\mathbb{F}$, every matrix $A \in \mathrm{Mat}_n(\mathbb{F})$ is conjugate to a unique rational canonical form.*

Clearly conjugate matrices have the same characteristic polynomial and the same minimal polynomial. So Lemma 5.11 has the following:

(5.13) Corollary (Theorem of Hamilton-Cayley). *Let $f(x)$ be the characteristic polynomial of a matrix A. Then $f(A) = 0$.*

(5.14) Example *The rational canonical forms in* $\mathrm{Mat}_2(\mathbb{F})$ *are of the following types:*
a) $t = 2$, $d_1(x) = d_2(x) = x - k$,

$$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix},$$

b) $t = 1$, $d_1(x) = x^2 + k_1x + k_0$,

$$\begin{pmatrix} 0 & -k_0 \\ 1 & -k_1 \end{pmatrix}.$$

(5.15) Example *The rational canonical forms in* $\mathrm{Mat}_3(\mathbb{F})$ *are of the following types:*
a) $t = 3$, $d_1(x) = d_2(x) = d_3(x) = x - k$,

$$\begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix},$$

b) $t = 2$, $d_1(x) = x - k$, $d_2(x) = (x - h)(x - k)$,

$$\begin{pmatrix} k & 0 & 0 \\ 0 & 0 & -kh \\ 0 & 1 & k + h \end{pmatrix},$$

c) $t = 1$, $d_1(x) = x^3 + k_2x^2 + k_1x + k_0$,

$$\begin{pmatrix} 0 & 0 & -k_0 \\ 1 & 0 & -k_1 \\ 0 & 1 & -k_2 \end{pmatrix}.$$

6 Jordan canonical forms

The rational canonical forms of matrices have the advantage of parametrizing the conjugacy classes of $\mathrm{Mat}_n(\mathbb{F})$ for any field $\mathbb{F}$. The disadvantage is that they say very little about eigenvalues and eigenspaces. For this reason, over an algebraically closed field, the Jordan canonical forms are more used and better known. They can be deduced from the primary decomposition of the $\mathbb{F}[x]$-modules associated to the rational canonical forms.

(6.1) Definition *For every* $\lambda \in \mathbb{F}$ *and every integer* $s \geq 0$ *we define inductively the* Jordan block $J(s, \lambda)$ *setting:*

$$J(0,\lambda) := \emptyset, \qquad J(1,\lambda) := (\lambda), \qquad J(s,\lambda) := \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 1 & & & \\ 0 & & & \\ \cdots & & J(s-1,\lambda) & \\ 0 & & & \end{pmatrix}, \ s > 1.$$

So, for example:

$$J(2,\lambda) = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}, \quad J(3,\lambda) = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}, \quad J(4,\lambda) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix}.$$

(6.2) Lemma *$J(s,\lambda)$ has λ as unique eigenvalue and corresponding eigenspace of dimension 1 in $\mathbb{F}^s$.*

Proof $J(s,\lambda)$ has characteristic polynomial $(x-\lambda)^s$, hence λ as unique eigenvalue. Elementary calculation shows that $\langle e_s \rangle$ is the corresponding eigenspace. ■

(6.3) Lemma *Let us consider the $\mathbb{F}[x]$-module*

$$V := \frac{\mathbb{F}[x]}{\langle (x-\lambda)^s \rangle}.$$

The Jordan block $J(s,\lambda)$ is the matrix of $\mu_x : V \to V$ with respect to the basis:

$$\mathcal{B}' := \left\{ I + (x-\lambda)^0, \quad I + (x-\lambda)^1, \quad \cdots, \quad I + (x-\lambda)^{s-1} \right\}.$$

In particular $J(s,\lambda)$ is conjugate to the companion matrix $C_{(x-\lambda)^s}$.

Proof For all $i \geq 0$ the following identity holds:

$$x(x-\lambda)^i = \lambda(x-\lambda)^i - \lambda(x-\lambda)^i + x(x-\lambda)^i = \lambda(x-\lambda)^i + (x-\lambda)^{i+1}.$$

It follows that, for $i \leq s-2$:

$\mu_x \left(I + (x-\lambda)^i \right) = I + x(x-\lambda)^i = \lambda \left(I + (x-\lambda)^i \right) \; + \; I + (x-\lambda)^{i+1}$,

$\mu_x \left(I + (x-\lambda)^{s-1} \right) = I + x(x-\lambda)^{s-1} = I + \lambda(x-\lambda)^{s-1} = \lambda \left(I + (x-\lambda)^{s-1} \right)$.

The last claim follows from Theorem 4.3. ■

(6.4) Corollary

(1) *Let* $d(x) = (x-\lambda_1)^{s_1} \dots (x-\lambda_m)^{s_m}$ *where* $\lambda_i \neq \lambda_j$ *for* $i \neq j$.

The companion matrix $C_{d(x)}$ *is conjugate to the matrix:*

$$J_{d(x)} := \begin{pmatrix} J(s_1,\lambda_1) & & \\ & \cdots & \\ & & J(s_m,\lambda_m) \end{pmatrix}. \tag{6.5}$$

(2) *Every rational canonical form* $C = \begin{pmatrix} C_{d_1(x)} & & \\ & \cdots & \\ & & C_{d_t(x)} \end{pmatrix}$ *is conjugate to*

$$J = \begin{pmatrix} J_{d_1(x)} & & \\ & \cdots & \\ & & J_{d_t(x)} \end{pmatrix} \quad \text{(Jordan form of } C\text{)}.$$

(6.6) Definition *In the above notation let* $\lambda_1, \dots, \lambda_m$ *be the distinct roots of* $d_t(x)$. *Set:*

$$d_i(x) = (x-\lambda_1)^{s_{i_1}} \ \dots \ (x-\lambda_m)^{s_{i_m}}, \quad 1 \leq i \leq t.$$

The factors of positive degree among

$$(x-\lambda_1)^{s_{1_1}}, \ \cdots, \ (x-\lambda_m)^{s_{1_m}}, \ \cdots, \ (x-\lambda_1)^{s_{t_1}}, \ \cdots, \ (x-\lambda_m)^{s_{t_m}}$$

(counted with their multiplicities) are called the elementary divisors *of* J.

(6.7) Example *If* $d_1(x) = (x-4), \quad d_2(x) = (x-3)(x-4)^2, \quad d_3(x) = (x-3)(x-4)^3$, *then the elementary divisors are:* $(x-4),\ (x-3),\ (x-4)^2,\ (x-3),\ (x-4)^3$.

So we have proved the following:

(6.8) Theorem *Let* $\mathbb{F}$ *be an algebraically closed field. Two matrices* A, B *in* $\mathrm{Mat}_n(\mathbb{F})$ *are conjugate if and only if they have the same Jordan form (up to a permutation of the blocks) or, equivalently, the same elementary divisors (counted with their multiplicities).*

We conclude this Section stating a useful result, not difficult to prove.

(6.9) Theorem *Let* $\mathbb{F}$ *be algebraically closed and let* $A \in \mathrm{Mat}_n(\mathbb{F})$. *The following conditions are equivalent:*

(1) A *is diagonalizable;*

(2) *the minimal polynomial of* A *has no multiple roots;*

(3) *every Jordan form of* A *is diagonal;*

(4) $\mathbb{F}^n$ *has a basis of eigenvectors of* A.

7 Exercises

(7.1) Exercise *Let $f : S \to R$ be a ring homomorphism. Show that every R-module M becomes an S-module by setting $sm := f(s)m, \quad \forall\ s \in S,\ m \in M$.*

(7.2) Exercise *Let p be a prime. Determine, up to isomorphisms, the abelian groups of order p^4.*

(7.3) Exercise *Determine, up to isomorphisms, the abelian groups of order 24 and order 100.*

(7.4) Exercise *Show that an euclidean domain is a principal ideal domain.*

(7.5) Exercise *Determine the primary decomposition and the normal form of the abelian group M having elementary divisors $2, 2, 4, 5, 5, 3, 9$. What is $Ann(M)$? What is the minimal number $d(M)$ of generators?*

(7.6) Exercise *Let D be a principal ideal domain and let d_1, d_2 be non-zero elements in D. Show that $Dd_1 = Dd_2$ if and only if $d_2 = \lambda d_1$ with λ invertible in D.*

(7.7) Exercise *Let M_1 and M_2 be R modules and $N_1 \leq M_1$, $N_2 \leq M_2$ be submodules. Show that:*

$$\frac{M_1 \oplus M_2}{N_1 \oplus N_2} \ \cong\ \frac{M_1}{N_1} \oplus \frac{M_2}{N_2}.$$

(7.8) Exercise *Suppose that R is a commutative ring. Let M be an R-module, m an element of M such that* Ann $(m) = \{0_R\}$. *Show that, for every ideal J of R:*

- *$Jm := \{jm \mid j \in J\}$ is a submodule of M;*
- *$\frac{Rm}{Jm} \cong \frac{R}{J}$ as R-modules.*

(7.9) Exercise *Calculate eigenvalues, eigenspaces, Jordan form and rational canonical form of each of the following matrices:*

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 1 & -1 & 0 \\ 2 & 1 & 1 & -1 \end{pmatrix}$$

Chapter II

The geometry of classical groups

We denote by V a vector space over the field $\mathbb{F}$. For simplicity we assume that its dimension is finite. Our main references here will be [11], [14], [15] and [21].

1 Sesquilinear forms

Let σ be an automorphism of $\mathbb{F}$ with $\sigma^2 = \mathrm{id}$. Set $\alpha^\sigma := \sigma(\alpha)$ for all $\alpha \in \mathbb{F}$.

(1.1) Definition *A σ-sesquilinear* form *on V is a map $(\ ,\ ) : V \times V \to \mathbb{F}$ such that, for every $\lambda, \mu \in \mathbb{F}$ and for every $u, v, w \in V$:*

(1) $(u, v + w) = (u, v) + (u, w)$,

(2) $(u + v, w) = (u, w) + (v, w)$,

(3) $(\lambda u, \mu v) = \lambda \mu^\sigma (u, v)$.

The form is said to be:

i) bilinear symmetric *if $\sigma = \mathrm{id}_\mathbb{F}$ and $(v, w) = (w, v)$, $\forall\ v, w \in V$;*

ii) bilinear antisymmetric *if $\sigma = \mathrm{id}_\mathbb{F}$ and $(v, v) = 0$, $\forall\ v \in V$;*

iii) hermitian *if $\sigma \neq \mathrm{id}_\mathbb{F}$, $\sigma^2 = \mathrm{id}_\mathbb{F}$ and $(v, w) = (w, v)^\sigma$, $\forall\ v, w \in V$;*

iv) non singular *if, for every $v \in V \setminus \{0_V\}$, there exists $u \in V$ such that $(u, v) \neq 0_\mathbb{F}$.*

(1.2) Definition *V is* non-singular *(or* non-degenerate*) when the form is non-singular.*

(1.3) Lemma *If the form is bilinear antisymmetric, then:*

$$(v, w) = -(w, v), \quad \forall\ v, w \in V.$$

Proof

$0 = (v+w, v+w) = (v,v)+(v,w)+(w,v)+(w,w) = (v,w)+(w,v) \Longrightarrow (v,w) = -(w,v)$.

■

(1.4) Definition *Let V, V' be vector spaces over $\mathbb{F}$, endowed with sesquilinear forms*

$$(\ ,\ ) : V \times V \to \mathbb{F}, \qquad (\ ,\ )' : V' \times V' \to \mathbb{F}.$$

(1) *An* isometry *from V to V' is an invertible element $f \in \mathrm{Hom}_{\mathbb{F}}(V, V')$ such that*

$$(f(v), f(w))' = (v, w), \quad \forall\ v, w \in V.$$

(2) *The spaces $(V, \mathbb{F}, (\ ,\ ))$ and $(V', \mathbb{F}, (\ ,\ )')$ are called* isometric *if there exists an isometry $f : V \to V'$.*

(1.5) Lemma *When $V = V'$, the set of isometries of V is a subgroup of $\mathrm{Aut}_{\mathbb{F}}(V)$, called the* group of isometries *of the form $(\ ,\ )$.*

The proof is left as an exercise.

(1.6) Theorem *(Witt's Extension Lemma) Let V be equipped with a non-degenerate form, either bilinear (symmetric or antisymmetric) or hermitian. Let U and W be subspaces and suppose that*

$$\tau : U \to W$$

is an isometry with respect to the restriction of the form to U and W, Then there exists an isometry $\hat{\tau} : V \to V$ which extends τ, namely such that $\hat{\tau}_U = \tau$.

For the proof of this important result see [1, page 81] or [14, page 367].

2 The matrix approach

Given a σ-sesquilinear form $(\ ,\ )$ on V, let us fix a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ of V over $\mathbb{F}$.

(2.1) Definition *The* matrix *J of the above form* with respect to *$\mathcal{B}$ is defined by*

$$J := ((v_i, v_j)), \quad 1 \le i, j \le n.$$

Given $v = \sum_{i=1}^n k_i v_i$, $w = \sum_{i=1}^n h_i v_i$ in V, it follows from the axioms that

$$(v,w) = \sum_{i,j=1}^{n} k_i h_j^{\sigma} (v_i, v_j) \ = \ v_{\mathcal{B}}^T J w_{\mathcal{B}}^{\sigma}, \quad \forall\ v, w \in V. \tag{2.2}$$

(2.3) Lemma *J is* the only *matrix of* $\mathrm{Mat}_n(\mathbb{F})$ *which satisfies* (2.2) *for the given form.*

Proof Let $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{F})$ satisfy $(v, w) = v_{\mathcal{B}}^T A w_{\mathcal{B}}^\sigma$ for all v, w in V.
Letting v, w vary in $\mathcal{B}$ and noting that $v_{i\mathcal{B}} = e_i$, $1 \le i \le n$ we have:

$$(v_i, v_j) \;=\; {v_i}_{\mathcal{B}}{}^T A {v_j}_{\mathcal{B}}{}^\sigma \;=\; {e_i}^T A e_j = a_{ij}, \quad 1 \le i, j \le n.$$

We conclude that $J = A$. ■

(2.4) Lemma *Let J be the matrix of a σ-sesquilinear form $(\ ,\ )$ on V.*

(1) *If $\sigma = \mathrm{id}_{\mathbb{F}}$, then the form is symmetric if and only if $J^T = J$;*

(2) *if $\sigma = \mathrm{id}_{\mathbb{F}}$, then the form is antisymmetric if and only if $J^T = -J$;*

(3) *if σ has order 2, then the form is hermitian if and only if $J^T = J^\sigma$.*

Moreover the form $(\ ,\ )$ is non-degenerate if and only if $\det J \neq 0$.

(2.5) Lemma *Let $J \in \mathrm{Mat}_n(\mathbb{F})$ be the matrix of a sesquilinear form on V with respect to a basis $\mathcal{B}$. Then $J' \in \mathrm{Mat}_n(\mathbb{F})$ is the matrix of the same form with respect to a basis $\mathcal{B}'$ if and only if J and J' are cogradient, namely if there exists P non-singular such that:*

$$J' = P^T J P^\sigma. \tag{2.6}$$

Proof Let J' be the matrix of the form with respect to $\mathcal{B}' = \{v'_1, \dots, v'_n\}$. Then:

$$v_{\mathcal{B}}^T J w_{\mathcal{B}}^\sigma = v_{\mathcal{B}'}^T J' w_{\mathcal{B}'}^\sigma, \quad \forall\ v, w \in V. \tag{2.7}$$

Setting $P := \big(\ (v'_1)_{\mathcal{B}} \,\big|\, \dots \,\big|\, (v'_n)_{\mathcal{B}}\ \big)$, we have $v_{\mathcal{B}} = P v_{\mathcal{B}'}$ for all $v \in V$. It follows:

$$v_{\mathcal{B}}^T J w_{\mathcal{B}}^\sigma = \left(v_{\mathcal{B}'}^T P^T\right) J \left(P^\sigma w_{\mathcal{B}'}^\sigma\right) = v_{\mathcal{B}'}^T \left(P^T J P^\sigma\right) w_{\mathcal{B}'}^\sigma, \quad \forall\ v, w \in V. \tag{2.8}$$

Comparing (2.7) with (2.8) we get $J' = P^T J P^\sigma$.
Vice versa, let $J' = P^T J P^\sigma$, for some non-singular P. Set $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ where $(v'_i)_{\mathcal{B}} = P e_i$. Then $\mathcal{B}'$ is a basis of V and $v_{\mathcal{B}} = P v_{\mathcal{B}'}$ for all $v \in V$. From (2.8) it follows that J' is the matrix of the form with respect to $\mathcal{B}'$. ■

(2.9) Theorem

(1) *Let J be the matrix of a sesquilinear form on $V = \mathbb{F}^n$ with respect to the canonical basis $\mathcal{B}$. Then its group of isometries is the subgroup:*

$$H := \left\{ h \in \mathrm{GL}_n(\mathbb{F}) \mid h^T J h^\sigma = J \right\}.$$

(2) *Let $\mathcal{B}'$ be another basis of $\mathbb{F}^n$. Then the group of isometries of the same form is:*

$$P^{-1} H P$$

where P is the matrix of the change of basis from $\mathcal{B}$ to $\mathcal{B}'$.

Proof

(1) If $\mathcal{B} = \{e_1, \ldots, e_n\}$ is the canonical basis, we have $v = v_{\mathcal{B}}$ for all $v \in V$. Thus:

$$(v, w) = v^T J w^\sigma, \quad \forall\ v, w \in V.$$

It follows that an element $h \in \mathrm{GL}_n(\mathbb{K})$ is an isometry if and only if:

$$v^T J w^\sigma = (hv)^T J (hw)^\sigma = v^T (h^T J h^\sigma) w^\sigma, \quad \forall\ v, w \in \mathbb{F}^n.$$

Equivalently h is an isometry if and only if

$$e_i^T J e_j = e_i^T (h^T J h^\sigma) e_j, \quad 1 \le i, j \le n \quad \Longleftrightarrow \quad J \ = \ h^T J h^\sigma.$$

(2) $J' = P^T J P^\sigma$ is the matrix of the form with respect to $\mathcal{B}'$. For every $h \in H$ we have:

$$(P^{-1} h P)^T J' \left(P^{-1} h P\right)^\sigma = J' \quad \Longleftrightarrow \quad h^T J h^\sigma = J.$$

■

3 Orthogonality

Let $(\ ,\ ) : V \times V \to \mathbb{F}$ be a bilinear (symmetric or antisymmetric) or an hermitian form.

(3.1) Definition *Two vectors $u, w \in V$ are said to be* orthogonal *if $(u, w) = 0_{\mathbb{F}}$.*

(3.2) Lemma *For every $W \subseteq V$ the subset*

$$W^\perp := \{ v \in V \mid (v, w) = 0,\ \forall\ w \in W \}$$

is a subspace, called the subspace orthogonal *to W.*

(3.3) Definition *Let W be a subspace of V. Then W is said to be*

(1) totally isotropic *(or* totally singular*) if* $W \leq W^{\perp}$*;*

(2) non-degenerate *if* $\mathrm{rad}(W) := W \cap W^{\perp} = \{0_V\}$.

Clearly V non singular $\iff \mathrm{rad}(V) = \{0_V\}$.

(3.4) Lemma *If* V *is non-degenerate then, for every subspace* W *of* V*:*

$$\dim W^{\perp} = \dim V - \dim W.$$

In particular:

(1) $\left(W^{\perp}\right)^{\perp} = W$*;*

(2) *the dimension of a totally isotropic space is at most* $\frac{\dim V}{2}$.

Proof Let $\mathcal{B}_W = \{w_1, \ldots, w_m\}$ be a basis of W. For every $v \in V$ we have:

$$v \in W^{\perp} \iff (w_i, v) = 0_{\mathbb{F}},\ 1 \leq i \leq m. \tag{3.5}$$

Extend $\mathcal{B}_W$ to a basis $\mathcal{B} = \{w_1, \ldots, w_m, w_{m+1}, \ldots, w_n\}$ of V and let J be the matrix of the form with respect to $\mathcal{B}$. From $(w_i)_{\mathcal{B}} = e_i$, $1 \leq i \leq m$, it follows:

$$v \in W^{\perp} \iff {e_i}^T J v_{\mathcal{B}}^{\sigma} = 0_{\mathbb{F}},\ 1 \leq i \leq m. \tag{3.6}$$

Since J is non-degenerate, its rows are independent. Hence the m equations of the linear homogeneous system (3.6) are independent. This system has n indeterminates, so the space of solutions has dimension $n - m$. We conclude that $W^{\perp}$ has dimension

$$n - m = \dim V - \dim W.$$

(1) $W \leq \left(W^{\perp}\right)^{\perp}$ and $\dim \left(W^{\perp}\right)^{\perp} = \dim\ V - \dim W^{\perp} = \dim W$.
(2) Let W be totally isotropic, i.e., $W \leq W^{\perp}$. Then:

$$\dim W \leq \dim W^{\perp} = \dim V - \dim W \quad \Longrightarrow \quad 2 \dim W \leq \dim V.$$

■

(3.7) Definition *Let* U, W *be subspaces of* V*. We write* $V = U \perp W$ *and say that* V *is an* orthogonal sum *of* U *and* W *if* $V = U \oplus W$ *and* U *is orthogonal to* W*, namely if:*

(1) $V = U + W$;

(2) $U \cap W = \{0_V\}$;

(3) $U \leq W^\perp$.

(3.8) Corollary *If V and W are non-degenerate, then*

$$V = W \perp W^\perp.$$

Moreover $W^\perp$ is non-degenerate.

Proof Since V is non-degenerate, Lemma 3.4 gives $\dim V = \dim\ W + \dim\ W^\perp$. Since W is non-degenerate, we have $W \cap W^\perp = \{0\}$. It follows $V = W \oplus W^\perp$. Finally $W^\perp$ is non-degenerate as $W^\perp \cap \left(W^\perp\right)^\perp = W^\perp \cap W = \{0\}$. ■

As a consequence of Witt's Lemma, we have the following:

(3.9) Corollary *Let V be endowed with a non-degenerate, either bilinear (symmetric or antisymmetric) or hermitian form. Then all the maximal totally isotropic subspaces have the same dimension, which is at most $\frac{\dim V}{2}$.*

Proof Let M be a totally isotropic subspace of largest possible dimension m. Clearly M is a maximal totally isotropic subspace. Take any totally isotropic subspace U. Since $\dim U \leq m$, there exists an injective $\mathbb{F}$-linear map $\tau : U \to M$. Now $\tau : U \to \tau(U)$ is an isometry, as the restriction of the form to U and to $\tau(U)$ is the zero-form. By theorem 1.6, there exists an isometry $\hat{\tau} : V \to V$ which extends τ. Thus $U \leq \hat{\tau}^{-1}(M)$ with $\hat{\tau}^{-1}(M)$ totally isotropic as $\hat{\tau}^{-1}$ is an isometry of V. If U is a maximal totally isotropic subspace, then $U = \hat{\tau}^{-1}(M)$ has dimension m. By Lemma 3.4 we have $m \leq \frac{\dim V}{2}$. ■

4 Symplectic spaces

(4.1) Definition *A vector space V over $\mathbb{F}$, endowed with a non-degenerate antisymmetric bilinear form is called* symplectic.

(4.2) Theorem *Let V be a symplectic space over $\mathbb{F}$, of dimension n. Then:*

(1) *$n = 2m$ is even;*

(2) *there exists a basis* $\mathcal{B}$ *of* V *with respect to which the matrix of the form is:*

$$J = \begin{pmatrix} \boldsymbol{0} & I_m \\ -I_m & \boldsymbol{0} \end{pmatrix}. \tag{4.3}$$

Proof Induction on n.

Suppose $n = 1$, $V = \mathbb{F}v$, $0 \neq v \in V$. For every $\lambda, \mu \in \mathbb{F}$: $(\lambda v, \mu v) = \lambda\mu(v, v) = 0_{\mathbb{F}}$, in contrast with the assumption that V is non degenerate. Hence $n \geq 2$.

Fix a non-zero vector $v_1 \in V$. There exists $w \in V$ such that $(v_1, w) \neq 0_{\mathbb{F}}$. In particular v_1 e w are linearly independent. Setting $w_1 := \lambda^{-1}w$, we have:

$$(v_1, w_1) = \left(v_1, \lambda^{-1}w\right) = \lambda^{-1}(v_1, w) = 1_{\mathbb{F}}.$$

If $n = 2$ our claim is proved since the matrix of the form w. r. to $\mathcal{B} = \{v_1,\ w_1\}$ is

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

If $n > 2$ we note that the subspace $W := \langle v_1, w_1 \rangle$ is non-singular. Thus:

$$V = W\ \perp\ W^{\perp}.$$

$W^{\perp}$ is non-degenerate, hence it is a symplectic space of dimension $n - 2$. By induction on n we have that $n - 2 = 2(m - 1)$ whence $n = 2m$, and moreover that $W^{\perp}$ admits a basis $\{v_2,\ \ldots,\ v_m,\ w_2,\ \ldots,\ w_m\}$ with respect to which the matrix of the form is

$$J_{W^{\perp}} = \begin{pmatrix} \mathbf{0} & I_{m-1} \\ -I_{m-1} & \mathbf{0} \end{pmatrix}.$$

Choosing $\mathcal{B} = \{v_1,\ v_2,\ \ldots,\ v_m,\ w_1,\ w_2,\ \ldots,\ w_m\}$ we obtain our claim. ■

(4.4) Definition *The group of isometries of a symplectic space* V *over* $\mathbb{F}$ *of dimension* $2m$ *is called the* symplectic group *of dimension* $2m$ *over* $\mathbb{F}$ *and indicated by* $\mathrm{Sp}_{2m}(\mathbb{F})$.

By the previous considerations, up to conjugation we may assume:

$$\mathrm{Sp}_{2m}(\mathbb{F}) = \left\{ g \in \mathrm{GL}_{2m}(\mathbb{F}) \mid g^T J g = J \right\}.$$

where J is as in (4.3). The subspace $\langle e_1, \ldots, e_m \rangle$, is a maximal totally isotropic space.

5 Some properties of finite fields

In contrast with the symplectic case, the classification of the non-singular, bilinear symmetric or hermitian forms, depends on the field $\mathbb{F}$ and may become very complicated. Thus our treatment will need further assumptions on $\mathbb{F}$. Since our interest is focused on finite fields, we will recall here a few specific facts, needed later, assuming the basic properties. As usual $\mathbb{F}_q$ denotes the finite field of order q, a prime power.

Consider the homomorphism $f : \mathbb{F}_q^* \to \mathbb{F}_q^*$ defined by $f(\alpha) := \alpha^2$. Clearly $\mathrm{Ker} f = \langle -1 \rangle$. If q is odd, $\mathrm{Ker} f$ has order 2. In this case $\mathrm{Im} f$, the set of non-zero squares in $\mathbb{F}_q$, has order $\frac{q-1}{2}$. Moreover, for any $\epsilon \in \mathbb{F}_q^* \setminus \mathrm{Im} f$, the coset $(\mathrm{Im} f)\,\epsilon \;=\; \left\{\alpha^2\epsilon \mid \alpha \in \mathbb{F}_q^*\right\}$ is the set of *non-squares*.
If q is even, $\mathrm{Ker} f$ has order 1. So f is surjective, i.e., every element of $\mathbb{F}_q$ is a square.

(5.1) Lemma *Every element of $\mathbb{F}_q$ is the sum of two squares.*

Proof By what observed above we may suppose q odd. Consider the set

$$X := \left\{\alpha^2 + \beta^2 \mid \alpha, \beta \in \mathbb{F}_q\right\}.$$

Note that $|X|$ does not divide $q = |\mathbb{F}_q|$, since:

$$|X| \geq \frac{q-1}{2} + 1 = \frac{q+1}{2} > \frac{q}{2}.$$

If every element of X were a square, X would be an additive subgroup of $\mathbb{F}_q$, in contrast with Lagrange's Theorem. So there exists a non-square $\epsilon \in X$. Write $\epsilon = \gamma^2 + \delta^2$. It follows that every non-square is in X. Indeed a non-square has shape $\alpha^2\epsilon = (\alpha\gamma)^2 + (\alpha\delta)^2$.

