

Linear and Bilinear Geometry

Yves Laszlo

Yves.Laszlo@universite-paris-saclay.fr

Beta version of January 7, 2025 with typos and mistakes



Contents

1	Introduction	13
1.1	Conventions	14
1.2	Prerequisites	14
1.3	Complement: Zorn's Lemma and application	15
I	Reduction of endomorphisms	17
2	Warm-up: review on basic linear algebra	19
2.1	Perspective	19
2.2	Euclidean plane	20
2.2.1	Euclidean Norm	20
2.2.2	Non oriented angle of pair of vectors or lines	21
2.2.3	Orthogonality in oriented Euclidean planes	22
2.2.4	Oriented angles of vectors	23
2.2.5	Isometries	24
2.2.6	Symmetric real matrices	26
2.3	General linear maps of the plane	26
2.3.1	Minimal polynomial	27
2.3.2	Cyclic vectors	27
2.4	Reminder on Gauss elimination method	27
2.4.1	Review of Transvections	28
2.4.2	Normal subgroups of $GL(V)$	29
2.4.3	Supplementary exercises	31
3	Generalities on modules	33
3.1	Perspective	33
3.2	Vocabulary and first examples	34
3.2.1	Quotient, cokernel	37
3.2.2	Properties to handle with caution	38

3.2.3	Cyclic modules	40
3.2.4	The $\mathbf{k}[T]$ -module V_a	41
3.3	Exact sequences and diagrams	42
3.3.1	Exact sequences	42
3.3.2	A fundamental exact sequence	43
3.3.3	Commutative diagrams	43
3.4	Functoriality and diagram chasing	44
3.5	Universal properties	47
3.5.1	Sum and product	48
3.5.2	Kernel and cokernel	48
3.6	Quotient rings	50
3.7	A variant of the Chinese remainder theorem	51
3.8	Additional Exercises	53
4	Equivalence Classes in $M_{p,q}(\mathbf{k}[T])$.	59
4.1	Perspective	59
4.2	Introduction	59
4.3	Elementary Divisors	60
4.3.1	Existence	60
4.3.2	What Uniqueness?	61
4.3.3	Equivalence Classes in $M_{p,q}(\mathbf{k}[T])$	64
4.4	Supplementary Section: Insight into K-Theory	64
4.5	Additional Exercises	65
5	Similarity classes of $M_n(\mathbf{k})$	67
5.1	Point of view	67
5.2	Introduction	68
5.2.1	Notations	69
5.3	Strategy	70
5.4	Invariance by \sim of $T \text{Id} - A$ of V_a	71
5.5	Similarity Invariants of $a \in \text{End}_{\mathbf{k}}(V)$	72
5.6	Calculation of V_a and Applications	74
5.7	Diagonalization	75
5.8	Cyclic Endomorphisms	76
5.9	Frobenius Decomposition	77
5.9.1	Equivalent Formulation	78
5.10	Summary	79
5.11	Application: Commutant	79

<i>CONTENTS</i>	5
5.12 Application: Jordan Reduction	80
5.12.1 Examples	82
5.12.2 Supplement on nilpotent matrices	84
5.13 Appendices	85
5.13.1 An algorithm from \sim to \approx	85
5.13.2 Jordan reduction by duality of nilpotents without modules	86
5.13.3 Frobenius decomposition without modules	87
5.14 Implementations in Sage	88
5.14.1 Elementary Divisor Calculations	88
5.14.2 Jordan-Chevalley Decomposition	89
5.15 Additional Exercises	91
6 Semisimplicity in $M_n(\mathbf{k})$	93
6.1 Perspective	93
6.2 Semisimplicity	94
6.2.1 General Semisimple Modules	94
6.2.2 Semisimple Modules over Principal R	95
6.2.3 «Reminder» on Perfect Fields	96
6.2.4 Criterion for Semisimplicity of V_a	97
6.3 Jordan-Chevalley Decomposition	98
6.3.1 Hensel's Lemma and Existence	98
6.3.2 Uniqueness	99
6.3.3 Similarity class of the components	100
6.3.4 Appendix: What about the algorithmic nature of the decomposition?	100
6.4 Additional Exercises	102
7 Reminder on Duality in Finite Dimension	103
7.1 Basic notions	103
7.2 Motivation	105
7.3 Formal Biorthogonality	105
7.4 Ante-dual Basis: Biduality	105
7.5 Orthogonal and Polar in Finite Dimension	106
7.6 Biduality Conventions (Finite Dimension)	107
7.7 Contravariance	108
7.7.1 Review of Transvections	109
7.8 Additional Exercises	110

8	Stable Subspaces	111
8.1	Perspective	111
8.2	Generalities	112
8.3	Characteristic Subspaces	112
8.3.1	Topological properties in the complex case	114
8.3.2	d -th roots in GL_n	116
9	Topology of Similarity Classes*	119
9.1	Perspective	119
9.2	Introduction	119
9.3	Closure of a Nilpotent Orbit	120
9.3.1	Order and Duality on Partitions	121
9.3.2	Rank and Nilpotent Orbits	123
9.3.3	A Nilpotent Matrix Deformation	123
9.4	Closure of an Arbitrary Orbit	124
9.5	Additional Exercises	125
II	Useful general algebra	127
10	Finiteness Properties of Modules	129
10.1	Introduction	129
10.2	Integrality	129
10.2.1	Principle of Extension of Algebraic Identities	130
10.2.2	An Application of Cayley-Hamilton	130
10.2.3	Rings of Integers	131
10.3	Noetherian Modules	131
10.3.1	Stability under exact sequences	133
10.3.2	Existence of Decomposition into Irreducibles in Noetherian Domains	133
10.3.3	Hilbert's Basis Theorem	134
10.4	Additional Exercises	134
11	Reminder on Unique Factorization Domains	137
11.1