■

$\mathrm{Aut}\,(\mathbb{F}_{p^a}) = \mathrm{Gal}_{\mathbb{F}_p}(\mathbb{F}_{p^a})$ has order a. So $\mathrm{Aut}\,(\mathbb{F}_{p^a})$ is generated by the Frobenius automorphism $\alpha \mapsto \alpha^p$, which has has order a. It follows that $\mathbb{F}_{p^a}$ has an automorphism σ of order 2 if and only if $a = 2b$ is even. In this case, we set $q = p^b$, so that $\mathbb{F}_{p^a} = \mathbb{F}_{q^2}$.
The automorphism $\sigma : \mathbb{F}_{q^2} \to \mathbb{F}_{q^2}$ of order 2 is the map: $\alpha \mapsto \alpha^q$. Moreover $\alpha\alpha^q \in \mathbb{F}_q$ for all $\alpha \in \mathbb{F}_{q^2}$, since $(\alpha\alpha^q)^q = \alpha\alpha^q$.

(5.2) Theorem *The* Norm *map $N : \mathbb{F}_{q^2} \to \mathbb{F}_q$ defined by $N(\alpha) := \alpha\alpha^q$, is surjective.*

Proof The restriction of N to $\mathbb{F}_{q^2}^*$ is a group homomorphism into $\mathbb{F}_q^*$. Its kernel consists of the roots of $x^{q+1} - 1$, hence has order $\leq q + 1$. Thus its image has order $q - 1$. ■

6 Unitary and orthogonal spaces

We recall that σ denotes an automorphism of the field $\mathbb{F}$ such that $\sigma^2 = \mathrm{id}$. More precisely, $\sigma = \mathrm{id}$ in the orthogonal case, $\sigma \neq \mathrm{id}$ in the hermitian case.

(6.1) Lemma *Consider a non-degenerate, bilinear symmetric or hermitian form* $(\ ,\ ) : V \times V \to \mathbb{F}$. *If char* $\mathbb{F} = 2$ *assume that the form is hermitian. Then* V *admits an orthogonal basis, i.e., a basis with respect to which the matrix of the form is diagonal.*

Proof We first show that there exists v such that $(v, v) \neq 0$. This is clear when $\dim V = 1$, since the form is non-degenerate. So suppose $\dim V > 1$.
For a fixed non-zero $u \in V$, there exists $w \in V$ such that $(u, w) \neq 0_{\mathbb{F}}$. Clearly we may assume $(u, u) = (w, w) = 0$. If $\mathrm{char}\,\mathbb{F} \neq 2$, setting $\lambda = (u, w)$, $\quad v = \lambda^{-1}u + w$ we have:

$$(v, v) = \lambda^{-1}(u, w) + \left(\lambda^{-1}\right)^{\sigma}(w, u) = \lambda^{-1}\lambda + (\lambda^{\sigma})^{-1}\lambda^{\sigma} = 2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}.$$

If $\mathrm{char}\,\mathbb{F} = 2$, the form is hermitian by assumption. So there exists $\alpha \in \mathbb{F}$ such that $\alpha^{\sigma} \neq \alpha$. In this case, setting $v = \lambda^{-1}\alpha u + w$ we have $(v, v) = \alpha + \alpha^{\sigma} = \alpha - \alpha^{\sigma} \neq 0_{\mathbb{F}}$.
Induction on $\dim V$, applied to $\langle v \rangle^{\perp}$, gives the existence of an orthogonal basis of V. ∎

(6.2) Remark *The hypothesis char* $\mathbb{F} \neq 2$ *when the form is bilinear symmetric, is necessary. Indeed the matrix* $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *defines a non-degenerate symmetric form on* $V = \mathbb{F}_2^2$. *Since* $(v, v) = 0$ *for all* v, *no orthogonal basis can exist.*

Even the existence of an orthogonal basis is far from a complete classification as shown, for example, by a Theorem of Sylvester ([14, Theorem 6.7 page 359]).

(6.3) Example *By the previous theorem, the symmetric matrices*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

are pairwise not cogradient in $\mathrm{Mat}_3(\mathbb{R})$.

6.1 Unitary spaces

(6.4) Definition *A space* V, *with a non-degenerate hermitian form, is called* unitary.

(6.5) Theorem *Let* V *be a unitary space. Suppose that, for all* $v \in V$, *there exists* $\mu \in \mathbb{F}$ *such that* $N(\mu) := \mu\mu^{\sigma} = (v, v)$. *Then there exists an orthonormal basis of* V, *i.e., a basis with respect to which the matrix of the hermitian form is the identity.*

In particular such basis exists for $\mathbb{F} = \mathbb{C}$, σ the complex conjugation, and for $\mathbb{F} = \mathbb{F}_{q^2}$.

Proof By Lemma 6.1 there exists $v \in V$ with $(v, v) \neq 0$. Under our assumptions there exists $\mu \in \mathbb{F}$ such that $\mu\mu^\sigma = (v, v)$. Substituting v with $\mu^{-1}v$ we get $(v, v) = 1$. For $n = 1$ the claim is proved. So suppose $n > 1$. The subspace $\langle v \rangle$ is non-degenerate. It follows that $V = \langle v \rangle \perp \langle v \rangle^\perp$. As $\langle v \rangle^\perp$ is non-degenerate of dimension $n - 1$, our claim follows by induction. ■

(6.6) Definition *The group of isometries of a unitary space V over $\mathbb{F}$ of dimension n, called the* unitary group *of dimension n over $\mathbb{F}$, is indicated by* $\mathrm{GU}_n(\mathbb{F})$.

By Theorem 6.5, if $\mathbb{F} = \mathbb{C}$ and σ is the complex conjugation or $\mathbb{F} = \mathbb{F}_{q^2}$, we may assume:

$$\mathrm{GU}_n(\mathbb{F}) = \left\{ g \in \mathrm{GL}_n(\mathbb{F}) \mid g^T g^\sigma = I_n \right\}.$$

(6.7) Remark *There are fields which do not admit any automorphism of order 2: so there are no unitary groups over such fields. To the already mentioned examples of $\mathbb{R}$ and $\mathbb{F}_{p^{2b+1}}$, we add the algebraic closure $\overline{F}_p$ of $\mathbb{F}_p$, as shown below.*

By contradiction suppose there exists an automorphism σ of order 2 of $\mathbb{F} := \overline{F}_p$.
Let $\alpha \in \mathbb{F}$ be such that $\sigma(\alpha) \neq \alpha$. Since α is algebraic over $\mathbb{F}_p$, we have that $\mathbb{K} = \mathbb{F}_p(\alpha)$ is finite of order p^n for some n. Thus $\mathbb{K}$ is the splitting field of $x^{p^n} - x$. It follows that $\mathbb{K}$ is fixed by σ and $\sigma_{|\mathbb{K}}$ has order 2. Thus $n = 2m$, $|\mathbb{K}| = q^2$ with $q = p^m$ and $\sigma(\alpha) = \alpha^q$. Now consider the subfield $\mathbb{L}$ of $\mathbb{F}$ of order q^4. Again $\mathbb{L}$ is fixed by σ and $\sigma(\beta) = \beta^{q^2}$ for all β in $\mathbb{L}$. From $\mathbb{K} \leq \mathbb{L}$ we get the contradiction $\alpha \neq \sigma(\alpha) = \alpha^{q^2} = \alpha$.

6.2 Quadratic Forms

(6.8) Definition *A* quadratic form *on V is a map $Q : V \to \mathbb{F}$ such that:*

(1) *$Q(\lambda v) = \lambda^2 Q(v)$ for all $\lambda \in \mathbb{F}$, $v \in V$;*

(2) *the* polar *form $(v, w) := Q(v + w) - Q(v) - Q(w)$, $\forall\ v, w \in V$, is bilinear.*

Q is non-degenerate *if its polar form is non-degenerate.*

Note that:

$$Q(0_V) = Q(0_\mathbb{F} 0_V) = (0_\mathbb{F})^2 Q(0_V) = 0_\mathbb{F}. \tag{6.9}$$

Q uniquely determines its polar form $(\ ,\ )$ which is clearly symmetric. Moreover

$$2Q(v) = (v, v), \quad \forall\ v \in V. \tag{6.10}$$

Indeed: $Q(2v) = Q(v + v) = Q(v) + Q(v) + (v, v)$ gives $4Q(v) = 2Q(v) + (v, v)$.
It follows from (6.10) that, if char $(\mathbb{F}) = 2$, the polar form $(\ ,\ )$ is antisymmetric.
On the other hand, if car $\mathbb{F} \neq 2$, every symmetric bilinear form $(\ ,\ )$ is the polar form of the quadratic form Q defined by:

$$Q(v) := \frac{1}{2}(v, v), \quad \forall\ v \in V.$$

Direct calculation shows that Q is quadratic and that

$$Q(v + w, v + w) - Q(v) - Q(w) = (v, w).$$

By the above considerations, in characteristic $\neq 2$, the study of quadratic forms is equivalent to the study of symmetric bilinear forms. But, for a unified treatment, we study the orthogonal spaces via quadratic forms.

6.3 Orthogonal spaces

(6.11) Definition *Let (V, Q) and (V', Q') be vector spaces over $\mathbb{F}$, endowed with quadratic forms Q and Q' respectively. An* isometry *from V to V' is an invertible element $f \in \mathrm{Hom}_{\mathbb{F}}(V, V')$ such that*

$$Q'(f(v)) = Q(v), \quad \forall\ v \in V.$$

The spaces (V, Q) and (V', Q') are isometric *if there exists an isometry $f : V \to V'$.*

Clearly, when $V = V'$, $Q = Q'$, the isometries of V form a subgroup of $\mathrm{Aut}_{\mathbb{F}}(V)$.

(6.12) Definition *Let Q be a non degenerate quadratic form on V.*

(1) *(V, Q) is called an* orthogonal space*;*

(2) *the group of isometries of (V, Q), called the* orthogonal group *relative to Q, is denoted by $O_n(\mathbb{F}, Q)$, where $n = \dim V$.*

Note that, in an orthogonal space, we may consider orthogonality with respect to the polar form, which is non-singular by definition of orthogonal space.

(6.13) Lemma *Suppose char* $\mathbb{F} = 2$.

(1) *any orthogonal space* (V, Q) *over* $\mathbb{F}$ *has even dimension;*

(2) *the orthogonal group* $O_{2m}(\mathbb{F}, Q)$ *is a subgroup of the symplectic group* $\mathrm{Sp}_{2m}(\mathbb{F})$.

Proof
(1) The polar form of any quadratic form is antisymmetric by (6.10), hence degenerate in odd dimension.
(2) The polar form associated to Q is non-degenerate, antisymmetric and it is preserved by every $f \in O_{2m}(\mathbb{F}, Q)$. Indeed:

$$(v, w) := Q(v + w) - Q(v) - Q(w) = Q(f(v + w)) - Q(f(v)) - Q(f(w)) =$$

$$Q(f(v) + f(w)) - Q(f(v)) - Q(f(w)) = (f(v), f(w)), \quad \forall\ v, w \in V.$$

■

(6.14) Lemma *Let* (V, Q) *be an orthogonal space of dimension* ≥ 2. *If* $Q(v_1) = 0$ *for some non-zero vector* $v_1 \in V$, *then there exists* $v_{-1} \in V \setminus \langle v_1 \rangle$ *such that:*

$$Q(x_1 v_1 + x_{-1} v_{-1}) = x_1 x_{-1}, \quad \forall\ x_1, x_{-1} \in \mathbb{F}. \tag{6.15}$$

The subspace $\langle v_1, v_{-1} \rangle$ *is non-singular.*

Proof $Q(v_1) = 0$ gives $(v_1, v_1) = 2Q(v_1) = 0$. As the polar form of Q is non-degenerate, there exists $u \in V$ with $(v_1, u) \neq 0$. In particular v_1 and u are linearly independent. Set

$$v_{-1} := (v_1, u)^{-1} u - (v_1, u)^{-2} Q(u) v_1.$$

Then $v_{-1} \notin \langle v_1 \rangle$ and:

$$(v_1, v_{-1}) = 1, \quad Q(v_{-1}) = (v_1, u)^{-2} Q(u) - (v_1, u)^{-2} Q(u) = 0.$$

Using the assumption $Q(v_1) = 0$ we get (6.15). The subspace is non-singular as the matrix of the polar form with respect to $\{v_1, v_{-1}\}$ is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ■

(6.16) Definition *An orthogonal space* (V, Q) *is called* anisotropic *if* $Q(v) \neq 0$ *for all non-zero vectors* $v \in V$.

Non-singular anisotropic spaces exist.

(6.17) Example *Let V be a separable, quadratic field extension of $\mathbb{F}$. Then*

$$|\mathrm{Gal}_{\mathbb{F}}(V)| = \dim_{\mathbb{F}} V = 2 \quad \Longrightarrow \quad \mathrm{Gal}_{\mathbb{F}}(V) = \langle \sigma \rangle, \quad \mathbb{F} = V_{\langle \sigma \rangle}.$$

The Norm *map* $N_{\mathbb{F}}^{V} : V \to \mathbb{F}$ *defined by:*

$$N_{\mathbb{F}}^{V}(v) := vv^{\sigma}, \quad \forall\ v \in V$$

is a non-degenerate anisotropic quadratic form on V.

More details are given in the next Lemma.

(6.18) Lemma *Let $f(t) = t^2 + at + b \in \mathbb{F}[t]$ be separable, irreducible and consider*

$$V = \frac{\mathbb{F}[t]}{\langle t^2 + at + b \rangle} = \{x_1 + x_{-1}t \mid x_1,\ x_{-1} \in \mathbb{F}\}$$

with respect to the usual sum of polynomials and product modulo $f(t)$. Then :

$$N_{\mathbb{F}}^{V}(x_1 + x_{-1}t) = x_1^2 - ax_1x_{-1} + bx_{-1}^2, \quad \forall\ x_1, x_{-1} \in \mathbb{F}. \tag{6.19}$$

With respect to the basis $\{1, t\}$, the polar form of $N_{\mathbb{F}}^{V}$ is the non-singular matrix

$$J = \begin{pmatrix} 2 & -a \\ -a & 2b \end{pmatrix}.$$

Proof Let $\mathrm{Gal}_{\mathbb{F}}(V) = \langle \sigma \rangle$. Then t and t^{σ} are the roots of $f(t)$ in V. Thus

$$t + t^{\sigma} = -a, \quad tt^{\sigma} = b, \quad x^{\sigma} = x,\ \forall\ x \in \mathbb{F}.$$

It follows:

$$N_{\mathbb{F}}^{V}(x_1 + x_{-1}t) = (x_1 + x_{-1}t)(x_1 + x_{-1}t^{\sigma}) = -ax_1x_{-1} + x_1^2 + bx_{-1}^2.$$

J is non-degenerate since Det $(J) = 4b - a^2 \neq 0$ by the irreducibility of $t^2 + at + b$ (and its separability when char $\mathbb{F} = 2$). ■

(6.20) Remark *If $\mathbb{F} = \mathbb{F}_q$ then $V = \mathbb{F}_{q^2}$ and the map $N_{\mathbb{F}}^{V} : \mathbb{F}_{q^2} \to \mathbb{F}_q$ coincides with $v \mapsto vv^q = v^{q+1}$. As shown in Section 5 it is surjective. It follows that the map $\begin{pmatrix} x_1 \\ x_{-1} \end{pmatrix} \mapsto x_1^2 - ax_1x_{-1} + bx_{-1}^2$ from $\mathbb{F}_q^2$ to $\mathbb{F}_q$ is surjective.*

The anisotropic orthogonal spaces are only those of Example 6.17. We first show:

(6.21) Theorem *Let (W, Q) be an anisotropic orthogonal space of dimension 2.*

(1) *For each non-zero vector $v_1 \in W$ there exists $v_{-1} \in W \setminus \{v_1\}$ such that*

$$Q(x_1v_1 + x_{-1}v_{-1}) = Q(v_1)\left(x_1^2 + \zeta x_{-1}^2 + x_1x_{-1}\right) \quad \forall\ x_1,\ x_{-1} \in \mathbb{F} \tag{6.22}$$

where $t^2 - t + \zeta$ is irreducible in $\mathbb{F}[t]$.

(2) *If the map $\mathbb{F}^2 \to \mathbb{F}$ defined by $\begin{pmatrix} x_1 \\ x_{-1} \end{pmatrix} \mapsto x_1^2 + \zeta x_{-1}^2 + x_1x_{-1}$ is onto, the space (W, Q) is isometric to $(V, N_{\mathbb{F}}^V)$, where $V = \frac{\mathbb{F}[t]}{\langle t^2 - t + \zeta\rangle}$.*

In particular:

- *if $\mathbb{F}$ is algebraically closed, no such W exists;*
- *if $\mathbb{F} = \mathbb{F}_q$, all orthogonal anisotropic 2-dimensional spaces are isometric.*

Proof

(1) We first show that there exists $w \in W \setminus \langle v_1\rangle$ such that $(v_1, w) \neq 0$. Indeed, if $(v_1, v_1) \neq 0$, then $W = \langle v_1\rangle \oplus \langle v_1\rangle^\perp$ and we take $w = v_1 + u$ with $u \in \langle v_1\rangle^\perp$. If $(v_1, v_1) = 0$, then $\langle v_1\rangle \leq \langle v_1\rangle^\perp \neq W$ and we take $w \in W \setminus \langle v_1\rangle^\perp$.

Now set:

$$v_{-1} := Q(v_1)(v_1, w)^{-1}w, \quad \zeta = \frac{Q(v_{-1})}{Q(v_1)}.$$

It follows $(v_1, v_{-1}) = Q(v_1)$ and, for all $x_1, x_{-1} \in \mathbb{F}$:

$$Q(x_1v_1 + x_{-1}v_{-1}) = x_1^2Q(v_1) + x_{-1}^2Q(v_{-1}) + x_1x_{-1}Q(v_1) = Q(v_1)\left(x_1^2 + \zeta x_{-1}^2 + x_1x_{-1}\right).$$

In particular, for $x_{-1} = 1$, we get $x_1v_1 + v_{-1} \neq 0$, whence:

$$0 \neq Q(x_1v_1 + v_{-1}) = Q(v_1)\left(x_1^2 + x_1 + \zeta\right), \quad \forall\ x_1 \in \mathbb{F}.$$

Thus $t^2 + t + \zeta$ is irreducible in $\mathbb{F}[t]$, since it has no roots in $\mathbb{F}$. It follows that $t^2 - t + \zeta$ is also irreducible.

(2) There exists $\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \in \mathbb{F}^2$ such that $\lambda^2 + \zeta\mu^2 + \lambda\mu = Q(v_1)^{-1}$. Substituting v_1 with $\lambda v_1 + \mu v_{-1}$ in point (1), we may suppose $Q(v_1) = 1$. Then (6.22) gives $Q(x_1v_1 + x_{-1}v_{-1}) = x_1^2 + \zeta x_{-1}^2 + x_1x_{-1}$. We conclude that the map $f = W \to \frac{\mathbb{F}[t]}{\langle t^2 - t + \zeta\rangle}$ defined by:

$$x_1v_1 + x_{-1}v_{-1} \mapsto x_1 + x_{-1}t \tag{6.23}$$

is an isometry in virtue of (6.19).

Finally, suppose $\mathbb{F} = \mathbb{F}_q$ and let $\left(V, N^V_{\mathbb{F}_q}\right)$ $\left(V', N^{V'}_{\mathbb{F}_q}\right)$ be 2-dimensional anisotropic orthogonal spaces. Since V and V' are finite fields of the same order, there exists a field automorphism $f : V \to V'$ such that $f_{|\mathbb{F}_q} = \mathrm{id}$. From

$$f(v)f(v^q) = f(vv^q) = vv^q, \quad \forall\, v \in V$$

we conclude that f is an isometry. ■

(6.24) Corollary *Let (V, Q) be an orthogonal space, with $V = \mathbb{F}_q^{2m}$.*

(1) *There exists a basis $\mathcal{B} = \{v_1 \ldots, v_m, v_{-1} \ldots, v_{-m},\}$ of V such that either $Q = Q^+$ or $Q = Q^-$ where, for all $v = \sum_{i=1}^m x_i v_i + x_{-i} v_{-i} \in V$:*

- *$Q^+(v) = \sum_{i=1}^m x_i x_{-i}$;*
- *$Q^-(v) = \sum_{i=1}^m x_i x_{-i} + x_m^2 + \zeta x_{-m}^2$, with $t^2 - t + \zeta$ a fixed, separable irreducible polynomial in $\mathbb{F}_q[t]$ (arbitrarily chosen with these properties).*

(2) *Q^+ has polar form $\sum_{i=1}^m (x_i y_{-i} + x_{-i} y_i)$, with matrix $J_1 = \begin{pmatrix} \boldsymbol{0} & I_m \\ I_m & \boldsymbol{0} \end{pmatrix}$;*

Q^- has polar form $\sum_{i=1}^m (x_i y_{-i} + x_{-i} y_i) + 2\,(x_m y_m + \zeta x_{-m} y_{-m})$, with matrix

$$J_2 = \begin{pmatrix} \boldsymbol{0} & I_{m-1} & 0 & 0 \\ I_{m-1} & \boldsymbol{0} & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2\zeta \end{pmatrix}.$$

(3) *(V, Q^+) is not isometric to (V, Q^-).*

The corresponding groups of isometries are indicated by $O^+_{2m}(q)$ and $O^-_{2m}(q)$.

Proof

(1) Let $m = 1$. If V is non-anisotropic, Lemma 6.14 gives $Q = Q^+$. If V is anisotropic, Theorem 6.21 gives $Q = Q^-$. So assume $m > 1$.

Step 1. We claim that there exists a non-zero vector $v_1 \in V$ such that $Q(v_1) = 0$.
By the same argument used in the proof of point (1) of Theorem 6.21, there exists a non-singular 2-dimensional subspace $W = \langle v_m, v_{-m} \rangle$. We may assume that W is anisotropic. Hence (W, Q) is isometric to $\left(\mathbb{F}_{q^2}, N^{\mathbb{F}_{q^2}}_{\mathbb{F}_q}\right)$ and

$$Q\,(x_m v_m + x_{-m} v_{-m}) = x_m x_{-m} + x_m^2 + \zeta x_{-m}^2, \quad \forall\, x_m, x_{-m} \in \mathbb{F}_q$$

for some irreducible polynomial $t^2 - t + \zeta \in \mathbb{F}[t]$.

Take a non-zero vector w in $W^{\perp}$. By the surjectivity of the norm for finite fields, there exist $u \in W$ such that $Q(u) = -Q(w)$. Then $v_1 = u + w \neq 0$, since $W \cap W^{\perp} = \{0\}$. Moreover, from $(u, w) = 0$, we get: $Q(v_1) = Q(u + w) = Q(u) + Q(w) = 0$.

Step 2. By Lemma 6.14 there exists a non-singular 2-dimensional subspace $\langle v_1, v_{-1}\rangle$ such that $Q(x_1 v_1 + x_{-1} v_{-1}) = x_1 x_{-1}$. We get:

$$V = \langle v_1, v_{-1}\rangle \oplus \langle v_1, v_{-1}\rangle^{\perp}.$$

By induction, $\langle v_1, v_{-1}\rangle^{\perp}$ has a basis $\mathcal{B}' = \{v_2 \ldots, v_m, v_{-2} \ldots, v_{-m},\}$ such that the restriction of Q to $\langle v_1, v_{-1}\rangle^{\perp}$ is either Q^+ or Q^-. This gives (1).
(2) Routine calculation using (1).
(3) V is a direct sum of mutually orthogonal 2-dimensional spaces:

$$V = \langle v_1, v_{-1}\rangle \perp \cdots \perp \langle v_m, v_{-m}\rangle$$

with the further property $(v_i, v_i) = 0$, $1 \leq i \leq m-1$. For Q^+ we have also $(v_m, v_m) = 0$, so that $\langle v_1, \ldots, v_m\rangle$ is a totally isotropic space of largest possible dimension $m = \frac{n}{2}$ (see Lemma 3.9). For Q^- the space $W = \langle v_1, \ldots, v_{m-1}\rangle$ is totally isotropic. It follows:

$$W \oplus \langle v_m, v_{-m}\rangle = W^{\perp}.$$

Let $\widehat{W}$ be a totally isotropic space which contains W. Then

$$W = W + \left(\widehat{W} \cap \langle v_m, v_{-m}\rangle\right) = W + \{0\} = W$$

since $\langle v_m, v_{-m}\rangle$ is anisotropic. We conclude that $W = \widehat{W}$, i.e., W is a maximal isotropic space of dimension $m-1$. So Q^+ and Q^- cannot be isometric. ■

(6.25) Theorem *Let (V, Q) be an orthogonal space, with $V = \mathbb{F}_q^{2m+1}$, q odd. There exists a basis of V such that the matrix of the polar form is one of the following:*

$$I_{2m+1} = \begin{pmatrix} 1 & & \\ & \ldots & \\ & & 1 \end{pmatrix}, \quad J = \begin{pmatrix} I_{2m} & \\ & \epsilon \end{pmatrix}, \tag{6.26}$$

where ϵ is a fixed non-square in $\mathbb{F}_q^*$ (arbitrarily chosen with this property). The two polar forms I_{2m+1} and J give rise to non-isometric orthogonal spaces, but their groups of isometries are conjugate, hence isomorphic. Both groups are indicated by $O_{2m+1}(q)$.

Proof We first show that, if an orthogonal space V over $\mathbb{F}_q$, has dimension > 1, then there exists $v_1 \in V$ with $(v_1, v_1) = 1$. By Lemma 6.1, there exists v_1 such that $(v_1, v_1) \neq 0$. Thus $(v_1, v_1) = \rho^2$ or $(v_1, v_1) = \rho^2\epsilon$ for some $\rho \in \mathbb{F}_q^*$. Substituting v_1 with $\rho^{-1}v_1$, if necessary, we have $(v_1, v_1) \in \{1, \epsilon\}$. If $(v_1, v_1) = \epsilon$, set $\lambda^2 + \mu^2 = \epsilon^{-1}$. Again by Lemma 6.1, applied to $\langle v_1\rangle^\perp$, there exists $v_2 \in \langle v_1\rangle^\perp$ such that $(v_2, v_2) \neq 0$. If $(v_2, v_2) = 1$, we substitute v_1 by v_2. If $(v_2, v_2) = \epsilon$, we substitute v_1 by $\lambda v_1 + \mu v_2$.
Now we prove our claim. If $m = 1$ we can take $\mathcal{B} = \{v_1\}$ with $(v_1, v_1) \in \{1, \epsilon\}$. If $m > 1$ we take v_1 with $(v_1, v_1) = 1$. Then $V = \langle v_1\rangle \perp \langle v_1\rangle^\perp$ and our claim follows by induction on $\dim V$ applied to $\langle v_1, v_2\rangle^\perp$.
I_{2m+1} and J define non isometric spaces because the dimension of a maximal isotropic space are, respectively, m and $m-1$. So J is not cogradient to I_{2m+1}. Also ϵI_{2m+1} is not cogredient to I_{2m+1}, otherwise we would have $\epsilon I_{2m+1} = P^T I_{2m+1} P$, a contradiction as $\epsilon^{2m+1} = \det(\epsilon I_{2m+1})$ is not a square. Since, over $\mathbb{F}_q$, there are only 2 non-isometric orthogonal spaces, J is cogredient to ϵI_{2m+1}. Now I_{2m+1} and ϵI_{2m+1} have the same group of isometries, since:

$$h^T(\epsilon I_{2m+1})h = \epsilon I_{2m+1} \iff h^T I_{2m+1} h = I_{2m+1}.$$

We conclude that the groups of isometries of I_{2m+1} and J are conjugate. ■

7 Exercises

(7.1) Exercise *Show that* $\mathrm{SL}_2(\mathbb{F}) = \mathrm{Sp}_2(\mathbb{F})$ *over any field* $\mathbb{F}$.

(7.2) Exercise *Let* $(V, Q, \mathbb{F})$ *be an orthogonal space. Suppose* $V = V_1 \perp V_2$.
Show that, for each $v = v_1 + v_2$ *with* $v_1 \in V_1$, $v_2 \in V_2$*:*

$$Q(v) = Q(v_1) + Q(v_2).$$

(7.3) Exercise *Let* V *be a quadratic extension of* $\mathbb{F}$ *and* $\langle\sigma\rangle = \mathrm{Gal}_{\mathbb{F}}(V)$.
Show that the map $N_{\mathbb{F}} : V \to \mathbb{F}$*, defined by* $N_{\mathbb{F}}^V(v) := vv^\sigma$ *is a quadratic form on* V.

(7.4) Exercise *In Lemma* 6.18 *show that the quadratic form*

$$N_{\mathbb{F}}^V(x_1 + x_{-1}t) = x_1^2 - ax_1x_{-1} + b$$

has matrix $J = \begin{pmatrix} 2 & -a \\ -a & 2b \end{pmatrix}$ *with respect to the basis* $\{1, t\}$.

(7.5) Exercise *Say whether the matrices*

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad J' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

are cogredient. In case they are, indicate a non-singular matrix P *such that* $P^T J P = J'$.

(7.6) Exercise *Let* V *be an anisotropic* 2*-dimensional orthogonal space over* $\mathbb{F}_q$, q *odd. Show that there exists a basis for which the polar form has matrix:* $\begin{pmatrix} 1 & 0 \\ 0 & -\epsilon \end{pmatrix}$, *where* ϵ *is a non square in* $\mathbb{F}_q$.

(7.7) Exercise *Let* q *be odd. Show that* -1 *is a square in* $\mathbb{F}_q$ *if and only if*

$$q \equiv 1 \pmod 4.$$

(7.8) Exercise *Let* q *be odd and* $\epsilon \in \mathbb{F}_q$ *be a non-square. Show that the matrices*

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix}$$

are not cogredient (equivalently define non-isometric orthogonal spaces).

(7.9) Exercise *Let* q *be odd and* $\epsilon \in \mathbb{F}_q$ *be a non-square. Show that the matrix* $J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *is respectively cogredient to*

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ if } q \equiv 1 \pmod 4, \quad \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix} \text{ if } q \equiv 3 \pmod 4.$$

(7.10) Exercise *Let* W *be a totally isotropic subspace of an orthogonal space* V. *Suppose*

$$V = W \oplus U$$

with U *anisotropic. Show that* W *is a maximal isotropic subspace of* V.

(7.11) Exercise *Let* q *be odd,* $V = \mathbb{F}_q^n$ *be a quadratic space, with* $n = 2m$. *Using the classification of quadratic spaces given in this Chapter, show that there exists a basis of* V *with respect to which the polar form has matrix* J_1 *or* J_2 *where*

$$J_1 = \begin{pmatrix} \boldsymbol{0} & I_m \\ I_m & \boldsymbol{0} \end{pmatrix}, \quad J_2 = \begin{pmatrix} \boldsymbol{0} & I_{m-1} & & \\ I_{m-1} & \boldsymbol{0} & & \\ & & 1 & \\ & & & -\epsilon \end{pmatrix}.$$

Chapter III

The finite simple classical groups

Apart from the general reference given in the Introduction, in this Chapter we mainly refer to [5], [11], [15], [22].

1 A criterion of simplicity

(1.1) Definition *A subgroup M of a group $G \neq \{1\}$ is said to be* maximal *if $M \neq G$ and there exists no subgroup $\widehat{M}$ such that $M < \widehat{M} < G$.*

If M is maximal in G, then every conjugate gMg^{-1} of M is maximal in G. Indeed

$$gMg^{-1} < N < G \implies M < g^{-1}Ng < G.$$

Let G be a subgroup of $\mathrm{Sym}(X)$. For any $\alpha \in X$, the set

$$G_\alpha := \{x \in G \mid x(\alpha) = \alpha\}$$

is a subgroup, called the *stabilizer* of α in G. If $\beta = g(\alpha)$ then $G_\beta = gG_\alpha g^{-1}$.

(1.2) Definition *Let $k \in \mathbb{N}$. $G \leq \mathrm{Sym}(X)$ is called:*

- k-transitive *if, for any two k-tuples of pairwise distinct elements in X:*

  $$(\alpha_1, \ldots, \alpha_k)\,,\ (\beta_1, \ldots, \beta_k)$$

  there exists $g \in G$ such that $g(\alpha_i) = \beta_i$, $1 \leq i \leq k$;

- transitive *if it is 1-transitive;*

- primitive *if it is transitive and G_α is a maximal subgroup of G for (any) $\alpha \in X$.*

To prove that G is transitive on X it is enough to fix $\gamma \in X$ and show that, for any $\alpha \in X$, there exists $g \in G$ such that $g(\gamma) = \alpha$. Actually a more general fact holds:

(1.3) Lemma *Let $G \leq \mathrm{Sym}(X)$ and $(\gamma_1, \ldots, \gamma_k)$ be a fixed k-tuple of distinct elements in X. If, for every k-tuple $(\alpha_1, \ldots, \alpha_k)$ of distinct elements in X there exists $g \in G$ such that $g(\gamma_i) = \alpha_i$, $1 \leq i \leq k$, then G is k-transitive.*

Proof Given $(\alpha_1, \ldots, \alpha_k)$, $(\beta_1, \ldots, \beta_k)$ let $g_1, g_2 \in G$ be such that:

$$g_1(\gamma_i) = \alpha_i, \quad g_2(\gamma_i) = \beta_i, \quad 1 \leq i \leq k.$$

Then $g_2 g_1^{-1}(\alpha_i) = \beta_i$, $1 \leq i \leq k$. ■

(1.4) Lemma *If $G \leq \mathrm{Sym}(X)$ is 2-transitive, then G is primitive.*

Proof Let $G_\alpha < H \leq G$, with $\alpha \in X$. We want to show that $H = G$. To this purpose, choose $h \in H \setminus G_\alpha$ and set $\beta = h(\alpha)$. So $\beta \neq \alpha$. Now take any $g \in G$. If $g(\alpha) = \alpha$, then $g \in H$. Otherwise $g(\alpha) = \gamma \neq \alpha$ and there exists $\overline{h} \in G$ such that $\big(\overline{h}(\alpha), \overline{h}(\beta)\big) = (\alpha, \gamma)$ since G is 2-transitive. In particular $\overline{h} \in G_\alpha < H$. Moreover, from $\overline{h}(\beta) = \gamma$ we get $\overline{h}h(\alpha) = g(\alpha)$. Thus $g^{-1}\overline{h}h \in G_\alpha < H$. From $\overline{h}h \in H$ it follows $g \in H$. So $G = H$. ■

(1.5) Definition *The* derived subgroup *G' of an abstract group G is the subgroup generated by all* commutators *$x^{-1}y^{-1}xy := (x, y)$, i.e.,:*

$$G' := \left\langle x^{-1}y^{-1}xy \mid x, y \in G \right\rangle.$$

If N is a (normal) subgroup of G, then $\frac{G}{N}$ is abelian if and only if $G' \leq N$.

(1.6) Definition *A group $S \neq \{1\}$ is* simple *if its normal subgroups are $\{1\}$ and S.*

The following Theorem provides a fundamental tool by which the simplicity of the classical groups can be proved.

(1.7) Theorem *(Iwasawa's Lemma). A subgroup S of $\mathrm{Sym}(X)$ is a simple group whenever the following conditions hold:*

- *S is primitive;*
- *$S = S'$, i.e., S is perfect;*

- *the stabilizer S_α of (any) $\alpha \in X$ contains a normal abelian subgroup A such that S is generated by the conjugates of A, i.e., $S = A^S := \langle A^s \mid s \in S \rangle$.*

Proof $X = \{s(\alpha) \mid s \in S\}$, by the transitivity of S. Let N be a normal subgroup of S. If $N \leq S_\alpha$, every $x = s(\alpha) \in X$ is fixed by $sNs^{-1} = N$, whence $N = \{\mathrm{id}\}$. So assume:

$$N \not\leq S_\alpha. \tag{1.8}$$

Since S_α normalizes N, the product $S_\alpha N = NS_\alpha$ is a subgroup of S. Moreover $S_\alpha \neq NS_\alpha$ in virtue of (1.8). By the maximality of S_α in the primitive group S we get

$$S_\alpha N = S. \tag{1.9}$$

From the assumptions $S = A^S$, A normal in S_α and N normal in S, it follows:

$$S = A^S = A^{S_\alpha N} = A^N \leq NA \leq S.$$

Thus $S = NA$ and

$$\frac{S}{N} = \frac{NA}{N} \cong \frac{A}{A \cap N} \quad \text{abelian} \quad \Longrightarrow \quad S' \leq N.$$

Finally, from $S' = S$ we conclude $S = N$. ■

2 The projective special linear groups

2.1 The action on the projective space

(2.1) Definition *The group of $n \times n$ invertible matrices, with entries in $\mathbb{F}$, is called the* general linear group *of degree n over $\mathbb{F}$, and indicated by $\mathrm{GL}_n(\mathbb{F})$ or $\mathrm{GL}_n(q)$ if $\mathbb{F} = \mathbb{F}_q$.*

We recall that, over the field $\mathbb{F}$, a matrix is invertible if and only if it has non-zero determinant. By the Theorem of Binet, the map

$$\delta : \mathrm{GL}_n(\mathbb{F}) \to \mathbb{F}^* \quad \text{such that} \quad A \mapsto \det A \tag{2.2}$$

is a homomorphism of groups. Clearly δ is surjective. Its kernel, consisting of the matrices of determinant 1, is called the *special linear group* of degree n over $\mathbb{F}$ and is indicated by $\mathrm{SL}_n(\mathbb{F})$ or $\mathrm{SL}_n(q)$ if $\mathbb{F} = \mathbb{F}_q$. It follows $\frac{\mathrm{GL}_n(\mathbb{F})}{\mathrm{SL}_n(\mathbb{F})} \sim \mathbb{F}^*$. In particular:

$$\frac{|\mathrm{GL}_n(q)|}{|\mathrm{SL}_n(q)|} = q - 1. \tag{2.3}$$

The center Z of $\mathrm{GL}_n(\mathbb{F})$ is defined as

$$Z := \{z \in \mathrm{GL}_n(\mathbb{F}) \mid zg = gz, \forall\ g \in \mathrm{GL}_n(\mathbb{F})\}\,.$$

Z consists of the scalar matrices. Via the homomorphism $g \mapsto Zg$ we have:

$$\begin{array}{lll} \mathrm{GL}_n(\mathbb{F}) \longrightarrow \frac{\mathrm{GL}_n(\mathbb{F})}{Z} := \mathrm{PGL}_n(\mathbb{F}) & & \text{(projective general linear group)} \\ \Big| \qquad\qquad \Big| & & \\ \mathrm{SL}_n(\mathbb{F}) \longrightarrow \frac{\mathrm{SL}_n(\mathbb{Z})Z}{Z} := \mathrm{PSL}_n(\mathbb{F}) & & \text{(projective special linear group).} \end{array}$$

Note that:

$$\frac{\mathrm{SL}_n(\mathbb{Z})Z}{Z} \cong \frac{\mathrm{SL}_n(\mathbb{F})}{Z \cap \mathrm{SL}_n(\mathbb{F})}.$$

From the above considerations:

$$|\mathrm{PGL}_n(q)| = \frac{|\mathrm{GL}_n(q)|}{q-1} = |\mathrm{SL}_n(q)|\,, \quad |\mathrm{PSL}_n(q)| = \frac{|\mathrm{SL}_n(q)|}{(n, q-1)}. \tag{2.4}$$

Consider the projective space $X := \mathcal{P}(\mathbb{F}^n)$, namely the set of 1-dimensional subspaces of $\mathbb{F}^n$. The group $\mathrm{PGL}_n(\mathbb{F})$ acts on X in a natural way. Indeed, the map

$$\begin{array}{ccc} \varphi : \mathrm{GL}_n(\mathbb{F}) & \longrightarrow & \mathrm{Sym}(X) \\ g & \mapsto & \begin{pmatrix} \langle v \rangle \\ \langle gv \rangle \end{pmatrix} \end{array}$$

is a homomorphism with Kernel $Z = \{\lambda I_n \mid \lambda \in \mathbb{F}^*\}$. It follows that

$$\mathrm{PGL}_n(\mathbb{F}) = \frac{\mathrm{GL}_n(\mathbb{F})}{Z} \ \cong\ \mathrm{Im}\varphi \ \leq\ \mathrm{Sym}(X).$$

So, up to the isomorphism induced by φ:

$$\mathrm{PSL}_n(\mathbb{F}) \leq \mathrm{PGL}_n(\mathbb{F}) \leq \mathrm{Sym}(X).$$

(2.5) Lemma *For $n \geq 2$ the group $\mathrm{PSL}_n(\mathbb{F})$ is a 2-transitive subgroup of $\mathrm{Sym}(X)$.*

Proof Let $\{e_1,\ \ldots,\ e_n\}$ be the canonical basis. Given a pair (v_1, v_2) of linearly independent vectors, there exist $s \in \mathrm{SL}_n(\mathbb{F})$ and $\lambda \in \mathbb{F}$ such that $(se_1, se_2) = (\lambda v_1, v_2)$. Indeed, we may extend $\{v_1, v_2\}$ to a basis $\{v_1, v_2, \ldots, v_n\}$ of $\mathbb{F}^n$ and consider the matrices:

$$b = \left(\ v_1 \ \middle|\ v_2 \ \middle|\ \ldots \ \middle|\ v_n\ \right), \quad s = \left(\ \det\ b^{-1}v_1 \ \middle|\ v_2 \ \middle|\ \ldots \ \middle|\ v_n\ \right).$$

Then $s \in \mathrm{SL}_n(\mathbb{F})$ and $se_1 = \lambda v_1$, with $\lambda = \det\ b^{-1}$, $se_2 = v_2$. It follows

$$(\langle se_1 \rangle\,,\ \langle se_2 \rangle) = (\langle v_1 \rangle\,, \langle v_2 \rangle).$$

By Lemma 1.3 the group $\mathrm{PSL}_2(\mathbb{F})$ is 2-transitive on X. ■

2.2 Root subgroups and the monomial subgroup

(2.6) Lemma *Each of the maps from* $(\mathbb{F}, +)$ *to* $(\mathrm{SL}_2(\mathbb{F}), \cdot)$ *defined by*

$$t \mapsto \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix},$$

is a group monomorphism.

Proof Straightforward calculation. ■

We interpret and generalize this Lemma. As usual we denote by $e_{i,j}$ the $n \times n$ matrix whose entries are all 0, except the entry (i, j) which is 1. Note that $e_{i,i}^2 = e_{ii}$ and $e_{i,j}^2 = 0$ for $i \neq j$. It follows that the map $f_{ij} : (\mathbb{F}, +) \to (\mathrm{SL}_n(\mathbb{F}), \cdot)$ such that, for all $t \in \mathbb{F}$:

$$t \mapsto \exp(te_{ij}) = I + te_{i,j},$$

is a group monomorphism for all $i \neq j$.

(2.7) Definition *For $i \neq j$ the image of f_{ij}, namely the subgroup $\{I + te_{i,j} \mid t \in \mathbb{F}\}$ is called a* root subgroup. *Its elements $I + te_{i,j}$ are called* elementary transvections.

More generally, each of the maps $(\mathbb{F}^{n-1}, +, 0) \to (\mathrm{SL}_n(\mathbb{F}), \cdot, I_n)$ defined by:

$$v \mapsto \begin{pmatrix} 1 & v^T \\ 0 & I_{n-1} \end{pmatrix}, \quad v \mapsto \begin{pmatrix} 1 & 0 \\ v & I_{n-1} \end{pmatrix}, \quad \forall\ v \in \mathbb{F}^{n-1} \tag{2.8}$$

is a group homomorphism. Since the additive group $\mathbb{F}^{n-1}$ is generated by the subgroups $\mathbb{F}e_i$, $1 \leq i \leq n-1$, the images of the maps in (2.8) are generated by elementary transvections.

For $n \geq 3$, every elementary transvection is a commutator. Indeed:

$$(e_{i,j}, e_{j,k}) = e_{i,k} \quad \text{whenever} \quad |\{i, j, k\}| = 3. \tag{2.9}$$

Any matrix whose columns are the vectors of the canonical basis (in some order) is called a *permutation matrix*. The map $\mathrm{Sym}(n) \to \mathrm{GL}_n(\mathbb{F})$ such that

$$\sigma \mapsto \pi_\sigma := \left(\ e_{\sigma(1)} \ \middle| \ \dots \ \middle| \ e_{\sigma(n)} \ \right)$$

is a monomorphism whose image is the group S_n of permutation matrices. For $n \geq 2$, the determinant map $\delta : S_n \to \langle -1 \rangle$ is an epimorphism with kernel $S_n \cap \mathrm{SL}_n(\mathbb{F})$.
If char $\mathbb{F} \neq 2$, then Ker $\delta \cong \mathrm{Alt}(n)$ has index 2 in S_n. If char $\mathbb{F} = 2$, then Ker $\delta = S_n$.
S_n normalizes the group of diagonal matrices $D \simeq (\mathbb{F}^*)^n$. In fact, for all i, j:

$$\pi_\sigma \, e_{i,j} \, \pi_\sigma^{-1} \ = \ e_{\sigma(i),\sigma(j)}. \tag{2.10}$$

(2.11) Definition *The product $M := DS_n$ of the diagonal and permutation subgroups is called the standard* monomial group.

The monomial subgroup M consists of the matrices whose columns are non-zero multiples of the vectors of the canonical basis (in some order). Clearly

$$\frac{M}{D} \cong \mathrm{Sym}(n).$$

(2.12) Lemma *$M \cap \mathrm{SL}_n(\mathbb{F})$ is generated by elementary transvections.*

Proof Suppose first $n = 2$. Then $M = DS_2 = D\left\langle \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle$. By the modular identity:

$$M \cap \mathrm{SL}_2(\mathbb{F}) = (D \cap \mathrm{SL}_2(\mathbb{F})) \left\langle \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \mid \alpha \in \mathbb{F}^* \right\} \left\langle \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\rangle.$$

So the claim is true by the following identities:

(1) $\begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \alpha - 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \alpha^{-1} - 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix}$;

(2) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$.

Then, for $n \geq 2$, the result follows easily. In fact $\mathrm{Sym}(n)$ is generated by transpositions and each matrix $\mathrm{diag}\left(\alpha_1, \ldots, \alpha_{n-1}, \prod_{i=1}^{n-1} \alpha_i^{-1}\right)$ in $D \cap \mathrm{SL}_n(\mathbb{F})$ can be written as

$$\left(\alpha_1, \ldots, 1, \alpha_1^{-1}\right) \ldots \left(1, \ldots, \alpha_{n-1}, \alpha_{n-1}^{-1}\right).$$

■

(2.13) Lemma *The group $\mathrm{SL}_n(\mathbb{F})$ is generated by the elementary transvections.*

Proof Fix $A = (a_{i,j}) \in \mathrm{SL}_n(\mathbb{F})$. We have to show that A is a product of elementary transvections. There exists an entry $a_{h,k} \neq 0$. Let $d = \mathrm{diag}(-1, 1, \ldots, 1)$ and note that, if $h \neq 1$, then $d\pi_{1h} \in M \cap \mathrm{SL}_n(\mathbb{F})$. Similarly, if $k \neq 1$, then $d\pi_{1k} \in M \cap \mathrm{SL}_n(\mathbb{F})$. If $a_{h,k} \neq a_{1,1}$, by Lemma 2.12 we may substitute A with $A' = \pi_{1h} A \pi_{k1}$, or $A' = Ad\pi_{k1}$ or $A' = d\pi_{1h} A$ according to $h \neq 1, k \neq 1$, or $h = 1, k \neq 1$ or $h \neq 1$, $k = 1$. Thus:

$$A' = \begin{pmatrix} \alpha & * \\ * & * \end{pmatrix}, \ \alpha = \pm a_{h,k} \neq 0.$$

Again by Lemma 2.12 we may substitute A' with:

$$A'' = \mathrm{diag}\left(\alpha^{-1}, \alpha, 1, \ldots, 1\right) A' = \begin{pmatrix} 1 & v^T \\ w & B \end{pmatrix}$$

where $v, w \in \mathbb{F}^{n-1},\ B \in \mathrm{SL}_{n-1}(\mathbb{F})$. By (2.8), we may substitute A'' with:

$$\begin{pmatrix} 1 & 0 \\ -w & 1 \end{pmatrix} A'' \begin{pmatrix} 1 & -v^T \\ 0 & I \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & B' \end{pmatrix}, \quad B' \in \mathrm{SL}_{n-1}(\mathbb{F}).$$

The claim now follows by induction on n. ■

2.3 Simplicity and order

(2.14) Theorem $\mathrm{PSL}_n(\mathbb{F})$ *is simple, except when* $n = 2$ *and* $\mathbb{F} = \mathbb{F}_2$ *or* $\mathbb{F} = \mathbb{F}_3$.

Proof $S = \mathrm{PSL}_n(\mathbb{F})$ is a 2-transitive subgroup of $\mathrm{Sym}(X)$ by Lemma 2.5, where $X = \mathcal{P}(\mathbb{F}^n)$ is the projective space. Hence S is a primitive subgroup of $\mathrm{Sym}(X)$ by Lemma 1.4. The preimage in $\mathrm{SL}_n(\mathbb{F})$ of the stabilizer $S_{\langle e_1 \rangle}$, namely the group

$$\left\{ \begin{pmatrix} \det a^{-1} & v^T \\ 0_{\mathbb{F}^{n-1}} & a \end{pmatrix} \mid a \in \mathrm{GL}_{n-1}(\mathbb{F}),\ v \in \mathbb{F}^{n-1} \right\}$$

contains the normal abelian subgroup

$$A := \left\{ \begin{pmatrix} 1 & v^T \\ 0 & I \end{pmatrix} \mid v \in \mathbb{F}^{n-1} \right\}.$$

It follows that the projective image of A is abelian and normal in $S_{\langle e_1 \rangle}$.
The group A is generated by the elementary transvections

$$\{I + tE_{12} \mid t \in \mathbb{F}\},\ \ldots,\ \{I + tE_{1n} \mid t \in \mathbb{F}\}.$$

By (2.10), every elementary transvection $I + te_{i,j}$ is conjugate to $I + tE_{1,2}$ under $DS_n \cap \mathrm{SL}_n(\mathbb{F})$. Thus the conjugates of A generate $\mathrm{SL}_n(\mathbb{F})$ by Lemma 2.13. Hence the conjugates of the projective image of A generate $\mathrm{PSL}_n(\mathbb{F}) = S$.
Finally suppose $|\mathbb{F}| \neq 2, 3$ if $n = 2$. Then $\mathrm{SL}_n(\mathbb{F}) = \mathrm{SL}_n(\mathbb{F})'$, whence $S = S'$: this fact follows from (2.9) for $n \geq 3$, from Lemma 2.12 for $n = 2$.
Our claim is proved in virtue of Iwasawa's Lemma (Theorem 1.7 of this Chapter).
For $|\mathbb{F}| = 2$ and $|\mathbb{F}| = 3$ we have, respectively, $|X| = 3$ and $|X| = 4$. Thus $\mathrm{PSL}_2(2) \leq \mathrm{Sym}(3)$ and $\mathrm{PSL}_2(3) \leq \mathrm{Sym}(4)$ cannot be simple. ■

(2.15) Theorem *When* $\mathbb{F} = \mathbb{F}_q$ *is finite, we have:*

$$|\mathrm{PSL}_n(q)| \ = \ \frac{1}{(n, q-1)}\, q^{\frac{n(n-1)}{2}} (q^2 - 1) \cdots (q^n - 1).$$

Proof The columns of every matrix $\left(\begin{array}{c|c|c} v_1 & \dots & v_n \end{array} \right)$ of $\mathrm{GL}_n(\mathbb{F})$ are a basis of $\mathbb{F}^n$ and, vice versa, the vectors of every basis $\{v_1, \dots, v_n\}$ can be taken as columns of a matrix in $\mathrm{GL}_n(\mathbb{F})$. So $|\mathrm{PSL}_n(q)|$ equals the number of basis of $V = \mathbb{F}_q^n$.
For v_1 one can choose any vector in $V \setminus \{0\}$: here there are $q^n - 1$ choices.
Once v_1 is fixed, v_2 must be chosen in $V \setminus \langle v_1 \rangle$: hence there are $q^n - q$ choices.
Then v_3 must be chosen in $V \setminus \langle v_1, v_2 \rangle$: this gives $q^n - q^2$ choices. And so on... Thus:

$$|\mathrm{GL}_n(q)| = (q^n - 1)(q^n - q)\left(q^n - q^2\right) \dots \left(q^n - q^{n-1}\right) = q^{\frac{n(n-1)}{2}} \prod_{i=1}^{n} \left(q^i - 1\right).$$

The claim follows from (2.4). ■

3 The symplectic groups

By Theorem 4.2 of Chapter II, up to conjugation under $\mathrm{GL}_{2m}(\mathbb{F})$, we may define the *symplectic group* $\mathrm{Sp}_{2m}(\mathbb{F})$ as

$$\mathrm{Sp}_{2m}(\mathbb{F}) = \left\{ g \in \mathrm{GL}_{2m}(\mathbb{F}) \mid g^T \begin{pmatrix} \mathbf{0} & I_m \\ -I_m & \mathbf{0} \end{pmatrix} g = \begin{pmatrix} \mathbf{0} & I_m \\ -I_m & \mathbf{0} \end{pmatrix} \right\}.$$

Direct calculation shows that $\mathrm{Sp}_2(\mathbb{F}) = \mathrm{SL}_2(\mathbb{F})$.

(3.1) Theorem *Let $m \geq 2$. Then:*

(1) $\mathrm{Sp}_{2m}(\mathbb{F})$ *is generated by the following matrices and their transposes:*

$$\begin{pmatrix} I_m + te_{i,j} & 0 \\ 0 & I_m - te_{j,i} \end{pmatrix} \begin{array}{c} 1 \leq i < j \leq m \\ t \in \mathbb{F} \end{array}, \quad \begin{pmatrix} I_m & te_{i,i} \\ 0 & I_m \end{pmatrix} \begin{array}{c} 1 \leq i \leq m \\ t \in \mathbb{F} \end{array};$$

(2) $\mathrm{Sp}_{2m}(\mathbb{F})' = \mathrm{Sp}_{2m}(\mathbb{F})$ *is perfect, except* $\mathrm{Sp}_4(\mathbb{F}_2) \cong \mathrm{Sym}(6)$*;*

(3) *the center of* $\mathrm{Sp}_{2m}(\mathbb{F})$ *is the subgroup generated by* $-I$*.*

In particular $\mathrm{Sp}_{2m}(\mathbb{F}) \leq \mathrm{SL}_{2m}(\mathbb{F})$ by (1).
For the original proof of (1) see [18]. The rest can be proved by direct calculation.

(3.2) Definition *The projective image of* $\mathrm{Sp}_{2m}(\mathbb{F})$*, namely the group*

$$\frac{\mathrm{Sp}_{2m}(\mathbb{F})Z}{Z} \cong \frac{\mathrm{Sp}_{2m}(\mathbb{F})}{\mathrm{Sp}_{2m} \cap Z} = \frac{\mathrm{Sp}_{2m}(\mathbb{F})}{\langle -I \rangle}$$

is called the projective symplectic group *and indicated by* $\mathrm{PSp}_{2m}(\mathbb{F})$*.*

$\mathrm{PSp}_{2m}(\mathbb{F})$, being a subgroup of $\mathrm{PSL}_n(\mathbb{F})$, acts on the projective space $X = \mathcal{P}(\mathbb{F}^n)$. Since all vectors are isotropic, all 1-dimensional subspaces $\langle v\rangle$ and $\langle w\rangle$ are isometric. By Witt's extension Lemma there exists $g \in Sp_{2m}(\mathbb{F})$ such that $\langle gv\rangle = \langle w\rangle$. So $\mathrm{PSp}_{2m}(\mathbb{F})$ is transitive on X. Again by Witt's Lemma , the stabilizer of $\langle v\rangle$ in $\mathrm{PSp}_{2m}(\mathbb{F})$ has 3 orbits on X, namely:

$$\{\langle v\rangle\}, \quad \{\langle w\rangle \mid (v,w) = 0\}, \quad \{\langle w\rangle \mid (v,w) \neq 0\}.$$

Using this information, one can prove the following

(3.3) Lemma $\mathrm{PSp}_{2m}(\mathbb{F})$ *is a primitive subgroup of* $\mathrm{Sym}(X)$*, where* $X = \mathcal{P}(\mathbb{F}^n)$.

(3.4) Theorem *Assume* $m \geq 2$ *and* $\mathbb{F} \neq \mathbb{F}_2$ *when* $m = 2$. *Then* $\mathrm{PSp}_{2m}(\mathbb{F})$ *is simple.*

Proof (sketch) Under our assumptions, the group $S = \mathrm{PSp}_{2m}(\mathbb{F})$ is perfect, by point (2) of Theorem 3.1, and acts primitively on the projective space $X = \mathcal{P}(\mathbb{F}^n)$ by the previous Lemma. In order to apply Iwasawa's Lemma to S, it is convenient to suppose that $Sp_{2m}(\mathbb{F})$ is the group of isometries of

$$J' = \begin{pmatrix} J_1 & \\ & J_2 \end{pmatrix}, \quad \text{where} \quad J_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} \mathbf{0} & I_{m-1} \\ -I_{m-1} & \mathbf{0} \end{pmatrix}$$

The linear preimage of the stabilizer $S_{\langle e_1\rangle}$ of $\langle e_1\rangle$ fixes $\langle e_1\rangle^\perp = \langle e_1, e_3, \ldots e_{2m}\rangle$ and induces the group $\mathrm{Sp}_{2(m-1)}(\mathbb{F})$ on $\frac{\langle e_1\rangle^\perp}{\langle e_1\rangle}$. So it consists of the matrices:

$$(3.5) \qquad \left\{ \begin{pmatrix} \alpha & \beta & \alpha u^T J_2 c \\ 0 & \alpha^{-1} & \mathbf{0}^T \\ \mathbf{0} & u & c \end{pmatrix} \mid 0 \neq \alpha,\ \beta \in \mathbb{F},\ u \in \mathbb{F}^{2m-2},\ c \in \mathrm{Sp}_{2m-2}(\mathbb{F}) \right\}.$$

Noting that

$$\begin{pmatrix} \alpha & \beta & \alpha u^T J_2 c \\ 0 & \alpha^{-1} & \mathbf{0}^T \\ \mathbf{0} & u & c \end{pmatrix}^{-1} = \begin{pmatrix} \alpha^{-1} & -\beta & -u^T J_2 \\ 0 & \alpha & \mathbf{0}^T \\ \mathbf{0} & -\alpha c^{-1} u & c^{-1} \end{pmatrix}$$

it is not difficult to check that the abelian group :

$$A = \left\{ \begin{pmatrix} 1 & \gamma & \mathbf{0}^T \\ 0 & 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{0} & I_{2m-2} \end{pmatrix} \mid \gamma \in \mathbb{F} \right\}$$

is normal in the preimage of $S_{\langle e_1\rangle}$ described by (3.5). One can also show that the conjugates of A generate $\mathrm{Sp}_{2m}(\mathbb{F})$. So the projective image of A is an abelian, normal subgroup of $S_{\langle e_1\rangle}$, whose conjugates generate S. Our claim follows from Theorem 1.7. ∎

(3.6) Theorem $|\mathrm{PSp}_{2m}(q)| = \frac{1}{(2,q-1)}\, q^{m^2}(q^2-1)(q^4-1)\cdots(q^{2m}-1)$.

Proof Each matrix of $Sp_{2m}(q)$ is a basis $\{v_1,\ldots,v_m,v_{-1},\ldots,v_{-m}\}$ of $\mathbb{F}^{2m}$ such that

$$(v_i,v_{-i}) = v_i^T J v_{-i} = 1, \quad (v_i,v_j) = v_i^T J v_j = 0 \quad j \neq -i.$$

$0 \neq v_1$ can be chosen in $(q^{2m}-1)$ ways (as $(v,v)=0$ for all v.
For any fixed v_1, the vector v_{-1} can be chosen in q^{2m-1} ways. Indeed it must satisfy

$$(v_1,v_{-1}) = v_1^T J v_{-1} = 1. \tag{3.7}$$

The space of solutions of the homogeneous equation in $2m$ indeterminates

$$v_1^T J v_{-1} = 0$$

has dimension $2m-1$. Hence the system (3.7) has q^{2m-1} solutions.

$$\mathbb{F}^n = \langle v_1, v_2\rangle \perp \langle v_2,\ldots,v_m,v_{-2},\ldots,v_{-m}\rangle\,.$$

Applying induction to the number of symplectic basis of $\langle v_2,\ldots,v_{-m}\rangle$ we get

$$|Sp_{2m}(q)| = (q^{2m}-1)q^{2m-1}\left(q^{(m-1)^2}(q^2-1)(q^4-1)\cdots(q^{2(m-1)}-1)\right).$$

■

4 The orthogonal groups

Given an orthogonal space (V,Q), with $V=\mathbb{F}^n$, we consider its group of isometries:

$$O_n(\mathbb{F},Q) := \{h \in \mathrm{GL}_n(\mathbb{F}) \mid Q(v) = Q(hv), \quad \forall\, v \in \mathbb{F}^n\}\,. \tag{4.1}$$

Any $h \in O_n(\mathbb{F},Q)$ preserves the non-degenerate symmetric bilinear form

$$(v,w) := Q(v+w) - Q(v) - Q(w), \quad \forall\, v,w \in \mathbb{F}^n. \tag{4.2}$$

Thus, if J denotes the matrix of (4.2) with respect to the canonical basis, we have:

$$h^T J h = J, \quad \forall\, h \in O_n(\mathbb{F},Q). \tag{4.3}$$

It follows, in particular, $(\det h)^2 = 1$, i.e., $\det h = \pm 1$ for all $h \in O_n(\mathbb{F},Q)$.

Suppose first char $\mathbb{F} \neq 2$. By the considerations at the beginning of Section 6.2, the isometries of J are precisely the isometries of Q. So we have the alternative definition:

$$O_n(\mathbb{F}, Q) := \left\{ h \in \mathrm{GL}_n(\mathbb{F}) \mid h^T J h = J \right\}, \quad \text{char } \mathbb{F} \neq 2. \tag{4.4}$$

In $O_n(\mathbb{F}, Q)$ there are matrices of determinant -1, as the reflections defined below. So the group of orthogonal transformations of determinant 1, namely the group

$$SO_n(\mathbb{F}, Q) := O_n(\mathbb{F}, Q) \cap \mathrm{SL}_n(\mathbb{F})$$

has index 2 in $O_n(\mathbb{F}, Q)$.
Now suppose char $\mathbb{F} = 2$. By Lemma 6.13 of Chapter II, we have $n = 2m$ and

$$O_{2m}(\mathbb{F}, Q) = SO_{2m}(\mathbb{F}, Q) \leq \mathrm{Sp}_{2m}(\mathbb{F}). \tag{4.5}$$

(4.6) Definition *For each $w \in \mathbb{F}^n$ with $Q(w) \neq 0$, the* reflection r_w *is the map*

$$v \mapsto v - \frac{(v, w)}{Q(w)}\, w, \quad \forall\ v \in \mathbb{F}^n.$$

It is immediate to see that $r_w \in O_n(\mathbb{F}, Q)$. Moreover:

(4.7) Theorem

(1) *the orthogonal group $O_n(\mathbb{F}, Q)$ is generated by the reflections;*

(2) *the center of $O_n(\mathbb{F}, Q)$ is generated by $-I$.*

But we are more interested in generators of the derived subgroup of $O_n(\mathbb{F}, Q)$, since this is the group whose projective image is generally simple.

(4.8) Definition *$\Omega_n(\mathbb{F}, Q)$ denotes the derived subgroup of $O_n(\mathbb{F}, Q)$ and $P\Omega_n(\mathbb{F}, Q)$ its projective image in* $\mathrm{PGL}_n(\mathbb{F})$.

Clearly $\Omega_n(\mathbb{F}, Q) \leq \mathrm{SO}_n(\mathbb{F}, Q)$. It can also be shown that:

$$|\mathrm{SO}_n(\mathbb{F}, Q) : \Omega_n(\mathbb{F}, Q)| \leq 2.$$

(4.9) Theorem *Let $m \geq 2$. Write $v = \sum_{i=1}^{m} (x_i e_i + x_{-i} e_{-i})$ if $v \in \mathbb{F}^{2m}$,*
$v = x_0 e_0 + \sum_{i=1}^{m} (x_i e_i + x_{-i} e_{-i})$ if $v \in \mathbb{F}^{2m+1}$.

- *If $Q(v) = \sum_{i_1}^{m} x_i x_{-i}$, then $\Omega_n(\mathbb{F}, Q) := \Omega_n^+(\mathbb{F})$ is generated by the following matrices and their transposes:*

$$\begin{pmatrix} I_m + t e_{i,j} & 0 \\ 0 & I_m - t e_{j,i} \end{pmatrix}, \quad \begin{pmatrix} I_m & t\,(e_{i,j} - e_{j,i}) \\ 0 & I_m \end{pmatrix}, \quad t \in \mathbb{F},\ i < j \leq m.$$

- *If* $Q(v) = x_0^2 + \sum_{i_1}^m x_i x_{-i}$ *and char* $\mathbb{F} \neq 2$*, then* $\Omega_n(\mathbb{F}, Q)$ *is generated by the following matrices and their transposes:*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & I_m + te_{j,i} & 0 \\ 0 & 0 & I_m - te_{i,j} \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & -te_i^T \\ 2e_i & I_m & -t^2 e_{i,i} \\ 0 & 0 & I_m \end{pmatrix}, \quad t \in \mathbb{F},\ j < i \leq m.$$

Note that the matrices of the corresponding polar forms are respectively

$$J_{2m} = \begin{pmatrix} 0 & I_m \\ I_m & 0 \end{pmatrix}, \quad J_{2m+1} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & I_m \\ 0 & I_m & 0 \end{pmatrix}.$$

In what follows, let $t^2 + t + \zeta$ be an irreducible polynomial in $\mathbb{F}[t]$, with roots $\alpha \neq \overline{\alpha}$ in

$$\mathbb{K} := \mathbb{F}(\alpha).$$

(4.10) Lemma *Consider the space* $(\mathbb{F}^2, Q_\zeta)$ *with* $Q_\zeta(v) = x_1^2 + x_1 x_{-1} + \zeta x_{-1}^2$ *for each* $v = \begin{pmatrix} x_1 \\ x_{-1} \end{pmatrix}$*. Set* $P = \begin{pmatrix} 1 & -\alpha \\ 1 & -\overline{\alpha} \end{pmatrix}$*. Then*

$$\mathrm{O}_2(\mathbb{F}, Q_\zeta) = P^{-1}\mathrm{O}_2^+(\mathbb{K})P \cap \mathrm{SL}_2(\mathbb{F})$$

where $\mathrm{O}_2^+(\mathbb{K})$ *is the group of isometries of* Q*, with* $Q(v) = x_1 x_{-1}$*.*
In particular, up to conjugation:

- $O_2^+(q) = \left\langle \begin{pmatrix} \beta & 0 \\ 0 & \beta^{-1} \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle$ *with* β *of order* $q-1$*;*
- $O_2^-(q) = \left\langle \begin{pmatrix} \frac{-\overline{\alpha}\gamma + \alpha\gamma^{-1}}{\alpha - \overline{\alpha}} & \frac{\zeta(\gamma - \gamma^{-1})}{\alpha - \overline{\alpha}} \\ \frac{-\gamma + \gamma^{-1}}{\alpha - \overline{\alpha}} & \frac{\alpha\gamma - \overline{\alpha}\gamma^{-1}}{\alpha - \overline{\alpha}} \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \right\rangle$ *with* $\gamma \in \mathbb{F}_{q^2}$ *of order* $q+1$*.*

Proof We pass from the canonical basis $\{e_1, e_2\}$ of $\mathbb{K}^2$ to the basis $\mathcal{B} = \left\{P^{-1}e_1, P^{-1}e_2\right\}$. For any v as in the statement, its coordinate vector $v_\mathcal{B}$ with respect to $\mathcal{B}$ becomes:

$$v_\mathcal{B} = Pv = \begin{pmatrix} x_1 - \alpha x_{-1} \\ x_1 - \overline{\alpha} x_{-1} \end{pmatrix}.$$

With this change of coordinates, the form Q such that $Q(v) = x_1 x_{-1}$ becomes Q_ζ, as:

$$Q(Pv) = (x_1 - \alpha x_{-1})(x_1 - \overline{\alpha} x_{-1}) = x_1^2 + x_1 x_{-1} + \zeta x_{-1}^2 = Q_\zeta(v).$$

Since $\mathrm{O}_2^+(\mathbb{K})$ preserves the quadratic form Q, its conjugate $P^{-1}\mathrm{O}_2^+(\mathbb{K})P$ preserves Q_ζ. Indeed, let $A \in \mathrm{O}_2^+(\mathbb{K})$. Then, for all $v \in \mathbb{K}^2$:

$$Q_\zeta(v) = Q(Pv) = Q(APv) = Q\left(PP^{-1}APv\right) = Q_\zeta\left((P^{-1}AP)v\right).$$

The rest follows by calculation. ■

(4.11) Remark *The space $(\mathbb{F}^2, Q_\zeta)$ is anisotropic, but $(\mathbb{K}^2, Q_\zeta)$ is not, since $t^2+t+\zeta$ is reducible over $\mathbb{K}$. In fact, by the previous Lemma, $(\mathbb{K}^2, Q_\zeta)$ is isometric to $(\mathbb{K}^2, Q)$.*

When $n = 2m$, let $t^2 + t + \zeta = (t - \alpha)(t - \overline{\alpha})$ be as in the Lemma 4.10 and set

$$Q_\zeta = \sum_{i=1}^{m} x_i x_{-i} + x_m^2 + \zeta x_{-m}^2.$$

$\Omega_n(\mathbb{F}, Q_\zeta)$ is a subgroup of a conjugate of $\Omega_n^+(\mathbb{K})$. Indeed, let $S = \mathrm{diag}\,(I_{n-2}, P)$ with P as in Lemma 4.10. then:

$$\Omega_n(\mathbb{F}, Q_\zeta) = S^{-1}\Omega_n^+(\mathbb{K})S \cap \mathrm{SL}_n(\mathbb{F}).$$

Recall that, when $\mathbb{F} = \mathbb{F}_q$ then, up to conjugation:

$$\Omega_n(\mathbb{F}_q, Q_\zeta) = \Omega_n^-(q).$$

For $n \geq 3$ the center of $\Omega_n(\mathbb{F}, Q)$ is $\Omega_n(\mathbb{F}, Q) \cap \langle -I \rangle$. Thus the projective image

$$P\Omega_{2m}^+(\mathbb{F}, Q) := \frac{\Omega_n(\mathbb{F}, Q)}{\Omega_n(\mathbb{F}, Q) \cap \langle -I \rangle}.$$

(4.12) Theorem *The groups $P\Omega_{2m}^+(q)$, $P\Omega_{2m}^-(q)$, for all q and $m \geq 3$, are simple. The groups $P\Omega_{2m+1}(q)$, for q odd and $m \geq 2$, are simple.*

The proof is based on Iwasawa's Lemma, since $P\Omega_{2m}^+(\mathbb{F}, Q)$ is perfect and acts as a primitive group on the set of isotropic 1-dimensional subspaces.

$$\begin{aligned}
|P\Omega_{2m+1}(q)| &= \tfrac{1}{(2,q-1)}\, q^{m^2}(q^2-1)(q^4-1)\cdots(q^{2m}-1) \\
\left|P\Omega_{2m}^+(q)\right| &= \tfrac{1}{(4,q^m-1)}\, q^{m(m-1)}(q^2-1)(q^4-1)\cdots(q^{2m-2}-1)(q^m-1) \\
\left|P\Omega_{2m}^-(q)\right| &= \tfrac{1}{(4,q^m+1)}\, q^{m(m-1)}(q^2-1)(q^4-1)\cdots(q^{2m-2}-1)(q^m+1).
\end{aligned}$$

5 The unitary groups

Let $\mathbb{F}$ have an automorphism σ of order 2 and f be a non-singular hermitian form on $\mathbb{F}^n$ with matrix J with respect to the canonical basis. The unitary group is defined as:

$$\mathrm{GU}_n(\mathbb{F}, f) = \left\{ g \in \mathrm{GL}_n(\mathbb{F}) \mid g^T J g^\sigma = J \right\}.$$

In particular, when $\mathbb{F} = \mathbb{F}_{q^2}$ or $\mathbb{F} = \mathbb{C}$ and σ is the complex conjugation, we may assume $J = I$ by the classification of hermitian form over these fields.
The center Z of $\mathrm{GU}_n(\mathbb{F}, f)$ consists of the scalar matrices αI such that

$$\alpha\alpha^\sigma = 1.$$

In particular the center of $\mathrm{GU}_n(q^2)$ has order $q+1$. (Exercise).

$$\mathrm{SU}_n(\mathbb{F}, f) := \mathrm{GU}_n(\mathbb{F}, f) \cap \mathrm{SL}_n(\mathbb{F}).$$

The projective image of $\mathrm{SU}_n(\mathbb{F}, f)$ in $\mathrm{PGL}_n(\mathbb{F})$, namely the group

$$\mathrm{PSU}_n(\mathbb{F}, f) := \frac{\mathrm{SU}_n(\mathbb{F}, f)Z}{Z} \cong \frac{\mathrm{SU}_n(\mathbb{F}, f)}{\mathrm{SU}_n(\mathbb{F}, f) \cap Z}$$

is called the *projective special unitary group.*

(5.1) Lemma $\mathrm{SL}_2(q) \cong \mathrm{SU}_2(q^2)$.

Proof Let $\gamma \in \mathbb{F}_{q^2}$ be such that $\gamma^{q-1} = -1$. Then $J = \begin{pmatrix} 0 & \gamma \\ -\gamma & 0 \end{pmatrix}$ defines a non-singular hermitian form. Direct calculation shows that, for all $a, b, c, d \in \mathbb{F}_{q^2}$ such that $ad - bc = 1$,

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} J \begin{pmatrix} a^q & b^q \\ c^q & d^q \end{pmatrix} = J \quad \Longleftrightarrow \quad a, b, c, d \in \mathbb{F}_q.$$

■

(5.2) Theorem *For $n \geq 3$ the groups $\mathrm{PSU}_n(\mathbb{F})$ are simple, except when $(n, \mathbb{F}) = (3, \mathbb{F}_4)$.*

Again the proof is based on Iwasawa's Lemma and the primitive action on the set of 1-dimensional isotropic subspaces.
In the finite case:

$$\left|\mathrm{PSU}_n(q^2)\right| = \frac{1}{(n, q+1)}\, q^{\frac{n(n-1)}{2}}(q^2-1)(q^3+1)(q^4-1)\cdots(q^n-(-1)^n).$$

6 The list of finite classical simple groups

Up to isomorphisms, the list is the following:

- $\mathrm{PSL}_n(q) = A_{n-1}(q)$, $n \geq 2$, except $\mathrm{PSL}_2(2) \cong \mathrm{Sym}(3)$, $\mathrm{PSL}_2(3) \cong \mathrm{Alt}(4)$;
- $\mathrm{PSp}_{2m}(q) = C_m(q)$, $m \geq 2$, except $\mathrm{PSp}_4(2) \cong \mathrm{Sym}(6)$;

- $\mathrm{PSp}_4(2)' \cong \mathrm{Alt}(6)$;
- $P\Omega_{2m+1}(q) = B_m(q)$, q odd, $m \geq 2$;
- $P\Omega^+_{2m}(q) = D_m(q)$, $P\Omega^-_{2m}(q) = {}^2D_m(q)$, $m \geq 3$;
- $\mathrm{PSU}_n(q^2) = {}^2A_{n-1}(q)$, $n \geq 3$, except $\mathrm{PSU}_3(4) \cong 3^2.Q_8$.

The lower bounds for n and m above are due to exceptional isomorphisms, such as:

- $\mathrm{SL}_2(q) \cong \mathrm{Sp}_2(q) \cong \mathrm{SU}_2(q^2)$;
- $\Omega^\pm_2(q) \cong C_{\frac{q\mp 1}{(2,q-1)}}$ (cyclic group);
- $P\Omega^+_4(q) \cong \mathrm{PSL}_2(q) \times \mathrm{PSL}_2(q)$;
- $P\Omega^-_4(q) \cong \mathrm{PSL}_2(q^2)$;
- $P\Omega^+_6(q) \cong \mathrm{PSL}_4(q)$;
- $P\Omega^-_6(q) \cong \mathrm{PSU}_4(q^2)$;

7 Exercises

(7.1) Exercise *Let G be a subgroup of $\mathrm{Sym}(X)$, $g \in G$ and $\alpha, \beta \in X$. Show that, if $\beta = g(\alpha)$ then $G_\beta = gG_\alpha g^{-1}$.*

(7.2) Exercise

- *Let N be a normal subgroup of G such that the factor group $\frac{G}{N}$ is abelian. Show that $G' \leq N$.*
- *Let N be a subgroup of G such that $G' \leq N$. Show that N is normal and $\frac{G}{N}$ is abelian.*

(7.3) Exercise *Assuming $\alpha\beta\gamma = 1$, write $\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$ and $\begin{pmatrix} 0 & \alpha & 0 \\ 0 & 0 & \beta \\ \gamma & 0 & 0 \end{pmatrix}$ as products of elementary transvections.*

(7.4) Exercise *Show that the map $\left(\mathbb{F}^2, +, 0\right) \to (\mathrm{SL}_3(\mathbb{F}), \cdot, I)$ defined by:*

$$\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ t_1 & 1 & 0 \\ t_2 & 0 & 1 \end{pmatrix}$$

is a homomorphism of groups. Write the matrix on the right (and its transpose) as a product of elementary transvections.

(7.5) Exercise *Show that* $\mathrm{SL}_2(\mathbb{F}) = \mathrm{SL}_2(\mathbb{F})'$ *except when* $|\mathbb{F}| = 2, 3$.

(7.6) Exercise *Show that the center* Z *of* $\mathrm{SL}_n(\mathbb{F})$ *consists of scalar matrices.*

(7.7) Exercise *Show that:* $|Z \cap \mathrm{SL}_n(q)| = (n, q-1)$.

(7.8) Exercise *Show that any matrix* $m \in \mathrm{Mat}_n(\mathbb{F})$ *is conjugate to its transpose. (***Hint***: start from a companion matrix) and deduce that:*

- *any symplectic transformation* $g \in \mathrm{Sp}_{2m}(\mathbb{F})$ *is conjugate to* g^{-1} *under* $\mathrm{GL}_{2m}(\mathbb{F})$*;*
- *any orthogonal transformation* $g \in O_n(\mathbb{F}, Q)$ *is conjugate to* g^{-1} *under* $\mathrm{GL}_n(\mathbb{F})$.

(7.9) Exercise *Let* $\mathbb{F}^n$ *be an orthogonal space with respect to* Q*. Show that, for every* $0 \neq w \in \mathbb{F}^n$ *the reflection* r_w *is a linear transformation of determinant* -1*, and an isometry of* Q*. Write the matrix of* r_w *with respect to a basis* $w, w_2, \ldots, w_n$ *where* $w_2, \ldots, w_n$ *is a basis of* $\langle w \rangle^\perp$.

Chapter IV

Some facts from representation theory

This deep and important theory cannot be developed in these notes. We just give some basic results and refer, for a systematic exposition, to books like [6], [7], [13].

1 Irreducible and indecomposable modules

We consider the space $\mathbb{F}^n$ of column vectors as a *left module* over the ring $\mathrm{Mat}_n(\mathbb{F})$ with respect to the usual product of matrices. Let A be a subset of $\mathrm{Mat}_n(\mathbb{F})$.

(1.1) Definition *A subspace W of $\mathbb{F}^n$ is* A-invariant *if $AW \leq W$, i.e., if:*

$$aw \in W, \quad \forall\ a \in A,\ \forall\ w \in W.$$

Clearly W is A-invariant if and only if it is $\mathbb{F}A$-invariant, where $\mathbb{F}A$ denotes the linear subspace of $\mathrm{Mat}_n(\mathbb{F})$ generated by A. Moreover, when A is a subring of $\mathrm{Mat}_n(\mathbb{F})$, then W is A-invariant if and only if it is a module over A.

(1.2) Lemma *Let $\mathbb{F} \leq \mathbb{K}$, a field extension. If $w_1, \ldots, w_m$ are linearly independent vectors of $\mathbb{F}^n$, then they are linearly independent in $\mathbb{K}^n$.*

Proof There exists $P \in \mathrm{GL}_n(\mathbb{F})$ such that $Pw_j = e_j$, $1 \leq j \leq m$. So assume $\sum_{i=1}^m k_i w_i = 0$, with $k_i \in \mathbb{K}$. Multiplying by P we get $\sum_{i=1}^m k_i e_i = 0$, whence $k_1 = \cdots = k_m = 0$. ∎

A subspace W of $\mathbb{F}^n$ can be extended to the subspace $W \otimes_{\mathbb{F}} \mathbb{K}$ of $\mathbb{K}^n$ defined as the subspace of $\mathbb{K}^n$ generated by any basis $\mathcal{B} = \{w_1, \ldots, w_m\}$ of W, namely:

$$W \otimes_{\mathbb{F}} \mathbb{K} \ = \ \left\{ \sum_{j=1}^m k_j w_j \mid k_j \in \mathbb{K} \right\} \quad \text{(tensor product)}.$$

$\mathcal{B}$ is a basis of $W \otimes_{\mathbb{F}} \mathbb{K}$, by Lemma 1.2. Thus, if W is an A-module, also $W \otimes_{\mathbb{F}} \mathbb{K}$ becomes an A-module via the action:

$$a \sum_{j=1}^{m} k_j w_j = \sum_{j=1}^{m} k_j \, aw_j \quad \forall \; a \in A.$$

(1.3) Definition *Let A be a subring (or a subgroup) of* $\mathrm{Mat}_n(\mathbb{F})$ *and W be an A-invariant subspace of $\mathbb{F}^n$. The* A-module W *is said to be:*

(1) indecomposable*, if there is no decomposition $W = W_1 \oplus W_2$ into proper A-invariant subspaces W_1, W_2;*

(2) irreducible*, if the only A-invariant subspaces of W are $\{0_{\mathbb{F}^n}\}$ and W;*

(3) absolutely irreducible*, if $W \otimes_{\mathbb{F}} \mathbb{K}$ is irreducible for any field extension $\mathbb{K}$ of $\mathbb{F}$.*

Accordingly, a subring (or a subgroup) A of $\mathrm{Mat}_n(\mathbb{F})$ is said to be:

- *indecomposable*, if $\mathbb{F}^n$ is indecomposable as an A-module;
- *irreducible*, if $\mathbb{F}^n$ is irreducible as an A-module;
- *absolutely irreducible*, if $\mathbb{F}^n$ is absolutely irreducible as an A-module.

Clearly an irreducible group is indecomposable. The converse is not true in general, as shown in Example 1.5 below. It is true when G is finite and $\mathbb{F}$ has characteristic p where $p = 0$ or p does not divide $|G|$ (see Theorem 1.11).

(1.4) Example *The subgroup G of* $\mathrm{GL}_2(\mathbb{R})$*, generated by the matrix* $g = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$*, is irreducible but not absolutely irreducible.*

Indeed g has no eigenvalue in $\mathbb{R}$. Thus $\mathbb{R}^2$ has no 1-dimensional G-submodule. But g has eigenvalues in $\mathbb{C}$. Thus, for example, $\left\langle \begin{pmatrix} 1 \\ i \end{pmatrix} \right\rangle$ is G-invariant in $\mathbb{C}^2$.

(1.5) Example *The subgroup* $G = \left\{ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \mid t \in \mathbb{F} \right\}$ *of* $\mathrm{GL}_2(\mathbb{F})$ *is reducible, but indecomposable, for any field $\mathbb{F}$.*

G is reducible because $\langle e_1 \rangle$ is G-invariant. Suppose $\mathbb{R}^2 = \langle v_1 \rangle \oplus \langle v_2 \rangle$ where each $\langle v_i \rangle$ is G-invariant. Then v_1, v_2 should be a basis of eigenvectors of G. Since every $g \in G$ has only the eigenvalue 1, one gets $Gv_1 = v_1$, $Gv_2 = v_2$, whence the contradiction $G = I$.

(1.6) Example *The subgroup G of $\mathrm{GL}_2(\mathbb{R})$, generated by the matrices*

$$g_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ g_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

is absolutely irreducible.

Indeed the only 1-dimensional g_1-invariant subspaces are its eigenspaces, namely $\langle e_1 \rangle$ and $\langle e_2 \rangle$, but they are not g_2-invariant.

(1.7) Lemma $\mathrm{Mat}_n(\mathbb{F})$ *is absolutely irreducible for any field $\mathbb{F}$. Moreover its center Z coincides with the field $\mathbb{F}I_n$ of scalar matrices.*

Proof Set $A = \mathrm{Mat}_n(\mathbb{F})$ and let $\{0\} \neq W$ be an A-invariant subspace of $\mathbb{K}^n$, where $\mathbb{K}$ is a field extension of $\mathbb{F}$. Take $0 \neq w \in W$. Then there exists a non-zero component α_i of w. From $e_{i,i} \in A$, it follows that $e_{i,i}w = \alpha_i e_i \in W$. Hence $e_i \in W$. Considering in A the permutation matrices $\pi_{(i,j)}$ we get that $\pi_{(i,j)}e_i = e_j \in W$ for $1 \leq j \leq n$. So W contains the canonical basis, whence $W = \mathbb{K}^n$.
By direct calculation one sees that a matrix commutes with all matrices $e_{ij} \in \mathrm{Mat}_n(\mathbb{F})$ if and only if it is scalar. ■

(1.8) Theorem *Let G be one of the following classical groups:*

$$\mathrm{SL}_n(\mathbb{F}),\ \mathrm{SU}_n(\mathbb{F}),\ \mathrm{Sp}_n(\mathbb{F}),\ n = 2m,\ \Omega_n(\mathbb{F}, Q).$$

Then $\mathbb{F}G = \mathrm{Mat}_n(\mathbb{F})$, except when $G = \Omega_2(\mathbb{F}, Q)$. In particular G is absolutely irreducible and its centralizer in $\mathrm{Mat}_n(\mathbb{F})$ consists of the scalar matrices.

Proof One can see that in each case, provided $G \neq \Omega_2(\mathbb{F}, Q)$, the group G contains n^2 linearly independent matrices (for instance the generators of these groups given in the previous Chapter). Hence the subspace $\mathbb{F}G$ generated by G coincides with $\mathrm{Mat}_n(\mathbb{F})$, which is absolutely irreducible. ■

(1.9) Lemma *(Schur's Lemma) Let $A \leq \mathrm{Mat}_n(\mathbb{F})$ be irreducible. Then*

$$C = C_{\mathrm{Mat}_n(\mathbb{F})}(A) := \{c \in \mathrm{Mat}_n(\mathbb{F}) \mid ca = ac,\ \forall\ a \in A\}$$

is a division algebra over $\mathbb{F}I_n$. In particular, if commutative, C is a field.

Proof It is easy to see that C is a subalgebra of $\mathrm{Mat}_n(\mathbb{F})$, which contains $Z = \mathbb{F}I_n$. Consider a non-zero matrix $c \in C$. The subspace $c\mathbb{F}^n$ is A-invariant, as:

$$a\,(c\mathbb{F}^n) = (ac)\mathbb{F}^n = (ca)\mathbb{F}^n = c(a\mathbb{F}^n) \leq c\mathbb{F}^n, \quad \forall\ a \in A.$$

$0_{\mathrm{Mat}_n(\mathbb{F})} \neq c \Longrightarrow c\mathbb{F}^n \neq \{0_{\mathbb{F}^n}\}$. It follows $c\mathbb{F}^n = \mathbb{F}^n$, by the irreducibility of A. Since the multiplication by c is surjective, it is injective. Thus c has inverse c^{-1}. Clearly $c^{-1} \in C$. ■

Up to here we considered the natural $\mathrm{Mat}_n(\mathbb{F})$-module $\mathbb{F}^n$. But we may also consider the left regular module ${}_{\mathrm{Mat}_n(\mathbb{F})}\mathrm{Mat}_n(\mathbb{F})$ and compare these two modules.

(1.10) Lemma *Let A be a subring of $\mathrm{Mat}_n(\mathbb{F})$, acting irreducibly on $\mathbb{F}^n$, and let $\{0\} \neq W \leq \mathrm{Mat}_n(\mathbb{F})$ be a minimal A-invariant subspace, in the regular action of $\mathrm{Mat}_n(\mathbb{F})$ on itself. Then there exists a vector e_i of the canonical basis such that $\mathbb{F}^n = We_i$. Moreover W is isomorphic to $\mathbb{F}^n$, as an A-module. In particular $dim_{\mathbb{F}}W = n$.*

Proof Choose $0 \neq w \in W$. Then w has a non-zero column we_i. The subspace We_i of $\mathbb{F}^n$ is such that $A(We_i) \leq We_i$. From $we_i \in We_i$ it follows $We_i \neq \{0\}$. Hence $We_i = \mathbb{F}^n$, by the irreducibility of A. Finally, the map $f : W \to \mathbb{F}^n$ defined by $w \mapsto we_i$ is an $\mathbb{F}$-isomorphism such that $f(aw) = af(w)$ for all $a \in A$. ■

(1.11) Theorem *(Maschke) Let $G \leq \mathrm{GL}_n(\mathbb{F})$ be a finite group, where $\mathbb{F}$ has characteristic 0 or a prime p which does not divide $|G|$. Then every G-invariant subspace W of $\mathbb{F}^n$ has a G-invariant complement.*

Proof Let $\mathbb{F}^n = W \oplus U$, where U is an $\mathbb{F}$-complement of W, and call $\pi : \mathbb{F}^n \to U$ the projection. Consider $\psi : \mathbb{F}^n \to \mathbb{F}^n$ defined by:

$$\psi(v) := \frac{1}{|G|}\sum_{x\in G} x^{-1}\pi(xv), \quad \forall\ v \in \mathbb{F}^n.$$

The image of ψ, namely $\psi(\mathbb{F}^n)$, is G-invariant, since for all $g \in G$ and $v \in \mathbb{F}^n$:

$$\psi(gv) := \frac{1}{|G|}\sum_{x\in G} x^{-1}\pi(xgv) = \frac{1}{|G|}g\sum_{x\in G}\left(g^{-1}x^{-1}\right)\pi(xgv) = g\psi(v).$$

Moreover, from $u - \pi(u) \in W$ for all $u \in \mathbb{F}^n$, it follows that:

$$v - \psi(v) = \frac{1}{|G|}\sum_{x\in G} x^{-1}xv - \frac{1}{|G|}\sum_{x\in G} x^{-1}\pi(xv) = \frac{1}{|G|}\sum_{x\in G} x^{-1}\left(xv - \pi(xv)\right) \in W.$$

Thus $v = (v - \psi(v)) + \psi(v)$ for all $v \in \mathbb{F}^n$, gives $\mathbb{F}^n = W + \psi(\mathbb{F}^n)$.
For all $w \in W$ and all $x \in G$ we have $\pi(xw) = 0$. So $\psi(w) = 0$, whence $\psi(v - \psi(v)) = 0$, for all v. This gives $\psi^2 = \psi$ and $W \cap \psi(\mathbb{F}^n) = \{0\}$. Indeed, from $w = \psi(v) \in W \cap \psi(V)$, we have $\psi(v) = \psi^2(v) = \psi(w) = 0$.
We conclude that $\psi(\mathbb{F}^n)$ is a G-invariant complement of W in $\mathbb{F}^n$. ■

2 Representations of groups

(2.1) Definition *Let H be an abstract group.*

(1) *A* representation *of H of degree n over $\mathbb{F}$ is a homomorphism $f : H \to \mathrm{GL}_n(\mathbb{F})$. The representation f is said to be* irreducible *if $\mathbb{F}^n$ is an irreducible $f(H)$-module.*

(2) *The* character χ *of f is the map $\chi : H \to \mathbb{F}$ such that*

$$\chi(h) := \mathrm{tr}(f(h)), \quad \forall\ h \in H.$$

(3) *Two representations $f_i : H \to \mathrm{GL}_n(\mathbb{F})$, $i = 1, 2$ are said to be* equivalent *if there exists $P \in \mathrm{GL}_n(\mathbb{F})$ such that*

$$Pf_1(h) = f_2(h)P, \quad \forall\ h \in H. \tag{2.2}$$

Since conjugate matrices have the same trace, equivalent representations have the same characters.

(2.3) Definition *Let H be an abstract group. The* group algebra $\mathbb{F}H$ *is defined as follows. The elements of H are a basis of $\mathbb{F}H$ as a vector space over $\mathbb{F}$. The product in $\mathbb{F}H$ is the extension, by linearity, of the product in H.*

In particular, by definition, the elements of $\mathbb{F}H$ are the formal linear combinations

$$\sum_{h \in H} \alpha_h h, \quad \alpha_h \in \mathbb{F}$$

with a finite number of non-zero coefficients. By definition, $\dim_{\mathbb{F}} \mathbb{F}H = |H|$.
The extension to $\mathbb{F}H$, by linearity, of any representation $f : H \to \mathrm{GL}_n(\mathbb{F})$ gives rise to an algebra homomorphism $f : \mathbb{F}H \to \mathrm{Mat}_n(\mathbb{F})$. Vice versa, if $f : \mathbb{F}H \to \mathrm{Mat}_n(\mathbb{F})$ is an algebra homomorphism, its restriction $f_H : H \to \mathrm{GL}_n(\mathbb{F})$ is a representation of H.

(2.4) Remark *If $f : H \to \mathrm{GL}_n(\mathbb{F})$ is a representation, then $\mathbb{F}^n$ is an H-module with respect to $hv := f(h)v$, for all $v \in \mathbb{F}^n$. Vice versa, if $\mathbb{F}^n$ is an $\mathbb{F}H$-module, the map $f : H \to \mathrm{GL}_n(\mathbb{F})$ such that $f(h) = \left(\, he_1 \, \middle| \, \ldots \, \middle| \, he_n \, \right)$ is a representation.*

(2.5) Lemma *Two representations $f_1 : H \to \mathrm{GL}_n(\mathbb{F})$ and $f_2 : H \to \mathrm{GL}_n(\mathbb{F})$ are equivalent if and only if the corresponding $\mathbb{F}H$-modules $V_i = \mathbb{F}^n$ are isomorphic, $i = 1, 2$.*

Proof Suppose first that f_1 and f_2 equivalent and let $P \in \mathrm{GL}_n(\mathbb{F})$ be as in point (3) of Definition 2.1. Then the the multiplication by P, namely the map $\mu_P : V_1 \to V_2$, is an $\mathbb{F}H$-isomorphism. Indeed μ_P is $\mathbb{F}$-linear and, for all $v \in \mathbb{F}^n$ and all $h \in H$:

$$\mu_P(f_1(h)v) = Pf_1(h)v = f_2(h)Pv = f_2(h)\mu_P(v).$$

Vice versa, if there exists an $\mathbb{F}H$-isomorphism $\sigma : V_1 \to V_2$ and $P \in \mathrm{GL}_n(\mathbb{F})$ is the matrix of σ with respect to the canonical basis, then $Pf_1(h) = f_2(h)P$ for all $h \in H$. Thus f_1 and f_2 are equivalent. ■

Given two representations $f_i : H \to \mathrm{GL}_{n_i}(\mathbb{F})$, $i = 1, 2$, we may consider their *sum*, namely the representation $f : H \to \mathrm{GL}_{n_1+n_2}(\mathbb{F})$, defined by:

$$f(h) := \begin{pmatrix} f_1(h) & 0 \\ 0 & f_2(h) \end{pmatrix}, \ \forall \ h \in H.$$

Set $M_i = \mathrm{Mat}_{n_i}(\mathbb{F})$. Clearly the subspace

$$M_1 \oplus M_2 := \left\{ \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \mid A_1 \in \mathrm{Mat}_{n_1}(\mathbb{F}), \ A_2 \in \mathrm{Mat}_{n_2}(\mathbb{F}) \right\}$$

is an $f(H)$-module. Moreover the projections

$$\pi_i : M_1 \oplus M_2 \to \mathrm{Mat}_{n_i}(\mathbb{F})$$

are $f(H)$-homomorphisms. In particular $f(H)\,\mathrm{Ker}\ \pi_i = \mathrm{Ker}\ \pi_i$, for $i = 1, 2$.

(2.6) Lemma *In the above notation, suppose that the representations*

$$f_i : H \to \mathrm{GL}_{n_i}(\mathbb{F}), \quad i = 1, 2$$

are irreducible and inequivalent. Let $0 \neq M$ be a minimal subspace of $M_1 \oplus M_2$ such that $f(H)M = M$. Then either $\pi_1(M) = 0$ or $\pi_2(M) = 0$.

Proof Suppose, by contradiction, $M \not\leq \mathrm{Ker}\ \pi_i$, for $i = 1, 2$. It follows that the $f(H)$-module $\mathrm{Ker}\ \pi_i \cap M$ is zero, $i = 1, 2$, by the minimality of M. Thus the restrictions

$$\pi_{i|M} : M \to \pi_i(M), \quad i = 1, 2$$

are $\mathbb{F}$-isomorphisms. In particular $n_1 = n_2 = \dim_{\mathbb{F}} M$. Again by the minimality of M, each $\pi_i(M)$ is a minimal $f_i(H)$-submodule of $\mathrm{Mat}_{n_i}(\mathbb{F})$. It follows from Lemma 1.10 of this Chapter, with $A = f_i(H)$, $W = \pi_i(M)$, that there exist $f_i(H)$ isomorphisms $\tau_i : \pi_i(M) \cong \mathbb{F}^{n_i}$, $i = 1, 2$. Thus $\tau_2\tau_1^{-1} : \mathbb{F}^{n_1} \to \mathbb{F}^{n_2}$ is a isomorphism of the $f_1(H)$-module $\mathbb{F}^{n_1}$ onto the $f_2(H)$-module $\mathbb{F}^{n_2}$, a contradiction. ■

Note that, if G is a group and V is a G-module, then $GW \leq W$ if and only if $GW = W$, for any subspace W of V. Indeed $W = 1_G W \leq GW$.

(2.7) Theorem *Let $f_i : H \to \mathrm{GL}_{n_i}(\mathbb{F})$ be irreducible pairwise inequivalent representations of a group H, with $\mathbb{F}$ algebraically closed. Suppose that $m_i \in \mathrm{Mat}_{n_i}(\mathbb{F})$, $1 \leq i \leq s$, are such that*

$$\sum_{i=1}^{s} \mathrm{tr}\,(m_i f_i(h)) = 0_{\mathbb{F}},\ \forall\ h \in H.$$

Then each $m_i = 0_{\mathrm{Mat}_{n_i}(\mathbb{F})}$, for $i = 1, \ldots, s$.

Proof Induction on s. Suppose $s = 1$ and put $n = n_1$, $f = f_1$. The set

$$M = \{m \in \mathrm{Mat}_n(\mathbb{F}) \mid \mathrm{tr}(mf(h)) = 0,\ \forall\ h \in H\}$$

is a subspace. Moreover $f(H)M = M$ since for all $h_1, h \in H$, $m \in M$:

$$\mathrm{tr}(f(h_1)m\, f(h)) = \mathrm{tr}(f(h)\, f(h_1)m) = \mathrm{tr}(f(hh_1)m) = 0.$$

We want to show that $M = \{0_{\mathrm{Mat}_n(\mathbb{F})}\}$. If this is false, we may choose a non-zero subspace U of M of minimal dimension with respect to the property $f(H)U = U$. By Lemma 1.10 we have $\dim U = n$ and $Uv = \mathbb{F}^n$ for some v. If $\{u_1, \ldots u_n\}$ is a basis of U, then $\{u_1 v, \ldots u_n v\}$ is a basis of $\mathbb{F}^n$. Up to conjugation we may suppose that

$$\{u_1 v, \ldots u_n v\} = \{e_1, \ldots, e_n\} \quad \text{(canonical basis)}.$$

For all $w \in \mathbb{F}^n$ we consider the matrix A_w with columns $A_w e_i = u_i w$, i.e.,:

$$A_w = \left(\ u_1 w \ \middle|\ \ldots\ \middle|\ u_n w\ \right).$$

Let λ_w be an eigenvalue of A_w, with eigenvector $\sum_{i=1}^{n} \rho_i e_i \neq 0_{\mathbb{F}^n}$. Then:

$$0_{\mathbb{F}^n} = (A_w - \lambda_w I) \sum_{i=1}^{n} \rho_i e_i =$$

$$\sum_{i=1}^{n} \rho_i \left(A_w - \lambda_w I\right) e_i = \sum_{i=1}^{n} \rho_i \left(u_i w - \lambda_w u_i v\right) = \sum_{i=1}^{n} \rho_i u_i \left(w - \lambda_w v\right).$$

It follows that the vectors

$$u_1(w-\lambda_w v),\ \ldots,\ u_n(w-\lambda_w v)$$

are linearly dependent. Hence the space $U(w-\lambda_w v)$, generated by them, has dimension less then n. Since it is $f(H)$-invariant, the irreducibility of $\mathbb{F}^n$ gives:

$$U(w-\lambda_w v)=\{0_{\mathbb{F}^n}\},\forall\ w\in\mathbb{F}^n.$$

In particular $u_i(e_j-\lambda_{e_j}v)=0_{\mathbb{F}^n}$ for all i,j. Thus, setting $\lambda_{e_j}=\lambda_j$:

$$u_ie_j=\lambda_j u_iv=\lambda_j e_i,\quad 1\leq i,j\leq n. \tag{2.8}$$

This tells us:

$$u_i=\left(\ \lambda_1 e_i\ \middle|\ \ldots\ \middle|\ \lambda_n e_i\ \right),\quad 1\leq i\leq n.$$

$$0=\operatorname{tr}(u_i\,\mathrm{id}_G)=\operatorname{tr}(u_i)=\lambda_i,\quad 1\leq i\leq n.$$

And now (2.8) gives that u_i has all columns equal to zero, hence $u_i=0_{\mathbb{F}^n}$ for all i-s, against the assumption that $u_1,\ldots,u_n$ are linearly independent. We conclude $M=\{0_{\mathrm{Mat}_n(\mathbb{F})}\}$ and the first step of induction is proved.

Now suppose $s>1$. Set $n=\sum_{i=1}^s n_i$ and consider the sum $f:H\to\mathrm{GL}_n(\mathbb{F})$ of the representations f_i, defined by:

$$f(h):=\begin{pmatrix} f_1(h) & & \\ & \ldots & \\ & & f_s(h)\end{pmatrix},\ \forall\ h\in H.$$

Let M be the following subset of $\mathrm{Mat}_{n_1}(\mathbb{F})\oplus\cdots\oplus\mathrm{Mat}_{n_s}(\mathbb{F})$:

$$M:=\left\{m=\begin{pmatrix} c_1 & & \\ & \ldots & \\ & & c_s\end{pmatrix}\ |\ \operatorname{tr}(mf(h))=\sum_{i=1}^{s}\operatorname{tr}(c_if_i(h))=0,\quad\forall\ h\in H\right\}.$$

Clearly M is an $f(H)$-invariant subspace and we want to show that $M=\{0_{\mathrm{Mat}_n(\mathbb{F})}\}$. If this is false, we may choose a non-zero subspace U of M of minimal dimension with respect to the property $f(H)U=U$. By the assumption that the representations $f_i:H\to\mathrm{GL}_{n_i}(\mathbb{F})$ are irreducible and pairwise inequivalent, Lemma 2.6 tells us that $\pi_i(U)=0_{\mathrm{Mat}_{n_i}(\mathbb{F})}$ for at least one i. We may suppose $i=1$. This means that, for all

$$\begin{pmatrix} u_1 & & \\ & \ldots & \\ & & u_s\end{pmatrix}\in U$$

we have $u_1 = 0_{\mathrm{Mat}_{n_1}(\mathbb{F})}$. It follows

$$0_{\mathbb{F}} = \sum_{i=1}^{s} \mathrm{tr}\left(u_i f_i(h)\right) = \sum_{i=2}^{s} \mathrm{tr}\left(u_i f_i(h)\right), \quad \forall\ h \in H.$$

By induction $u_2 = \cdots = u_s = 0$, whence $U = \left\{0_{\mathrm{Mat}_n(\mathbb{F})}\right\}$, a contradiction. ■

(2.9) Corollary *Let $f_i : G \to \mathrm{GL}_{n_i}(\mathbb{F})$, $i \leq s$, be pairwise inequivalent, absolutely irreducible representations of a group G with k conjugacy classes. Then $s \leq k$.*

Proof We may suppose $\mathbb{F}$ algebraically closed. Choose representatives $g_1, \ldots, g_k$ of the conjugacy classes of G and consider the s vectors of $\mathbb{F}^k$:

$$v_1 = \begin{pmatrix} \mathrm{tr}\left(f_1(g_1)\right) \\ \cdots \\ \mathrm{tr}\left(f_1(g_k)\right) \end{pmatrix}, \ \ldots\ , \quad v_s = \begin{pmatrix} \mathrm{tr}\left(f_s(g_1)\right) \\ \cdots \\ \mathrm{tr}\left(f_s(g_k)\right) \end{pmatrix}.$$

Suppose $\sum_{i=1}^{s} \alpha_i v_i = 0_{\mathbb{F}^n}$ for some $\alpha_i \in \mathbb{F}$. It follows

$$\sum_{i=1}^{s} \alpha_i \,\mathrm{tr}\left(f_i(g_j)\right) = \sum_{i=1}^{s} \mathrm{tr}\left(\alpha_i f_i(g_j)\right) = 0_{\mathbb{F}}, \quad 1 \leq j \leq k.$$

Every $g \in G$ is conjugate to a g_j and $\mathrm{tr}(g) = \mathrm{tr}(g_j)$. Thus:

$$\sum_{i=1}^{s} \mathrm{tr}\left(\alpha_i\, f_i(g)\right) = 0_{\mathbb{F}}, \quad \forall\ g \in G.$$

By the previous Theorem $\alpha_i = 0$ for all $i \leq s$. This means that the vectors $v_1, \ldots, v_s$ are linearly independent in $\mathbb{F}^k$. We conclude $s \leq k$. ■

(2.10) Theorem *Let G be a subgroup of $\mathrm{GL}_n(\mathbb{F})$ and denote by $\mathbb{F}G$ the linear subspace of $\mathrm{Mat}_n(\mathbb{F})$ generated by G. The following conditions are equivalent:*

(1) *G is absolutely irreducible;*

(2) *$\mathbb{F}G = \mathrm{Mat}_n(\mathbb{F})$ (equivalently, $\dim_{\mathbb{F}} \mathbb{F}G = n^2$);*

(3) *G is irreducible and $C_{\mathrm{Mat}_n(\mathbb{F})}(G) = \mathbb{F}I_n$.*

Proof

(1) $\Longrightarrow$ (2) Substituting $\mathbb{F}$ with its algebraic closure, if necessary, we may suppose $\mathbb{F}$ algebraically closed. Let $g_1, \ldots, g_m$ be a basis of $\mathbb{F}G$ and consider the orthogonal space

$\mathbb{F}G^{\perp}$ with respect to the bilinear form $(g_1, g_2) = \mathrm{tr}(g_1 g_2)$ (see (3.7) in the Exercises of this Chapter). Since this form is non-degenerate, $FG^{\perp}$ has dimension $n^2 - m$. Thus, if $m < n^2$, there exists a non-zero matrix m such that $\mathrm{tr}\, mg = 0$ for all $g \in G$, in contrast with Theorem 2.7.

(2) $\Longrightarrow$ (3) Any G-invariant subspace would be $\mathrm{Mat}_n(\mathbb{F})$- invariant, against the irreducibility of $\mathrm{Mat}_n(\mathbb{F})$. The last claim follows from the fact that the center of $\mathrm{Mat}_n(\mathbb{F})$ consists of scalar matrices.

(3) $\Longrightarrow$ (1) [6, Theorem 29.13]. ■

Point (1) of the following Lemma explains why, in the study of classical groups, one is interested in the groups of isometries of non-degenerate forms. By point (2) an absolutely irreducible group can fix at most one form, necessarily non-degenerate, up to scalars.

(2.11) Lemma *Let $J \in \mathrm{Mat}_n(\mathbb{F})$ be such that $J^T = J^{\sigma}$ (σ a field automorphism), or $J^T = -J$ and let $G \leq \mathrm{GL}_n(\mathbb{F})$ be a group of isometries of J, namely*

$$g^T J g^{\sigma} = J, \quad \forall\ g \in G.$$

(1) *if $\det J = 0$, then G is reducible;*

(2) *if G is absolutely irreducible, and J' is such that $g^T J' g^{\sigma} = J'$ for all $g \in G$, then $J' = \lambda J$ for some $0 \neq \lambda \in \mathbb{F}$.*

Proof

(1) The 0-eigenspace W of J is non-zero. W is G^{σ}-invariant, since:

$$Jg^{\sigma}w = \left(g^{-1}\right)^T Jw = 0 \implies g^{\sigma}w \in W, \quad \forall\ g \in G,\ w \in W.$$

From the fact that W is G^{σ}-invariant, it follows that $W^{\sigma^{-1}}$ is G-invariant.

(2) $Jg^{\sigma}J^{-1} = \left(g^{-1}\right)^T = J' g^{\sigma} J'^{-1}$ for all g gives $J'^{-1}J \in C_{\mathrm{Mat}_n(\mathbb{F})}\left(G^{\sigma}\right) = \mathbb{F}I_n$. ■

3 Exercises

(3.1) Exercise *Let $\mathbb{K}$ be a field extension of $\mathbb{F}$. Show that any subset $\{w_1, \ldots, w_m\}$ of $\mathbb{F}^n$ which is linearly independent over $\mathbb{F}$ is also linearly independent over $\mathbb{K}$.*

(3.2) Exercise *Let $G = \mathrm{GL}_4(\mathbb{F})$ and $W = \langle e_1, e_2 \rangle \leq \mathbb{F}^4$. Determine:*

i) *the stabilizer G_W of W in G;*

ii) *the kernel of the restriction map defined by $h \mapsto h_W$ for all $h \in G_W$;*

iii) *the group $(G_W)^W$ induced by G_W on W.*

(3.3) Exercise *Let $W \leq \mathbb{F}^n$, $G \leq \mathrm{GL}_n(\mathbb{F})$. Suppose that $\dim W = m > \frac{n}{2}$ and that G_W acts irreducibly on W. Show that W is the only G_W-invariant subspace of dimension m. Deduce that $C_G(G_W) \leq G_W$.*

(3.4) Exercise *Show that $C_{\mathrm{Mat}_2(\mathbb{F})}(\mathrm{SL}_2(\mathbb{F})) = \mathbb{F}I_2$.*

(3.5) Exercise *Show, by induction on n, that $C_{\mathrm{Mat}_n(\mathbb{F})}(\mathrm{SL}_n(\mathbb{F})) = \mathbb{F}I_n$.*

Hint. For $n \geq 3$, start with any $(n-1)$-dimensional subspace W. Consider its stabilizer H in $\mathrm{SL}_n(\mathbb{F})$ and note that $H^W \cong \mathrm{GL}_{n-1}(\mathbb{F})$ acts irreducibly on W. Deduce that, for every $c \in C_{\mathrm{Mat}_n(\mathbb{F})}(H)$ and for every $w \in W$

$$cw = \lambda_c w, \quad \lambda_c \in \mathbb{F}.$$

Take another $(n-1)$-dimensional subspace $W' \neq W$. Again, for all $w' \in W'$:

$$cw' = \mu_c w', \quad \mu_c \in \mathbb{F}.$$

The conclusion follows easily from $\mathbb{F}^n = W + W'$.

(3.6) Exercise *Show that the map $(\ ,\ ) : \mathrm{Mat}_n(\mathbb{F}) \times \mathrm{Mat}_n(\mathbb{F}) \to \mathbb{F}$ defined by:*

$$(A, B) := \mathrm{tr}(AB) \tag{3.7}$$

is bilinear and that it is non-degenerate.

(3.8) Exercise *Let $G = \mathrm{Sym}(3)$ and set:*

$$\sigma = \mathrm{id} + (123) + (132) + (12) + (13) + (23),$$
$$\tau = \mathrm{id} + (123) + (132) - (12) - (13) - (23),$$
$$\rho_1 = \mathrm{id} + (12) - (13) - (123), \quad \rho_2 = \mathrm{id} + (23) - (13) - (132),$$
$$\zeta_1 = \mathrm{id} + (12) - (23) - (132), \quad \zeta_2 = \mathrm{id} + (12) - (13) + (123) - (132).$$

i) *Show that, with respect to the product $j(fg) := (jf)g$ for $j \in \{1, 2, 3\}$, $f, g \in G$:*

$$\mathbb{C}G = \mathbb{C}\sigma \oplus \mathbb{C}\tau \oplus (\mathbb{C}\rho_1 + \mathbb{C}\rho_2) \oplus (\mathbb{C}\zeta_1 + \mathbb{C}\zeta_2)$$

is a decomposition of the group algebra $\mathbb{C}G$ into 4 minimal left ideals.

ii) *Calculate explicitly the representations* f_i *of* G *afforded by these ideals and show that they are irreducible (Clearly it is enough to write* $f_i(12)$ *and* $f_i(13)$ *for* $i = 1, 2, 3$*).*

iii) *Show that* 3 *of them, say* f_1, f_2, f_3 *are inequivalent, of respective degrees* 1, 1, 2*.*

iv) *Conclude that* f_1, f_2, f_3 *are the only irreducible representations of* G *over* $\mathbb{C}$ *(use Corollary* 2.9*).*

Chapter V

Groups of Lie type

1 Lie Algebras

Our main references here will be [10] and the book of R. Carter[5].

(1.1) Definition *A* Lie algebra L *is a vector space* L*, over a field* $\mathbb{F}$*, endowed with a bilinear map* $L \times L \to L$:

$$(x, y) \mapsto [xy] \quad \text{(Lie product)}$$

for which the following conditions hold. For all $x, y, z \in L$*:*

(1) $[xx] = 0$*;*

(2) $[x[yz]] + [y[zx]] + [z[xy]] = 0$ *(*Jacobi identity*).*

By (1) any Lie product is anticommutative, namely $[xy] = -[yx]$. Indeed:

$$0 = [(x + y)(x + y)] = [xx] + [xy] + [yx] + [yy] = [xy] + [yx].$$

(1.2) Definition *Let* $\mathcal{B} = \{x_1, \ldots, x_n\}$ *be a basis of* L *over* $\mathbb{F}$*. The* structure constants *of* L *(with respect to* $\mathcal{B}$*) are the elements* $a_{ij}^k \in \mathbb{F}$ *defined by:*

$$[x_i x_j] = \sum_{k=1}^{n} a_{ij}^k x_k.$$

Every Lie product over L is determined by its structure constants by the bilinearity.

(1.3) Definition

(1) *A subspace* I *of* L *is called an* ideal *if* $[ix] \in I$ *for all* $i \in I$*,* $x \in L$*;*

(2) *L is* simple *if $L \neq \{0\}$ and it has no proper ideal.*

(1.4) Definition *A linear map $\delta : L \to L$ is called a* derivation *if it satisfies*

$$\delta([yz]) = [\delta(y)z] + [y\delta(z)], \ \forall \ y, z \in L.$$

(1.5) Example *For each $x \in L$ the derivation* $\operatorname{ad} x : L \to L$ *defined by:*

$$\operatorname{ad} x(y) := [xy], \forall \ y \in L.$$

The linearity of $\operatorname{ad} x$ is an immediate consequence of the bilinearity of the Lie product. The map $\operatorname{ad} x$ is a derivation by axioms (1) and (2) of Definition 1.1 of Lie product.

(1.6) Definition *Let L, L' be Lie algebras over $\mathbb{F}$. A map $\varphi : L \to L'$ is called a* homomorphism *if, for all $x, y \in L$:*

$$\varphi([xy]) = [\varphi(x)\varphi(y)].$$

An isomorphism *is a bijective homomorphism. An isomorphism $\varphi : L \to L$ is called an* automorphism *of L. The group of automorphisms of L is indicated by* $\operatorname{Aut}(L)$.

2 Linear Lie Algebras

An *associative algebra* A, over a field $\mathbb{F}$, is a ring A, which is a vector space over $\mathbb{F}$, satisfying the following axiom. For all $\lambda \in \mathbb{F}$ and for all $x, y \in A$:

$$\lambda(xy) = (\lambda x)y = x(\lambda y).$$

(2.1) Lemma *Let A be an associative algebra over $\mathbb{F}$. Then A is a Lie algebra with respect to the product defined by:*

$$[x, y] := xy - yx, \quad \forall \ x, y \in A. \tag{2.2}$$

Proof Routine calculation. ■

(2.3) Definition *Let V be a vector space over $\mathbb{F}$.*

(1) *The associative algebra $End_{\mathbb{F}}(V)$, considered as a Lie algebra with respect to the product* (2.2)*, is called the* general linear *Lie algebra and indicated by $\mathcal{GL}(V)$;*

(2) *the matrix algebra* $\mathrm{Mat}_n(\mathbb{F})$*, considered as a Lie algebra with respect to* (2.2)*, is indicated by* $\mathcal{GL}_n(\mathbb{F})$*;*

(3) $\mathcal{GL}_n(\mathbb{F})$ *and its subalgebras are called the* linear *Lie algebras.*

Let $\mathcal{B}$ be a fixed basis of $V = \mathbb{F}^n$. The map $\Phi_{\mathcal{B}} : \mathcal{GL}(V) \simeq \mathcal{GL}_n(\mathbb{F})$ such that $\Phi_{\mathcal{B}}(\alpha)$ is the matrix of α with respect to $\mathcal{B}$ is an isomorphism of Lie algebras. Thus:

$$\mathcal{GL}(\mathbb{F}^n) \simeq \mathcal{GL}_n(\mathbb{F}).$$

A basis of $\mathcal{GL}_n(\mathbb{F})$ consists of the matrices having 1 in one position and 0 elsewhere, namely the matrices:

$$\{e_{ij} \mid 1 \leq i, j \leq n\}.$$

The structure constants, with respect to this basis, are all ± 1 or 0. More precisely:

$$[e_{ij}, e_{k\ell}] := e_{ij}e_{k\ell} - e_{k\ell}e_{ij} = \delta_{jk}e_{i\ell} - \delta_{\ell i}e_{kj}. \tag{2.4}$$

Conjugation by a fixed element of $\mathrm{GL}_n(\mathbb{F})$ is an automorphism of the associative algebra $\mathrm{Mat}_n(\mathbb{F})$ and also of the Lie algebra $\mathcal{GL}_n(\mathbb{F})$, as shown in the following:

(2.5) Lemma *For a fixed* $g \in \mathrm{GL}_n(\mathbb{F})$*, let* $\gamma_g : \mathcal{GL}_n(\mathbb{F}) \to \mathcal{GL}_n(\mathbb{F})$ *be defined by:*

$$\gamma_g(m) := g^{-1}mg, \quad \forall\ m \in \mathcal{GL}_n(\mathbb{F}).$$

Then γ_g *is an automorphism of the Lie algebra* $\mathcal{GL}_{\ell+1}(\mathbb{F})$*.*

Proof γ_g is linear since, for all $m_1, m_2, m \in \mathrm{GL}_n(\mathbb{F})$, $\lambda \in \mathbb{F}$:

$$\begin{array}{l} g^{-1}(m_1 + m_2)g = g^{-1}m_1g + g^{-1}m_2g \\ g^{-1}(\lambda m)g = \lambda g^{-1}mg \end{array}.$$

γ_g preserves the Lie product, i.e., $[g^{-1}m_1g, g^{-1}m_2g] = g^{-1}[m_1, m_2]g$. In fact:

$$g^{-1}m_1gg^{-1}m_2g - g^{-1}m_2gg^{-1}m_1g = g^{-1}\left(m_1m_2 - m_2m_1\right)g.$$

γ_g is bijective having $\gamma_{g^{-1}}$ as its inverse. ■

(2.6) Lemma *The trace map* $\mathrm{tr} : \mathcal{GL}_n(\mathbb{F}) \to \mathcal{GL}_1(\mathbb{F})$ *is a Lie algebras homomorphism. In particular its kernel is a subalgebra, indicated by* $\mathbf{A}_\ell$*.*

Proof For all $a, b \in \mathcal{GL}_n(\mathbb{F})$, $\lambda \in \mathbb{F}$:
$\mathrm{tr}(a + b) = \mathrm{tr}(a) + \mathrm{tr}(b)$,
$\mathrm{tr}(\lambda a) = \lambda\,\mathrm{tr}(a)$,
$\mathrm{tr}([a, b]) = \mathrm{tr}(ab - ba) = \mathrm{tr}(ab) - \mathrm{tr}(ba) = 0 = [\mathrm{tr}(a), \mathrm{tr}(b)]$. ■

3 The classical Lie algebras

We give an explicit description of the *classical* Lie algebras over $\mathbb{C}$.

3.1 The special linear algebra $\mathbf{A}_\ell$

$\mathbf{A}_\ell$ is the subalgebra of $\mathcal{GL}_{\ell+1}(\mathbb{C})$ consisting of the matrices of trace 0, namely the kernel of the trace homomorphism $\mathrm{tr} : \mathcal{GL}_{\ell+1}(\mathbb{C}) \to \mathcal{GL}_1(\mathbb{C})$.
A basis of $\mathbf{A}_\ell$ is given by the matrices:

$$\{e_{i,i} - e_{i+1,i+1} \mid 1 \leq i \leq \ell\} \ \cup \ \{e_{ij} | 1 \leq i \neq j \leq \ell + 1\}\,. \tag{3.1}$$

Thus, for the dimension of the special linear algebra, we get:

$$\dim_{\mathbb{C}}(\mathbf{A}_\ell) = (\ell + 1)\ell + \ell = \ell^2 + 2\ell. \tag{3.2}$$

(3.3) Theorem $\mathrm{PGL}_{\ell+1}(\mathbb{C}) \leq \mathrm{Aut}(\mathbf{A}_\ell)$.

Proof By Lemma 2.5, for all $g \in \mathrm{GL}_{\ell+1}(\mathbb{C})$, the inner automorphism

$$\gamma_g : \mathcal{GL}_{\ell+1}(\mathbb{C}) \to \mathcal{GL}_{\ell+1}(\mathbb{C})$$

is an automorphism of the Lie algebra $\mathcal{GL}_{\ell+1}(\mathbb{C})$. For all $m \in \mathbf{A}_\ell$ we have $\mathrm{tr}(\gamma_g(m)) = \mathrm{tr}(m) = 0$, i.e., $\gamma_g(\mathbf{A}_\ell) \leq \mathbf{A}_\ell$. Since $\mathbf{A}_\ell$ has finite dimension and γ_g is injective, we get $\gamma_g(\mathbf{A}_\ell) = \mathbf{A}_\ell$. So the restriction of γ_g to $\mathbf{A}_\ell$ is an automorphism of $\mathbf{A}_\ell$. Hence we may consider the homomorphism $\gamma : \mathrm{GL}_{\ell+1}(\mathbb{C}) \ \to \ \mathrm{Aut}\,(\mathbf{A}_\ell)$ defined by: $g \mapsto \gamma_g$. The kernel of γ is the subgroup Z of scalar matrices. We conclude that:

$$\mathrm{PGL}_{\ell+1}(\mathbb{C}) := \frac{\mathrm{GL}_{\ell+1}(\mathbb{C})}{Z} \quad \simeq \quad \mathrm{Im}\ \gamma \leq \mathrm{Aut}(\mathbf{A}_\ell).$$

■

3.2 The symplectic algebra $\mathbf{C}_\ell$

Let us consider the antisymmetric, non-singular matrix:

$$s = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}. \tag{3.4}$$

The symplectic algebra $\mathbf{C}_\ell$ is the subalgebra of $\mathcal{GL}_{2\ell}(\mathbb{C})$ defined by:

$$\mathbf{C}_\ell := \{x \in \mathcal{GL}_{2\ell}(\mathbb{C}) \mid sx = -x^T s\}.$$

Partitioning x into $\ell \times \ell$ blocks, we have that $x \in \mathbf{C}_\ell$ if and only if it has shape:

$$x = \begin{pmatrix} m & n \\ p & -m^T \end{pmatrix} \quad \text{with} \quad n = n^T\,,\ p = p^T \ \text{symmetric.}$$

Thus, a basis of $\mathbf{C}_\ell$ is given by the matrices:

$$\left\{ \begin{pmatrix} e_{ij} & 0 \\ 0 & -e_{ji} \end{pmatrix} \mid 1 \leq i, j \leq \ell \right\} \cup \tag{3.5}$$

$$\left\{ \begin{pmatrix} 0 & e_{ii} \\ 0 & 0 \end{pmatrix} \mid 1 \leq i \leq \ell \right\} \cup \left\{ \begin{pmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{pmatrix} \mid 1 \leq i < j \leq \ell \right\} \cup \tag{3.6}$$

$$\{\text{the transposes of (3.6)}\}. \tag{3.7}$$

So, for the dimension of the symplectic algebra, we obtain:

$$\dim_{\mathbb{C}} \mathbf{C}_\ell = \ell^{\mathbf{2}} + \mathbf{2}\left(\mathbf{1} + \frac{\ell(\ell - \mathbf{1})}{\mathbf{2}}\right) = \mathbf{2}\ell^{\mathbf{2}} + \ell. \tag{3.8}$$

(3.9) Theorem $\mathrm{PSp}_{\ell+1}(\mathbb{C}) \leq \mathrm{Aut}(\mathbf{C}_\ell)$.

Proof Let $\mathrm{Sp}_{2\ell}(\mathbb{C})$ be the group of isometries of s in (3.4). Thus

$$sg = \left(g^{-1}\right)^T s, \quad \forall\ g \in \mathrm{Sp}_{2\ell}(\mathbb{C}).$$

Take γ_g as in Lemma 2.5. Then $\gamma_g(x) = g^{-1}xg \in \mathbf{C}_\ell$, for all $x \in \mathbf{C}_\ell$. Indeed:

$$s\left(g^{-1}xg\right) = g^T sxg = g^T\left(-x^T s\right)g = -g^T x^T\ \left(g^{-1}\right)^T\ s = -\left(g^{-1}xg\right)^T s.$$

So the restriction of γ_g to $\mathbf{C}_\ell$ is an automorphism of $\mathbf{C}_\ell$. Hence we may consider the homomorphism $\gamma : \mathrm{Sp}_{2\ell}(\mathbb{C}) \ \rightarrow\ \mathrm{Aut}\,(\mathbf{C}_\ell)$ defined by: $g \mapsto \gamma_g$. The kernel of γ is the subgroup $\langle -I \rangle$ of symplectic scalar matrices. We conclude that:

$$\mathrm{PSp}_{\ell+1}(\mathbb{C}) := \frac{\mathrm{Sp}_{2\ell}(\mathbb{C})}{\langle -I \rangle} \quad \simeq \quad \mathrm{Im}\ \gamma \leq \mathrm{Aut}(\mathbf{C}_\ell).$$

■

3.3 The orthogonal algebra $\mathbf{B}_\ell$

Let us consider the symmetric, non-singular matrix:

$$s = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_\ell \\ 0 & I_\ell & 0 \end{pmatrix}. \tag{3.10}$$

The orthogonal algebra $\mathbf{B}_\ell$ is the subalgebra of $\mathcal{GL}_{2\ell+1}(\mathbb{C})$ defined by:

$$\mathbf{B}_\ell := \{x \in \mathcal{GL}_{2\ell+1}(\mathbb{C}) \mid sx = -x^T s\}.$$

Partitioning x into blocks, one has that $x \in \mathbf{B}_\ell$ if and only if it has shape

$$x = \begin{pmatrix} 0 & -v_1^T & -v_2^T \\ v_2 & m & n \\ v_1 & p & -m^T \end{pmatrix} \quad \text{with} \quad n = -n^T\,,\ p = -p^T \ \text{antisymmetric.}$$

Thus the orthogonal algebra $\mathbf{B}_\ell$ has basis:

$$\left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & e_{ij} & 0 \\ 0 & 0 & -e_{ji} \end{pmatrix} | 1 \le i, j \le \ell \right\} \cup \tag{3.11}$$

$$\left\{ \begin{pmatrix} 0 & -e_i^T & 0 \\ 0 & 0 & 0 \\ e_i & 0 & 0 \end{pmatrix} | 1 \le i \le \ell \right\} \ \cup \ \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & e_{ij} - e_{ji} \\ 0 & 0 & 0 \end{pmatrix} | 1 \le i < j \le \ell \right\} \tag{3.12}$$

$$\cup \ \{\text{the transposes of 3.12}\}\,.$$

We conclude that the dimension of this orthogonal algebra is given by:

$$\dim_{\mathbb{C}} \mathbf{B}_\ell = \ell^2 + 2\left(l + \frac{\ell(\ell-1)}{2}\right) = 2\ell^2 + \ell. \tag{3.13}$$

(3.14) Theorem *Let $G \le \mathrm{GL}_{2\ell+1}(\mathbb{C})$ be the group of isometries of s in* (3.10)*. Then*

$$\frac{ZG}{Z} \le \mathrm{Aut}(\mathbf{B}_\ell)$$

where Z denotes the group of scalar matrices.

The proof is the same as that of Theorem 3.9.

3.4 The orthogonal algebra $\mathbf{D}_\ell$

Let us consider the symmetric, non-singular matrix:

$$s = \begin{pmatrix} 0 & I_\ell \\ I_\ell & 0 \end{pmatrix}. \tag{3.15}$$

The orthogonal algebra $\mathbf{D}_\ell$ is the subalgebra of $\mathcal{GL}_{2\ell}(\mathbb{C})$ defined by:

$$\mathbf{D}_\ell := \{x \in \mathcal{GL}_{2\ell}(\mathbb{C}) \mid sx = -x^T s\}.$$

Partitioning x into blocks, one has that $x \in \mathbf{D}_\ell$ if and only if it has shape:

$$x = \begin{pmatrix} m & n \\ p & -m^T \end{pmatrix} \quad \text{with} \quad n = -n^T,\ p = -p^T \text{ antisymmetric.}$$

Thus the orthogonal algebra $\mathbf{D}_\ell$ has basis:

$$\left\{ \begin{pmatrix} e_{ij} & 0 \\ 0 & -e_{ji} \end{pmatrix} | 1 \leq i, j \leq \ell \right\} \cup \tag{3.16}$$

$$\left\{ \begin{pmatrix} 0 & e_{ij} - e_{ji} \\ 0 & 0 \end{pmatrix} | 1 \leq i < j \leq \ell \right\} \cup \{\text{their transposes}\}. \tag{3.17}$$

We conclude that the dimension of this orthogonal algebra is given by:

$$\dim_{\mathbb{C}} \mathbf{D}_\ell = \ell^{\mathbf{2}} + \mathbf{2}\frac{\ell(\ell - \mathbf{1})}{\mathbf{2}} = \mathbf{2}\ell^{\mathbf{2}} - \ell. \tag{3.18}$$

(3.19) Theorem *Let $G \leq \mathrm{GL}_{2\ell}(\mathbb{C})$ be the group of isometries of s in* (3.15)*. Then*

$$\frac{ZG}{Z} \leq \mathrm{Aut}(\mathbf{D}_\ell)$$

where Z denotes the group of scalar matrices.

The proof is the same as that of Theorem 3.9.

4 Root systems

Let L be a finite dimensional simple Lie algebras over $\mathbb{C}$. By the classification due to Killing and Cartan, L is one of the 9 algebras denoted respectively by:

$$\mathbf{A}_\ell, \mathbf{B}_\ell,\ \mathbf{C}_\ell,\ \mathbf{D}_\ell,\ \mathbf{E}_6,\ \mathbf{E}_7,\ \mathbf{E}_8,\ \mathbf{F}_4,\ \mathbf{G}_2. \tag{4.1}$$

There exists a set $\Phi = \Phi(L)$ such that L admits a decomposition

$$L \ = \ \mathcal{H} \ \oplus \bigoplus_{r \in \Phi} L_r \quad \text{(Cartan decomposition)} \tag{4.2}$$

where $\mathcal{H}$ is an ℓ-dimensional abelian subalgebra (namely $[h_1 h_2] = 0$ for all $h_1, h_2 \in \mathcal{H}$) and, for each $r \in \Phi$, the following conditions hold:

(1) $L_r = \mathbb{C}v_r$ for some $v_r \in L$, i.e., L_r is a 1-dimensional space;

(2) $[hv_r] = r(h)v_r$ with $r(h) \in \mathbb{C}$, for all $h \in \mathcal{H}$;

(3) the map $\operatorname{ad} v_r : L \to L$ is nilpotent;

(4) there exists a unique $s \in \Phi$ (denoted by $-r$) such that $0 \neq [v_r v_s] \in \mathcal{H}$.

(4.3) Remark *Fix $y \in L$. Recalling that $\operatorname{ad} y(x) := [yx]$, for all $x \in L$, we have:*

- $\operatorname{ad} h(\mathcal{H}) = \{0\}$ *for all $h \in \mathcal{H}$ since $\mathcal{H}$ is abelian.*
- *v_r is an eigenvector of $\operatorname{ad} h$, with eigenvalue $r(h)$, by point (2) above.*

Every $r \in \Phi$ may be identified with the linear map $r : \mathcal{H} \to \mathbb{C}$ defined by $h \mapsto r(h)$. Clearly r is an element of the dual space $\mathcal{H}^*$ of $\mathcal{H}$, by the bilinearity of the Lie product. Moreover different elements of Φ give rise to different maps. So:

$$\Phi \subseteq \mathcal{H}^*.$$

Now, consider the bilinear, symmetric form: $L \times L \to \mathbb{C}$ defined by

$$(x, y) := \operatorname{tr}(\operatorname{ad} x \ \operatorname{ad} y) \quad \text{(Killing form)}.$$

Since this form is non-degenerate, its restriction to $\mathcal{H} \times \mathcal{H}$ induces the isomorphism of vector spaces $\varphi : \mathcal{H} \to \mathcal{H}^*$ where, for each $\overline{h} \in \mathcal{H}$:

$$\varphi(\overline{h})\,(h) := \operatorname{tr}(\operatorname{ad} \overline{h} \ \operatorname{ad} h), \quad \forall\ h \in \mathcal{H}.$$

Identifying each $r \in \Phi$ with its preimage in $\mathcal{H}$, we may assume:

$$\Phi \subseteq \mathcal{H}.$$

It can be shown that Φ contains a $\mathbb{C}$-basis

$$\Pi = \{r_1,\ \ldots,\ r_\ell\} \quad \text{(fundamental system)}$$

of $\mathcal{H}$ such that every $r \in \Phi$:

(1) is a linear combination of elements in Π with *rational* coefficients;

(2) these coefficients are either all positive, or all negative.

Property (2) defines an obvious partition of Φ into positive and negative roots:

$$\Phi \;=\; \Phi^+ \,\dot\cup\, \Phi^-.$$

By property (1), Φ is a subset of the real vector space:

$$\mathcal{H}_{\mathbb{R}} \;:=\; \mathbb{R}r_1 \oplus \;\cdots\oplus\; \mathbb{R}r_\ell \;\simeq \mathbb{R}^\ell.$$

$\mathcal{H}_{\mathbb{R}}$ is an *euclidean space* with respect to the Killing form as scalar product:

$$(x,y) := \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y), \quad \forall\; x,y \in \mathcal{H}_{\mathbb{R}}.$$

The *length* of a vector $x \in \mathcal{H}_{\mathbb{R}}$ and the *angle* $\widehat{xy}$ for $x, y \in \mathcal{H}_{\mathbb{R}} \setminus \{0\}$ are defined by:

$$|x| := \sqrt{(x,x)}, \quad \cos\widehat{xy} := \frac{(x,y)}{|x|\,|y|}.$$

(4.4) Definition *The numbers A_{rs} are defined by:*

$$A_{rs} := \frac{2(r,s)}{(r,r)}, \quad \forall\; r,s \in \Phi.$$

It turns out that all A_{rs} are in $\mathbb{Z}$. In particular, if $r, s \in \Phi$ are linearly independent and $r+s \in \Phi$, then $A_{rs} = p - q$ where $0 \le p, q \in \mathbb{N}$ and

$$-pr+s, \ldots, s, \ldots, qr+s \tag{4.5}$$

is the longest chain of roots through s involving r.

(4.6) Example *Take the root system Φ with $\Phi^+ = \{r_1, r_2, r_1+r_2, 2r_1+r_2\}$.*
Set $s = r_1 + r_2$, $t = 2r_1 + r_2$.

r, s	Longest chain	p, q	A_{rs}
$r_1,\ r_2$	$r_2,\ r_2+r_1,\ r_2+2r_1$	$0, 2$	-2
$r_1,\ r_1+r_2$	$-r_1+(r_1+r_2),\ (r_1+r_2),\ (r_1+r_2)\,r_1$	$1, 1$	0
$r_1,\ 2r_1+r_2$	$-2r_1+(2r_1+r_2),\ -r_1+(2r_1+r_2),\ t$	$2, 0$	2
$r_2,\ r_1$	$r_1,\ r_1+r_2$	$0, 1$	-1
$r_2,\ r_1+r_2$	$-r_2+(r_1+r_2),\ (r_1+r_2)$	$1, 0$	1
$r_2,\ 2r_1+r_2$	$2r_1+r_2$	$0, 0$	0 .

The *Cartan matrix* of L, with respect to a basis $\{r_1, \ldots, r_\ell\}$ of $\mathcal{H}_{\mathbb{R}}$, is defined as:

$$A \;:=\; \left(\frac{2(r_i,r_j)}{(r_i,r_i)}\right), \quad 1 \le i,j \le \ell. \tag{4.7}$$

A basis $\{r_1,\ \ldots\ ,\ r_\ell\}$ of $\mathcal{H}_{\mathbb{R}}$ can be normalized into the basis $\{h_{r_1},\ \ldots\ ,\ h_{r_\ell}\}$, where:

$$h_i := \frac{2r_i}{(r_i,r_i)},\ 1 \le i \le \ell.$$

4.1 Root system of type $\mathbf{A}_\ell$

Let$\{e_1,\ \ldots,\ e_{\ell+1}\}$ be an orthonormal basis of the euclidean space $\mathbb{R}^{\ell+1}$.
The following vectors of $\mathbb{R}^{\ell+1}$ form a fundamental system of type $\mathbf{A}_\ell$:

$$\Pi = \left\{ \underbrace{-e_1+e_2}_{r_1},\ \underbrace{-e_2+e_3}_{r_2},\ \ldots, \underbrace{-e_\ell+e_{\ell+1}}_{r_\ell} \right\}.$$

The full root system has order $\ell(\ell+1)$ and is as follows:

$$\Phi = \underbrace{\{-e_i+e_j, \mid 1\le i<j\le \ell+1\}}_{\Phi^+} \ \dot\cup\ \underbrace{\{e_i-e_j, \mid 1\le i<j\le \ell+1\}}_{\Phi^-}.$$

All roots $r\in\Phi$ have the same length $|r|=\sqrt{2}$ (for this root system).
Cartan matrix:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 2 & -1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 2 & -1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 & 2 & -1 \\ 0 & 0 & 0 & \ldots & 0 & -1 & 2 \end{pmatrix}.$$

4.2 Root system of type $\mathbf{B}_\ell$

Let $\{e_1,\ \ldots,\ e_\ell\}$ be an orthonormal basis of the euclidean space $\mathbb{R}^\ell$.
The following vectors form a fundamental system of type $\mathbf{B}_\ell$

$$\Pi = \left\{ \underbrace{e_1-e_2}_{r_1},\quad \underbrace{e_2-e_3}_{r_2},\quad \ldots,\quad \underbrace{e_{\ell-1}-e_\ell}_{r_{\ell-1}},\quad \underbrace{e_\ell}_{r_\ell} \right\}.$$

The full root system has order $2\ell^2$ and is as follows:

$$\Phi = \underbrace{\{e_i\pm e_j,\ e_i \mid 1\le i<j\le \ell\}}_{\Phi^+} \ \dot\cup\ \underbrace{\{-e_i\mp e_j,\ -e_i \mid 1\le i<j\le \ell\}}_{\Phi^-}.$$

For all $r\in\Phi$ we have $|r|\in\{\sqrt{2},\ 1\}$. So there are *long* and *short* roots. E.g. the r_i-s, $i\le \ell-1$, are long, r_ℓ is short.
Cartan matrix:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 2 & -1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 2 & -1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 & 2 & -1 \\ 0 & 0 & 0 & \ldots & 0 & -2 & 2 \end{pmatrix}.$$

4.3 Root system of type $\mathbf{C}_\ell$

Let $\{e_1,\ \ldots,\ e_\ell\}$ be an orthonormal basis of the euclidean space $\mathbb{R}^\ell$.
The following vectors form a fundamental system of type $\mathbf{C}_\ell$

$$\Pi = \left\{ \underbrace{e_1 - e_2}_{r_1},\quad \underbrace{e_2 - e_3}_{r_2},\quad \ldots,\quad \underbrace{e_{\ell-1} - e_\ell}_{r_{\ell-1}},\quad \underbrace{2e_\ell}_{r_\ell} \right\}.$$

The full root system has order $2\ell^2$ and is as follows:

$$\Phi = \underbrace{\{e_i \pm e_j,\ 2e_i \mid 1 \le i < j \le \ell\}}_{\Phi^+} \ \dot\cup\ \underbrace{\{-e_i \mp e_j,\ -2e_i,\ \mid 1 \le i < j \le \ell\}}_{\Phi^-}.$$

For all $r \in \Phi$ we have $|r| \in \{\sqrt{2},\ 2\}$. Here the r_i-s, $i \le \ell - 1$, are short, r_ℓ is long.
Cartan matrix:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 2 & -1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 2 & -1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 & 2 & -2 \\ 0 & 0 & 0 & \ldots & 0 & -1 & 2 \end{pmatrix}.$$

4.4 Root system of type $\mathbf{D}_\ell$

Let $\{e_1,\ \ldots,\ e_\ell\}$ be an orthonormal basis of the euclidean space $\mathbb{R}^\ell$.
The following vectors form a fundamental system of type $\mathbf{D}_\ell$

$$\Pi = \left\{ \underbrace{e_1 - e_2}_{r_1},\quad \underbrace{e_2 - e_3}_{r_2},\quad \ldots,\quad \underbrace{e_{\ell-1} - e_\ell}_{r_{\ell-1}},\quad \underbrace{e_{\ell-1} + e_\ell}_{r_\ell} \right\}.$$

The full root system has order $2\ell(\ell-1)$ and is as follows:

$$\Phi = \underbrace{\{e_i \pm e_j \mid 1 \le i < j \le \ell\}}_{\Phi^+} \ \dot\cup\ \underbrace{\{-e_i \mp e_j \mid 1 \le i < j \le \ell\}}_{\Phi^-}.$$

As in the case of $\mathbf{A}_\ell$ all roots have the same length. For this system $|r| = \sqrt{2}$.
Cartan matrix:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 2 & -1 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & -1 & 2 & -1 & -1 \\ 0 & 0 & 0 & \ldots & -1 & 2 & 0 \\ 0 & 0 & 0 & \ldots & -1 & 0 & 2 \end{pmatrix}.$$

5 Chevalley basis of a simple Lie algebra

Let $L = \mathcal{H} \oplus \bigoplus_{r\in\Phi} L_r$ be a simple Lie algebra over $\mathbb{C}$, with fundamental system Π. Chevalley has proved the existence of a basis of L

$$\{h_r \mid r \in \Pi\} \cup \{e_r \mid r \in \Phi\} \quad \text{(Chevalley basis)} \tag{5.1}$$

where $\mathcal{H} = \bigoplus_{r\in\Pi} \mathbb{C}h_r$ and $L_r = \mathbb{C}e_r$ for each r, satisfying the following conditions:

- $[h_r h_s] = 0$, for all $r, s \in \Pi$;
- $[h_r e_s] = A_{rs} e_s$, for all $r \in \Pi$, $s \in \Phi$, with A_{rs} as in Definition 4.4;
- $[e_r e_{-r}] = h_r$, for all $r \in \Phi$;
- $[e_r e_s] = 0$, for all $r, s \in \Phi$, $r + s \neq 0$ and $r + s \notin \Phi$;
- $[e_r e_s] = \pm(p + 1)e_{r+s}$, if $r + s \in \Phi$, with p as in (4.5).

In particular, with respect to a Chevalley basis, the multiplication constants of L are all in $\mathbb{Z}$, a crucial property for the definition of the groups of Lie type over any field $\mathbb{F}$.

(5.2) Lemma *Suppose that L is linear and that $\mathcal{H}$ consists of diagonal matrices. Then, for each $r \in \Phi$, we have $e_{-r} = e_r^T$.*

Proof For all $h \in \mathcal{H}$, $\operatorname{ad} h\,(e_r) = he_r - e_r h = r(h)e_r$. The condition $h = h^T$ gives:

$$\operatorname{ad} h\left({e_r}^T\right) = h{e_r}^T - {e_r}^T h = (e_r h - he_r)^T = -r(h){e_r}^T.$$

■

(5.3) Example *Chevalley basis of* $\mathbf{A}_1$.

$$\mathbf{A}_1 \;=\; \mathbb{C}\underbrace{\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}}_{h_{r_1}} \oplus \mathbb{C}\underbrace{\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}}_{e_{r_1}} \oplus \mathbb{C}\underbrace{\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}}_{e_{-r_1}}.$$

Let $h = \begin{pmatrix}a & 0\\ 0 & -a\end{pmatrix} \in \mathcal{H}$. With respect to the above basis:

$$(\operatorname{ad} h)_{|\langle e_{r_1}, e_{-r_1}\rangle} = \begin{pmatrix}2a & 0\\ 0 & -2a\end{pmatrix} \quad \Longrightarrow \begin{cases} r_1(h) &=& 2a\\ -r_1(h) &=& -2a\end{cases}$$

Since $2a = \operatorname{tr}\left(\operatorname{ad}\begin{pmatrix}1/4 & 0\\ 0 & -1/4\end{pmatrix}\operatorname{ad} h\right)$, the Killing form allows the identification:

$$r_1 = \begin{pmatrix}1/4 & 0\\ 0 & -1/4\end{pmatrix}.$$

Normalized basis of $\mathcal{H}$:

$$h_1 := \frac{2r_1}{(r_1, r_1)} = \frac{2r_1}{\operatorname{tr}(\operatorname{ad} r_1)^2} = \frac{2}{1/2} r_1 = 4r_1 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

Root system: $\Phi = \{r_1,\ -r_1\}$.

(5.4) Example *Chevalley basis of* $\mathbf{A}_2$.

$$\mathbf{A}_2 \ = \ \underbrace{\mathbb{C}h_{r_1} \oplus \mathbb{C}h_{r_2}}_{\mathcal{H}} \oplus \mathbb{C}e_{r_1} \oplus \ \mathbb{C}e_{r_2} \oplus \ \mathbb{C}e_s \oplus \ \mathbb{C}e_{-r_1} \oplus \ \mathbb{C}e_{-r_2} \oplus \ \mathbb{C}e_{-s}$$

where:

$$h_{r_1} = \begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}, \quad h_{r_2} = \begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix}, \quad e_{r_1} = \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}, \quad e_{r_2} = \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}$$

$$e_s = \begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}, \quad e_{-r_1} = e_{r_1}^T, \quad e_{-r_2} = e_{r_2}^T, \quad e_{-s} = e_s^T.$$

We justify and complete the notation. Let $h = \begin{pmatrix}a & 0 & 0\\ 0 & b & 0\\ 0 & 0 & -a-b\end{pmatrix} \in \mathcal{H}$.

With respect to the above ordered basis:

$$\operatorname{ad} h_{|\langle e_{r_1},\ e_{r_2},\ e_s\rangle} = \begin{pmatrix}a-b & 0 & 0\\ 0 & a+2b & 0\\ 0 & 0 & 2a+b++\end{pmatrix}$$

$$\Longrightarrow \quad \begin{cases} r_1(h) = a-b\\ r_2(h = a+2b \quad \text{giving} \quad s = r_1 + r_2.\\ s(h) = 2a+b\end{cases}$$

Since

$$a-b = \operatorname{tr}\left(\operatorname{ad}\begin{pmatrix}1/6 & 0 & 0\\ 0 & -1/6 & 0\\ 0 & 0 & 0\end{pmatrix}\operatorname{ad} h\right), \quad a+2b = \operatorname{tr}\left(\operatorname{ad}\begin{pmatrix}0 & 0 & 0\\ 0 & 1/6 & 0\\ 0 & 0 & -1/6\end{pmatrix}\operatorname{ad} h\right)$$

the Killing form allows the identifications:

$$r_1 = \begin{pmatrix}1/6 & 0 & 0\\ 0 & -1/6 & 0\\ 0 & 0 & 0\end{pmatrix}, \quad r_2 = \begin{pmatrix}0 & 0 & 0\\ 0 & 1/6 & 0\\ 0 & 0 & -1/6\end{pmatrix}.$$

Normalized basis of $\mathcal{H}$: $\left\{\frac{2r_1}{(r_1,r_1)} = h_{r_1},\ \frac{2r_2}{(r_2,r_2)} = h_{r_1}\right\}$ with h_{r_1}, h_{r_2} as above.
Root system $\Phi = \Phi^+ \cup\ \Phi^-$, with

$$\Phi^+ = \{r_1,\ r_2,\ r_1 + r_2\}\,, \quad \Phi^- = \{\ -r_1,\ -r_2,\ -r_1 - r_2\}\,.$$

(5.5) Example *As fundamental system of* $\mathbf{A}_\ell$ *one may take the* $\ell+1 \times \ell+1$ *matrices*

$$e_{r_1} = e_{1,2}, \quad e_{r_2} = e_{2,3}, \quad \ldots, \quad e_{r_\ell} = e_{\ell,\ell+1}.$$

(5.6) Example *Chevalley basis of* $\mathbf{C_2}$.

$$\mathbf{C}_2 \ = \ \underbrace{\mathbb{C}h_{r_1} \oplus \mathbb{C}h_{r_2}}_{\mathcal{H}} \oplus \mathbb{C}e_{r_1} \oplus\ \mathbb{C}e_{r_2} \oplus\ \mathbb{C}e_s \oplus\ \mathbb{C}e_t \oplus\ \mathbb{C}e_{-r_1} \oplus\ \mathbb{C}e_{-r_2} \oplus\ \mathbb{C}e_{-s} \oplus\ \mathbb{C}e_{-t}$$

where:

$$h_{r_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad h_{r_2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad e_{r_1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix},$$

$$e_{r_2} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad e_s = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad e_t = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$e_{-r_1} = e_{r_1}^T, \quad e_{-r_2} = e_{r_2}^T, \quad e_{-s} = e_s^T, \quad e_{-t} = e_t^T.$$

We justify and complete the notation. Let $h = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & -a & 0 \\ 0 & 0 & 0 & -b \end{pmatrix}$.

With respect to the above ordered basis:

$$(\operatorname{ad} h)_{\langle e_{r_1}, e_{r_2}, e_s, e_t\rangle} = \begin{pmatrix} a-b & 0 & 0 & 0 \\ 0 & 2b & 0 & 0 \\ 0 & 0 & a+b & 0 \\ 0 & 0 & 0 & 2a \end{pmatrix}$$

$$\Longrightarrow \quad \begin{cases} r_1(h) = a - b \\ r_2(h) = 2b \\ s(h) = a + b \\ t(h) = 2a \end{cases} \quad \text{giving} \quad \begin{cases} s = r_1 + r_2 \\ t = 2r_1 + r_2. \end{cases}$$

Since

$$-a + b = \operatorname{tr}\left(\operatorname{ad}\begin{pmatrix} -1/12 & 0 & 0 & 0 \\ 0 & 1/12 & 0 & 0 \\ 0 & 0 & 1/12 & 0 \\ 0 & 0 & 0 & -1/12 \end{pmatrix} \operatorname{ad} h\right),$$

$$2a = \mathrm{tr}\left(\mathrm{ad}\begin{pmatrix}1/6 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & -1/6 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\mathrm{ad}\, h\right)$$

the Killing form allows the identifications:

$$r_1 = \begin{pmatrix}-1/12 & 0 & 0 & 0\\ 0 & 1/12 & 0 & 0\\ 0 & 0 & 1/12 & 0\\ 0 & 0 & 0 & -1/12\end{pmatrix}, \quad r_2 = \begin{pmatrix}1/6 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & -1/6 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}.$$

$(r_1, r_1) = \frac{1}{6}$, $(r_2, r_2) = \frac{1}{3}$, $(r_1, r_2) = -\frac{1}{6}$. Cartan matrix $\begin{pmatrix}2 & -2\\ -1 & 2\end{pmatrix}$.

Normalized basis of $\mathcal{H}$: $\left\{h_{r_1} = \frac{2r_1}{(r_1,r_1)},\ h_{r_2} = \frac{2r_2}{(r_2,r_2)}\right\}$ with h_{r_1}, h_{r_2} as above.

Root system: $\Phi = \{r_1,\ r_2,\ r_1 + r_2,\ 2r_1 + r_2,\ -r_1,\ -r_2,\ -r_1 - r_2,\ -2r_1 - r_2\}$

The non-trivial products of basis elements are written below. They agree with the conditions for a Chevalley basis given at the beginning of this Section, and also with the values of A_{rs} given in Example 4.6.

$[\ ]$	e_{r_1}	e_{r_2}	$e_{r_1+r_2}$	$e_{2r_1+r_2}$	e_{-r_1}	e_{-r_2}	$e_{-r_1-r_2}$	$e_{-2r_1-r_2}$
h_{r_1}	$2e_{r_1}$	$-2e_{r_2}$	0	$2e_{2r_1+r_2}$	$-2e_{-r_1}$	$2e_{-r_2}$	0	$-2e_{-2r_1-r_2}$
h_{r_2}	$-e_{r_1}$	$2e_{r_2}$	$e_{r_1+r_2}$	0	e_{-r_1}	$-2e_{-r_2}$	$-e_{-r_1-r_2}$	0

$[\ ]$	e_{r_1}	e_{r_2}	$e_{r_1+r_2}$	$e_{2r_1+r_2}$	e_{-r_1}	e_{-r_2}	$e_{-r_1-r_2}$	$e_{-2r_1-r_2}$
e_{r_1}	0	$e_{r_1+r_2}$	$2e_{2r_1+r_2}$	0	h_{r_1}	0	$-2e_{-r_2}$	$-e_{-r_1-r_2}$
e_{r_2}	$-e_{r_1+r_2}$	0	0	0	0	h_{r_2}	e_{-r_1}	0
$e_{r_1+r_2}$	$-2e_{2r_1+r_2}$	0	0	0	$-2e_{r_2}$	e_{r_1}	$h_{r_1+r_2}$	e_{-r_1}
$e_{2r_1+r_2}$	0	0	0	0	$-e_{r_1+r_2}$	0	e_{r_1}	$h_{2r_1+r_2}$
e_{-r_1}	$-h_{r_1}$	0	$2e_{r_2}$	$e_{r_1+r_2}$	0	$-e_{-r_1-r_2}$	$-2e_{-2r_1-r_2}$	0
e_{-r_2}	0	$-h_{r_2}$	$-e_{r_1}$	0	$e_{-r_1-r_2}$	0	0	0
$e_{-r_1-r_2}$	$2e_{-r_2}$	$-e_{-r_1}$	$-h_{r_1+r_2}$	$-e_{r_1}$	$2e_{-2r_1-r_2}$	0	0	0
$e_{-2r_1-r_2}$	$e_{-r_1-r_2}$	0	$-e_{-r_1}$	$-h_{2r_1+r_2}$	0	0	0	0

6 The action of $\exp \mathrm{ad}\, e$, with e nilpotent

Let L be a linear Lie algebra over $\mathbb{C}$ and $e \in L$. Consider the map $\mathrm{ad}\, e : L \to L$, defined as $x \mapsto [ex]$. The following identity, which can be verified by induction, holds:

$$\frac{(\mathrm{ad}\, e)^k}{k!}(x) \ = \ \sum_{i=0}^{k} \frac{e^i}{i!}\ x\ \frac{(-e)^{k-i}}{(k-i)!}, \quad \forall\ k \in \mathbb{N}. \tag{6.1}$$

In particular, if e is a nilpotent matrix, then $\operatorname{ad} e$ is nilpotent and we may consider the linear map:

$$\exp \operatorname{ad} e := \sum_{k=0}^{\infty} \frac{(\operatorname{ad} e)^k}{k!}.$$

(6.2) Lemma *Let L be a subalgebra of the general linear Lie algebra $\mathcal{G}L_n(\mathbb{C})$ and let $e \in L$ be a nilpotent matrix. Then, for all $x \in L$:*

$$\exp \operatorname{ad} e(x) = (\exp e)\, x\, (\exp e)^{-1}. \tag{6.3}$$

In particular the map $\exp \operatorname{ad} e : L \to L$ is an automorphism of L.

For the proof, based on (6.1), see [5, Lemma 4.5.1, page 66]. The conclusion follows from Lemma 2.5 of this chapter.
In the next two examples we give a proof of (6.3) in the most frequent cases.

(6.4) Example *Let $e^2 = 0$. Then $\exp e = I + e$. Moreover:*

$$\begin{aligned}
&\operatorname{ad} e : x \mapsto [e, x] = ex - xe \\
&(\operatorname{ad} e)^2 : x \mapsto [e, ex - xe] = -2(exe) \\
&(\operatorname{ad} e)^3 : x \mapsto [e, -2exe] = 0.
\end{aligned}$$

Thus $\exp \operatorname{ad} e = I + \operatorname{ad} e + \frac{1}{2} (\operatorname{ad} e)^2$ and:

$$\exp \operatorname{ad} e(x) \;=\; x + (ex - xe) - exe = (I + e)\, x\, (I - e) = (\exp e)\, x\, (\exp e)^{-1}.$$

(6.5) Example *Let $e^3 = 0$. Then $\exp e = I + e + \frac{1}{2}e^2$. Moreover:*

$$\begin{aligned}
&\operatorname{ad} e : x \mapsto ex - xe \\
&(\operatorname{ad} e)^2 : x \mapsto [e, ex - xe] = e^2x - 2exe + xe^2 \\
&(\operatorname{ad} e)^3 : x \mapsto [e, e^2x - 2exe + xe^2] = -3e^2xe + 3exe^2 \\
&(\operatorname{ad} e)^4 : x \mapsto [e, -3e^2xe + 3exe^2] = 6e^2xe^2 \\
&(\operatorname{ad} e)^5 : x \mapsto [e, 6e^2xe^2] = 0.
\end{aligned}$$

Thus $\exp \operatorname{ad} e = I + \operatorname{ad} e + \frac{1}{2} (\operatorname{ad} e)^2 + \frac{1}{6} (\operatorname{ad} e)^3 + \frac{1}{24} (\operatorname{ad} e)^4$ and

$$\exp \operatorname{ad} e(x) \;=\; x + (ex - xe) + \left(\frac{1}{2}e^2x - exe + \frac{1}{2}xe^2\right) - \frac{1}{2}\left(e^2xe - exe^2\right) + \frac{1}{4}e^2xe^2 =$$

$$\left(I + e + \frac{1}{2}e^2\right) x \left(I - e + \frac{1}{2}e^2\right) \;=\; (\exp e)x\, (\exp e)^{-1}.$$

7 Groups of Lie type

Let L be a simple Lie algebra over $\mathbb{C}$, with Chevalley basis as in (5.1):

$$\{h_r \mid r \in \Pi\} \cup \{e_r \mid r \in \Phi\}\,.$$

For all $r \in \Phi$ and for all $t \in \mathbb{C}$, we set

$$x_r(t) := \exp\,(t\,ad)\;e_r \tag{7.1}$$

(7.2) Definition *The Lie group* $L(\mathbb{C})$ *is the subgroup of* $\mathrm{Aut}(L)$ *generated by the automorphisms* (7.1)*, namely the group:*

$$L(\mathbb{C}) \;:=\; \langle x_r(t) \mid t \in \mathbb{C},\; r \in \Phi\rangle.$$

Since the structure constants are integers, it is possible to define a Lie algebra $\mathbb{F}\otimes_{\mathbb{Z}} L = L_{\mathbb{F}}$ over any field $\mathbb{F}$. The matrix representing $x_r(t)$ with respect to a Chevalley basis has entries of the form at^i where $a \in \mathbb{Z}$ and $i \in \mathbb{N}$. Interpreting a as an element of $\mathbb{F}$, one can identify $x_r(t)$ with an element of $\mathrm{Aut}\,(L_{\mathbb{F}})$ and define the group $L(\mathbb{F})$ as

$$L(\mathbb{F}) \;:=\; \langle x_r(t) \mid t \in \mathbb{F},\; r \in \Phi\rangle \quad \text{(the group of type } L \text{ over } \mathbb{F}).$$

The identifications are as follows (see Section 3):

- $\mathbf{A}_\ell(\mathbb{F}) \cong \mathrm{PSL}_{\ell+1}(\mathbb{F})$;
- $\mathbf{B}_\ell(\mathbb{F}) \cong \mathrm{P\Omega}_{2\ell+1}(\mathbb{F}, f)$ where f is the quadratic form: $x_0^2 + \sum_{i=1}^{\ell} x_i x_{-i}$;
- $\mathbf{C}_\ell(\mathbb{F})(\mathbb{F}) \cong \mathrm{PSp}_{2\ell}(\mathbb{F})$;
- $\mathbf{D}_\ell(\mathbb{F}) \cong \mathrm{P\Omega}_{2\ell}(\mathbb{F}, f)$ where f is the quadratic form: $\sum_{i=1}^{\ell} x_i x_{-i}$.
- ${}^{\mathbf{2}}\mathbf{A}_\ell(\mathbb{F}) \cong \mathrm{PSU}_{\ell+1}(\mathbb{F})$;
- ${}^{\mathbf{2}}\mathbf{D}_\ell(\mathbb{F}) \cong P\Omega_{2\ell}(\mathbb{F}_0, f)$ where $\mathbb{F}$ has an automorphism σ of order 2, with fixed field $\mathbb{F}_0$, and f is the form $\sum_{i=1}^{\ell-1} x_i x_{-i} + (x_\ell - \alpha x_{-\ell})\,(x_\ell - \alpha^\sigma x_{-\ell})$, $\alpha \in \mathbb{F} \setminus \mathbb{F}_0$.

The consideration of groups of Lie type allows a unified treatment of important classes of groups, like finite simple groups. According to the Classification Theorem, every finite simple group S is isomorphic to one of the following:

- a cyclic group C_p, of prime order p;

- an alternating group $\mathrm{Alt}(n)$, $n \geq 5$;
- a group of Lie type $L(\mathbb{F}_q)$, where L is one of the algebras in (4.1);
- a twisted group of Lie type ${}^iL(\mathbb{F}_q)$, namely the subgroup of $L(\mathbb{F}_{q^i})$ consisting of the elements fixed by an automorphism of order i of $L(\mathbb{F}_{q^i})$;
- one of the 26 sporadic simple groups.

8 Uniform definition of certain subgroups

Let L be a simple Lie algebra over $\mathbb{C}$, with Cartan decomposition

$$L = \mathcal{H} \oplus \bigoplus_{r \in \Phi \subseteq \mathcal{H}} \mathbb{C}e_r.$$

We describe some kinds of important subgroups, which may be defined in a uniform way.

8.1 Unipotent subgroups

For each $r \in \Phi$, the map

$$t \mapsto x_r(t) :== \exp(t \operatorname{ad} e_r) \tag{8.1}$$

is a monomorphism from the additive group $(\mathbb{F}, +)$ into the multiplicative group $L(\mathbb{F})$.

(8.2) Definition

- *The image of the monomorphism* (8.1) *is denoted by* X_r *and called the* radical subgroup *corresponding to the root* r*;*
- *the subgroup generated by all radical subgroups corresponding to positive roots is denoted by* U^+ *;*
- *the subgroup generated by all radical subgroups corresponding to negative roots is denoted by* U^-*.*

Thus:

$$\begin{aligned} X_r &= \{x_r(t) \mid t \in \mathbb{F}\} \simeq (\mathbb{F}, +) \\ U^+ &= \langle x_r(t) \mid t \in \mathbb{F}, r \in \Phi^+ \rangle \\ U^- &= \langle x_r(t) \mid t \in \mathbb{F},\ r \in \Phi^- \rangle\ . \end{aligned}$$

U^+, U^- (and their conjugates in $L(\mathbb{F})$) are called *unipotent* subgroups. By definition

$$L(\mathbb{F}) = \left\langle U^+,\ U^- \right\rangle.$$

(8.3) Example *In* $A_\ell(\mathbb{F})$ *identified with* $\mathrm{PSL}_{\ell+1}(\mathbb{F})$*:*

- X_r *is the projective image of the group* $\{I + te_{i,j} \mid t \in \mathbb{F}\}$ *for some* $i \neq j$,
- U^+ *is the projective image of the subgroup of upper unitriangular matrices,*
- U^- *is the projective image of the subgroup of lower unitriangular matrices.*

8.2 The subgroup $\langle X_r, X_{-r}\rangle$

For each $r \in \Phi$, the group $\langle X_r, X_{-r}\rangle$ fixes every vector of the Chevalley basis (5.1) except e_r, h_r, e_{-r}. Multiplying e_r by an appropriate scalar, if necessary, we may assume:

- $x_r(t)\,(e_r) = e_r$;
- $x_r(t)\,(h_r) = h_r - 2t\,e_r$;
- $x_r(t)\,(e_{-r}) = -t^2\,e_r + t\,h_r + e_{-r}$;
- $x_{-r}(t)\,(e_r) = e_r - th_r - t^2 e_{-r}$;
- $x_{-r}(t)\,(h_r) = h_r + 2t\,e_r$;
- $x_{-r}(t)\,(e_{-r}) = e_{-r}$.

(8.4) Theorem *There exists an epimorphism* $\varphi_r : \mathrm{SL}_2(\mathbb{F}) \to \langle X_r, X_{-r}\rangle$ *under which:*

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mapsto x_r(t), \qquad \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \mapsto x_{-r}(t). \tag{8.5}$$

Proof The group $\mathrm{SL}_2(\mathbb{F})$ has a matrix representation of degree 3, deriving from its action on the space of homogeneous polynomials of degree 2 over $\mathbb{F}$ in the indeterminates x, y. With respect to the basis $-x^2, 2xy, y^2$, we have:

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -2t & -t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ -t & 1 & 0 \\ -t^2 & 2t & 1 \end{pmatrix}.$$

These are the matrices of the action of $x_r(t)$ and $x_{-r}(t)$ restricted to $\langle e_r, r, e_{-r}\rangle$ by the formulas before the statement. ∎

8.3 Diagonal and monomial subgroups

In $\mathrm{SL}_2(\mathbb{F})$, for all $\lambda \in \mathbb{F}$ we have:

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda^{-1}-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \lambda-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\lambda^{-1} \\ 0 & 1 \end{pmatrix}.$$

Hence, for all $r \in \Phi$ and all $\lambda \in \mathbb{F}$ we set:

$$h_r(\lambda) := \varphi_r \left(\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \right) = x_{-r}(\lambda^{-1}-1)\; x_r(1)\; x_{-r}(\lambda-1)\; x_r(-\lambda^{-1}).$$

(8.6) Definition *The* diagonal *subgroup H of $L(\mathbb{F})$ is defined by*

$$H := \langle h_r(\lambda) \mid 0 \neq \lambda \in \mathbb{F},\ r \in \Phi \rangle. \tag{8.7}$$

The group H normalizes both U^+ and U^-.

(8.8) Definition *The product U^+H is called a* Borel subgroup *and is denoted by B^+. Similarly the product U^-H is denoted by B^-.*

(8.9) Example *Identifying $\mathbf{A}_\ell(\mathbb{F})$ with the projective image of $\mathrm{SL}_{\ell+1}(\mathbb{F})$:*

- *B^+ is the image of the group of upper triangular matrices of determinant 1,*
- *B^- is the image of the group of lower triangular matrices of determinant 1.*

In $\mathrm{SL}_2(\mathbb{F})$ we have:

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

Hence, for all $r \in \Phi$ we set:

$$n_r = \varphi_r \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right) = x_{-r}(-1)\; x_r(1)\; x_{-r}(-1).$$

(8.10) Definition *The (standard)* monomial *subgroup N of $L(\mathbb{F})$ is defined by:*

$$N := \langle h_r(\lambda), n_r \mid r \in \Phi, \lambda \in \mathbb{F} \rangle. \tag{8.11}$$

H is a normal subgroup of N.

(8.12) Definition *The factor group $W(L) := \frac{N}{H}$ is called the* Weyl group *of L.*

$$
\begin{aligned}
W(\mathbf{A}_\ell) &\simeq \mathrm{Sym}\,(\ell+1)\,,\\
W(\mathbf{C}_\ell) &\simeq W(\mathbf{B}_\ell) \simeq C_2^\ell\,\mathrm{Sym}\,(\ell)\,,\\
W(\mathbf{D}_\ell) &\simeq C_2^{\ell-1}\,\mathrm{Sym}\,(\ell)\,.
\end{aligned}
$$

(8.13) Example *In the orthogonal algebra B_1 over $\mathbb{C}$, with $\Phi = \{r, -r\}$ and basis*

$$
h_r = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad e_r = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \\ -\sqrt{2} & 0 & 0 \end{pmatrix}, \quad e_{-r} = \begin{pmatrix} 0 & 0 & -\sqrt{2} \\ \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
$$

we have:

$$
x_r(t) = I + te_r + \frac{t^2}{2}e_r^2 = \begin{pmatrix} 1 & \sqrt{2}\,t & 0 \\ 0 & 1 & 0 \\ -\sqrt{2}\,t & -t^2 & 1 \end{pmatrix}; \quad x_{-r}(t) = x_r(t)^T;
$$

$$
h_r(\lambda) = \; x_{-r}(\lambda^{-1}-1)\; x_r(1)\; x_{-r}(\lambda-1)\; x_r(-\lambda^{-1}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda^{-2} & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix};
$$

$$
n_r = x_r(1)x_{-r}(-1)x_r(1) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix};
$$

$$
h_{-r}(\lambda) = h_r(\lambda)^{-1}, \quad n_r = n_r^{-1};
$$

$$
H = \langle h_r(\lambda) \mid r \in \Phi,\ \lambda \in \mathbb{C}^* \rangle = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu^{-1} \end{pmatrix} \mid \mu \in \mathbb{C}^* \right\};
$$

$$
N = \langle h_r(\lambda), n_r \mid r \in \Phi\ \lambda \in \mathbb{C}^* \rangle = \left\{ \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & \mu^{-1} \\ 0 & \mu & 0 \end{pmatrix} \mid \mu \in \mathbb{C}^* \right\};
$$

$$
W = \frac{N}{H} \cong \left\langle \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\rangle \cong \mathrm{Sym}(2).
$$

(8.14) Example *Identifying $\mathbf{A}_\ell(\mathbb{F})$ with the projective image of $\mathrm{SL}_{\ell+1}(\mathbb{F})$:*

- *H is the image of the subgroup of diagonal matrices of determinant 1;*
- *N is the image of the subgroup of* monomial *matrices of determinant 1;*
- *the factor group $\frac{N}{H}$ is isomorphic to the symmetric group $\mathrm{Sym}(\ell+1)$.*

9 Exercises

(9.1) Exercise *Let $\varphi : L \to L'$ be a homomorphism of Lie algebras. Show that its kernel is an ideal.*

(9.2) Exercise *Let L be a Lie algebra and $x \in L$. Show that the map $\operatorname{ad} x$ is a derivation.*

(9.3) Exercise *Write a basis of $\mathbf{C}_2$ and a basis of $\mathbf{C}_3$.*

(9.4) Exercise *Show that $\mathbf{C}_\ell(\mathbb{F})$ is a Lie subalgebra of $\mathcal{GL}_{2\ell}(\mathbb{F})$.*

(9.5) Exercise *Write a basis of $\mathbf{B}_1$ and a basis of $\mathbf{B}_2$.*

(9.6) Exercise *Write a basis of $\mathbf{D}_2$.*

(9.7) Exercise *Verify formula* (6.3) *assuming $e^4 = 0$.*

Chapter VI

Maximal subgroups of the finite classical groups

Here the main references are [1], [2] and [15].

1 Some preliminary facts

(1.1) Definition *Let $1 \neq G$ be a group. A subgroup M of G is said to be* maximal *if $M \neq G$ and there exists no subgroup H such that $M < H < G$.*

If G is finite, by order reasons every subgroup $H \neq G$ is contained in a maximal subgroup. If M is maximal in G, then also every conjugate gMg^{-1} of M in G is maximal. Indeed

$$gMg^{-1} < K < G \implies M < g^{-1}Kg < G.$$

For this reason the maximal subgroups are studied up to conjugation.

(1.2) Lemma *Let $G = G'$ and let M be a maximal subgroup of G. Then:*

(1) *M contains the center Z of G;*

(2) *$\frac{M}{Z}$ is maximal in $\frac{G}{Z}$;*

(3) *the preimage in G of every maximal subgroup of $\frac{G}{Z}$ is maximal in G.*

Proof
(1) Suppose $Z \not\leq M$. Then $M < ZM$ gives $ZM = G$, by the maximality of M. Hence M is normal in G and the factor group $\frac{G}{M}$ is abelian. In fact:

$$\frac{G}{M} = \frac{ZM}{M} \cong \frac{Z}{M \cap Z}.$$

It follows $G' \leq M$, a contradiction, as we are assuming $G' = G$.
Points (2) and (3) follow from the fact that the subgroups of $\frac{G}{Z}$ are those of the form $\frac{K}{Z}$, where K is a subgroup of G which contains Z. ■

(1.3) Lemma *If $Z(G) = \{1\}$ then G is isomorphic to a subgroup of* $\mathrm{Aut}(G)$.

Proof For every $g \in G$ the map $\gamma : G \to G$ defined by $x \mapsto gxg^{-1}$ is an automorphism of G (called *inner*). Consider the homomorphism $\varphi : G \to \mathrm{Aut}(G)$ defined by: $g \mapsto \gamma$. Ker $\varphi = Z(G)$. Thus, under our assumption, $G \cong \varphi(G) \leq \mathrm{Aut}(G)$. ■

2 Aschbacher's Theorem

Let $\overline{G}_0$ be one of the following groups, with the further assumption that it is simple:

$$\mathrm{PSL}_n(q),\ \mathrm{PSU}_n(q^2),\ \mathrm{PSp}_{2m}(q),\ P\Omega^{\pm}_{2m}(q),\ P\Omega_{2m+1}(q).$$

Suppose that $\overline{G}$ is a group such that $\overline{G}_0 \lhd \overline{G} \leq \mathrm{Aut}(\overline{G}_0)$. By the subgroup structure theorem due to Aschbacher, every maximal subgroup $\overline{H}$ of $\overline{G}$, not containing $\overline{G}_0$, belongs to a class in the table below:

Rough description of the classes of maximal subgroups

$\mathcal{C}_1$	Stabilizers of subspaces	
$\mathcal{C}_2$	Stabilizers of decompositions $V = \oplus_{i=1}^{t} V_i$,	$\dim V_i = m$
$\mathcal{C}_3$	Stabilizers of prime degree extension fields of $\mathbb{F}_q$	
$\mathcal{C}_4$	Stabilizers of tensor decompositions $V = V_1 \otimes V_2$	
$\mathcal{C}_5$	Stabilizers of prime index subfields of $\mathbb{F}_q$	
$\mathcal{C}_6$	Normalisers of symplectic $-$ type $r-$ groups, $(r, q) = 1$	
$\mathcal{C}_7$	Stabilizers of decompositions $\otimes_{i=1}^{t} V_i$,	$\dim V_i = m$
$\mathcal{C}_8$	Classical subgroups	
$\mathcal{S}$	Almost simple absolutely irreducible subgroups	
$\mathcal{N}$	Novelty subgroups	

The 8 classes $\mathcal{C}_i = \mathcal{C}_i(\overline{G})$ consist of "natural" subgroups of $\overline{G}$, which can be described in geometric terms. Class $\mathcal{N}$ exists only for $\overline{G}_0 = P\Omega_8^{\pm}(p^a)$ or $\overline{G}_0 = \mathrm{PSp}_{2m}(2^a)'$ (see [4]). We will describe the structure of the groups in some of these classes in the case:

$$\overline{G} = \overline{G}_0 = \mathrm{PSL}_n(q).$$

It is easier to describe the linear preimages of such groups. To this purpose we set $V = \mathbb{F}^n$, with canonical basis $\{e_1, \ldots, e_n\}$, and $G = \mathrm{SL}_n(q)$.

3 The reducible subgroups $\mathcal{C}_1$

If W is a subspace of V, then its *stabilizer* $G_W := \{g \in G \mid gW = W\}$ is a subgroup of G. If W' is a subspace of V and $\dim W = \dim W'$, there exists $g \in G$ such that $gW = W'$. It follows that $G_{W'} = gG_Wg^{-1}$. So, if W is a subspace of dimension m, up to conjugation we may suppose:

$$W = \langle e_1, \ldots, e_m \rangle, \quad G_W = \left\{ \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \mid \det(C) = \det(A)^{-1} \right\}.$$

To see its structure we factorize G_W as follows:

$$G_W = U \; C_{q-1} \; (\mathrm{SL}_m(q) \times \mathrm{SL}_{n-m}(q)) \tag{3.1}$$

where

$$U \;=\; \left\{ \begin{pmatrix} I_m & B \\ 0 & I_{n-m} \end{pmatrix} \mid B \in \mathrm{Mat}_{m,\, n-m}(q) \right\} \;\cong\; (\mathbb{F}_q, +)^{m(n-m)}$$

$U \lhd G_W$,

$$C_{q-1} \;=\; \left\{ \begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & I_{m-1} & 0 & 0 \\ 0 & 0 & \alpha^{-1} & 0 \\ 0 & 0 & 0 & I_{n-m-1} \end{pmatrix} \mid \alpha \in \mathbb{F}_q^* \right\} \;\cong\; (\mathbb{F}_q^*, \cdot)$$

cyclic, and

$$\mathrm{SL}_m(q) \times \mathrm{SL}_{n-m}(q) \;=\; \left\{ \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \mid X \in \mathrm{SL}_m(q), Y \in \mathrm{SL}_{n-m}(q) \right\}.$$

Actually we may suppose $m \leq \frac{n}{2}$ since, considering the transpose of G_W, namely

$$G_W^T = \left\{ \begin{pmatrix} A & 0 \\ B^T & C \end{pmatrix} \mid \det(C) = \det(A)^{-1} \right\}$$

we obtain the stabilizer of a subspace of dimension $n - m \geq \frac{n}{2}$, namely of:

$$\langle e_{m+1}, \ldots, e_n \rangle .$$

(3.2) Definition *The groups in class $\mathcal{C}_1$ are called* parabolic subgroups.

They are the only subgroups in the classes $\mathcal{C}_i$, $1 \leq i \leq 8$, which contain a Sylow p-subgroup of $\mathrm{SL}_n(q)$, $q = p^a$. When W is chosen as above, the Sylow p-subgroup consists of the upper unitriangular matrices, namely:

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

4 The imprimitive subgroups $\mathcal{C}_2$

Let $n = mt$, $1 \leq m < n$ and consider a decomposition $\mathcal{D}$ of V as a direct sum

$$V = V_1 \oplus \cdots \oplus V_t$$

of t subspaces V_i, all of the same dimension m.

(4.1) Definition *The stabilizer $N_{\mathrm{GL}_n(q)}(\mathcal{D})$ of the above decomposition is the subgroup of G which permutes the spaces V_i among themselves, i.e.,*

$$N_{\mathrm{GL}_n(q)}(\mathcal{D}) := \{ g \in G \mid gV_i = V_j,\ 1 \leq i, j \leq t \} .$$

We study first the structure of $N_{\mathrm{GL}_n(q)}(\mathcal{D})$. Up to conjugation we may assume:

$$V_1 = \langle e_1, \ldots, e_m \rangle ,\ \ldots ,\quad V_t = \left\langle e_{(t-1)m+1}, \ldots, e_n \right\rangle .$$

For each $g \in N_{\mathrm{GL}_n(q)}(\mathcal{D})$, let φ_g be the permutation induced by g on the set $\{V_1, \ldots, V_t\}$. The map

$$\begin{array}{rccc} \varphi : & N_{\mathrm{GL}_n(q)}(\mathcal{D}) & \to & \mathrm{Sym}(t) \\ & g & \mapsto & \varphi_g \end{array}$$

is a homomorphism and

$$\mathrm{Ker}\,\varphi \ = \ \bigcap_{i=1}^{t} G_{V_i} = \left\{ \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ \ldots & \ldots & \ldots & \\ & & & A_t \end{pmatrix} \mid A_i \in \mathrm{GL}_m(q) \right\} \ \cong \ \mathrm{GL}_m(q)^t .$$

Denote by H the subgroup of $\mathrm{GL}_t(q)$ consisting of all permutation matrices.

Then the group:

$$\widehat{H} := H \otimes I_m = \{h \otimes I_m \mid h \in H\} \ \leq \ \mathrm{GL}_n(q)$$

permutes the V_i-s in all possible ways. Hence $\widehat{H} \leq N_{\mathrm{GL}_n(q)}(\mathcal{D})$ and

$$\varphi(\widehat{H}) = \mathrm{Sym}(t).$$

It follows:

$$N_{\mathrm{GL}_n(q)}(\mathcal{D}) = (\mathrm{Ker}\,\varphi) \ \varphi\left(\widehat{H}\right) \ \cong \ \mathrm{GL}_m(q)^t \,\mathrm{Sym}(t) \ = \ \mathrm{GL}_m(q) \wr \mathrm{Sym}(t).$$

Finally we have to determine $N_G(\mathcal{D}) = N_{\mathrm{GL}_n(q)}(\mathcal{D}) \cap \mathrm{SL}_n(q)$. To this purpose, let

$$\sigma = \begin{pmatrix} 0 & 1 & \\ -1 & 0 & \\ & & I_{n-2} \end{pmatrix}.$$

Then $\langle \sigma, \mathrm{Alt}(t) \rangle$ is a subgroup of $N_G(\mathcal{D})$ which maps onto $\mathrm{Sym}(t)$. It follows that

$$N_G(\mathcal{D}) = (\mathrm{Ker}\varphi \cap \mathrm{SL}_n(q)) \, \langle \sigma, \mathrm{Alt}(t) \rangle \, .$$

Note that $\mathrm{Ker}\,\varphi \ \cap \mathrm{SL}_n(q)$ can be factorized as the product of the group:

$$\left\{ \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ \cdots & \cdots & \cdots & \\ & & & B_t \end{pmatrix} \mid B_i \in \mathrm{SL}_m(q) \right\} \ \cong \ \mathrm{SL}_m(q)^t$$

and the group

$$\left\{ \begin{pmatrix} \mathrm{diag}(\alpha_1, \ldots, 1) & & & \\ & \mathrm{diag}(\alpha_2, \ldots, 1) & & \\ & & \cdots & \\ & & & \mathrm{diag}\left((\prod_{i=1}^{t-1} \alpha_i)^{-1}, \ldots, 1 \right) \end{pmatrix} \mid \alpha_i \in \mathbb{F}_q^* \right\}$$

is isomorphic to $(C_{q-1})^{t-1}$. Thus:

$$\frac{N_G(\mathcal{D})}{\mathrm{SL}_m(q)^t \ (C_{q-1})^{t-1}} \ \cong \ \mathrm{Sym}(t).$$

Equivalently:

$$N_G(\mathcal{D}) \ = \ \mathrm{SL}_m(q)^t \ (C_{q-1})^{t-1} \, . \, \mathrm{Sym}(t) \quad (\mathrm{non-split\ extension}).$$

(4.2) Remark *For $m = 1$, the subgroup $N_{\mathrm{GL}_n(q)}(\mathcal{D})$ coincides with the standard monomial subgroup.*

5 The irreducible subgroups $\mathcal{C}_3$

(5.1) Lemma *Let $\mathbb{K}$ be a subfield of the field $\mathbb{F}$. Two matrices $A, B \in \mathrm{Mat}_n(\mathbb{K})$ are conjugate under $\mathrm{GL}_n(\mathbb{K})$ if and only if they are conjugate under $\mathrm{GL}_n(\mathbb{F})$.*

Proof The rational canonical forms C_A e C_B of A and B respectively lie in $\mathrm{Mat}_n(\mathbb{K})$. If A, B are conjugate under $\mathrm{GL}_n(\mathbb{F})$, we have $C_A = C_B$. Hence A and B are conjugate also under $\mathrm{GL}_n(\mathbb{K})$, having the same rational canonical form. The converse is obvious. ■

(5.2) Lemma $\mathrm{Mat}_n(q)$ *contains a self-centralizing subalgebra $R \cong \mathbb{F}_{q^n}$. Moreover*

$$\frac{N_{\mathrm{GL}_n(q)}(R)}{C_{\mathrm{GL}_n(q)}(R)} \cong \mathrm{Gal}_{\mathbb{F}_q}(\mathbb{F}_{q^n}) \cong C_n \text{ (cyclic group of order } n).$$

Proof Let $p(t)$ be an irreducible polynomial of degree n in $\mathbb{F}_q[t]$. Denoting by A its companion matrix, we obtain the subring:

$$\mathbb{F}_q[A] = \mathbb{F}_q I_n + \mathbb{F}_q A + \cdots + \mathbb{F}_q A^{n-1} \;\cong\; \frac{\mathbb{F}_q[t]}{\langle p(t)\rangle} \;\cong \mathbb{F}_{q^n}.$$

Since $\mathbb{F}_q^n$ is an irreducible A-module, the centralizer C of A in $\mathrm{Mat}_n(q)$ is a field. The multiplicative group $C \setminus \{0\}$ is generated by a matrix $B \in \mathrm{Mat}_n(q)$. Since the minimal polynomial of B has degree $\leq n$, the dimension of C over $\mathbb{F}_q$ does not exceed n. We conclude that $C = \mathbb{F}_q[A]$. Thus we take $R = \mathbb{F}_q[A]$.
The Jordan form of A in $\mathrm{Mat}_n(q^n)$ is $J_A = \mathrm{diag}\left(\epsilon, \epsilon^q, \ldots, \epsilon^{q^{n-1}}\right)$ where ϵ is a root of $p(t)$ in $\mathbb{F}_{q^n}$. It follows that J_A is conjugate to $(J_A)^q$ in $\mathrm{GL}_n(q^n)$. By the previous Lemma, there exists $g \in \mathrm{GL}_n(q)$ such that $g^{-1}Ag = A^q$. Clearly g normalizes R. Moreover the automorphism $\gamma : R \to R$ such that $X \mapsto g^{-1}Xg$ for all $X \in R$, has order n. Hence it generates the Galois group $\mathrm{Gal}_{\mathbb{F}_q}(\mathbb{F}_{q^n})$.
Finally, let y be an element of the normalizer of R in $\mathrm{GL}_n(q)$. The map $\nu : R \to R$ such that $X \to y^{-1}Xy$ for all $X \in R$, is a field automorphism. The scalar matrices, which form the subfield of R of order q, are fixed by ν. We conclude that $\nu \in \mathrm{Gal}_{\mathbb{F}_q}(\mathbb{F}_{q^n})$. ■

The subgroups of class $\mathcal{C}_3$ are $N(R) \cap \mathrm{SL}_n(q)$, where $N(R)$ is defined as in the previous Lemma.

6 Groups in class $\mathcal{S}$

They arise from absolutely irreducible representations of simple groups. We give only some examples.

6.1 The Suzuki groups $Sz(q)$ in $\mathrm{Sp}_4(q)$

The Suzuki groups ${}^2B_2(q) = Sz(q)$ are simple groups of order $q^2(q-1)(q^2+1)$, with $q = 2^{2r+1}$, $r \geq 1$. They were discovered by M.Suzuki in 1960. $Sz(q)$ was originally defined as the subgroup of $\mathrm{SL}_4(2^{2r+1})$ generated by:

$$T := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{6.1}$$

and by the groups:

$$Q := \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ \alpha^r & 1 & 0 & 0 \\ \beta & \alpha & 1 & 0 \\ \alpha^{2r+1} + \alpha^r\beta + \beta^{2r} & \alpha^{r+1} + \beta & \alpha^r & 1 \end{pmatrix} \mid \alpha, \beta \in \mathbb{F}_q \right\}. \tag{6.2}$$

T and Q fix the symplectic form T. Hence $Sz(q)$ is a subgroup of $\mathrm{Sp}_4(q)$, with respect to T. For $q \geq 8$ it is a maximal subgroup.

6.2 Representations of $\mathrm{SL}_2(\mathbb{F})$

Let $\mathbb{F}$ be a field of characteristic $p \geq 0$ and V be the vector space of homogeneous polynomials in two variables x, y, of degree $d-1$, over $\mathbb{F}$. Every matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{F})$$

acts in a natural way on the basis $\mathcal{B} = \left\{x^{d-1}, x^{d-2}y, \ldots, y^{d-1}\right\}$ of V, via:

$$x^i y^j \mapsto (a_{11}x + a_{21}y)^i (a_{12}x + a_{22}y)^j .$$

Call $\alpha : V \to V$ the extension by linearity of this action. The homomorphism

$$h_d : \mathrm{SL}_2(\mathbb{F}) \to \mathrm{SL}_d(\mathbb{F}) \tag{6.3}$$

such that each $A \in \mathrm{SL}_2(\mathbb{F})$ maps to the matrix of α with respect to $\mathcal{B}$, is a representation of degree d of $\mathrm{SL}_2(\mathbb{F})$. This representation is absolutely irreducible whenever $0 < d \leq p$ (see also [3]). When d is even and $\mathbb{F} = \mathbb{F}_q$, with q appropriate, it gives rise to maximal subgroups of $\mathrm{Sp}_d(q)$.

(6.4) Example *For $d = 4$, the homomorphism $h_4 : \mathrm{SL}_2(\mathbb{F}) \to \mathrm{SL}_4(\mathbb{F})$ acts as:*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a^3 & a^2b & ab^2 & b^3 \\ 3a^2c & a^2d + 2abc & 2abd + b^2c & 3b^2d \\ 3ac^2 & 2acd + bc^2 & ad^2 + 2bcd & 3bd^2 \\ c^3 & c^2d & cd^2 & d^3 \end{pmatrix}. \tag{6.5}$$

7 Exercises

(7.1) Exercise *Let W and W' be subspaces of $\mathbb{F}^n$. Show that there exists $g \in \mathrm{SL}_n(\mathbb{F})$ such that $gW = W'$ if and only if they have the same dimension.*

(7.2) Exercise *In $\mathrm{Mat}_3(7)$ find a field of order 7^3, its centralizer and its normalizer.*

(7.3) Exercise *Show that the representation (6.5) fixes a symplectic form.*

(7.4) Exercise *Write explicitly an absolutely irreducible representation of $\mathrm{SL}_2(7)$ of degree 6, fixing a symplectic form.*

References

[1] M.Aschbacher, Finite group theory, Cambridge University Press, 1986.

[2] J.Bray, D. Holt and C. Roney Dougal, The Maximal Subgroups of the Low-Dimensional Finite Classical Groups, London mathematical Society Lecture Note Series: 407, Cambridge University Press, 2013.

[3] R.Burkhardt, Die Zerlegungsmatrizen der Gruppen PSL$(2, p^f)$, J.Algebra **40**, 75–96 (1976).

[4] T.Burness, Simple groups, fixed point ratios and applications, http://seis.bristol.ac.uk/ tb13602/epfl.

[5] R.W.Carter, Simple groups of Lie type, John Wiley and sons, 1972.

[6] C.W.Curtis, I.Reiner, Representation Theory of Finite Groups and Associative Algebras, Wiley-Interscience, 1962.

[7] J.D.Dixon, The structure of Linear Groups, Van Nostrand Reinhold Company, 1971.

[8] B.Hartley, T.O.Hawkes, Rings, Modules and Linear Algebra, Chapman and Hall, 1970.

[9] I.N.Herstein, Algebra, Editori Riuniti, 1982.

[10] J.E.Humphreys, Introduction to Lie Algebras and Representation Theory, Second Printing Revised, Springer-Verlag, 1972.

[11] B.Huppert, Endliche Gruppen I, Springer Verlag, 1983.

[12] M.Isaacs, I.M. Algebra: a graduate course, Brooks/Cole Publishing Company, 1994.

[13] M.Isaacs, Character Theory of finite groups, Academic Press, 1976.

[14] N.Jacobson, Basic Algebra I, W.H.Freeman and company, San Francisco,1974.

[15] P.Kleidman and M.Liebeck, *The subgroup structure of the finite classical groups*, London Mathematical Society Lecture Note Series, 129. Cambridge University Press, 1990.

[16] S.Lang, Linear Algebra, Addison Wesley, 1966.

[17] S.Lang, Algebra, Revised third Edition, Springer, 2002.

[18] R.Ree, On some simple groups defined by Chevalley, Trans. Amer. Math. Soc., **84** (1957), 392-400.

[19] D.J.Robinson, A course in the Theory of Groups, Springer, 1982.

[20] M.C.Tamburini, Algebra 1, Algebra 2, Approfondimenti di Algebra, Istituzioni di Algebra Superiore. http://docenti.unicatt.it/ita/maria_clara_tamburini_ bellani/ Oppure a richiesta: mariaclara.tamburini@gmail.com

[21] D.Taylor, The geometry of classical groups, Heldermann Verlag, 1992.

[22] H.Wielandt, Finite Permutation Groups, Academic Press, 1964.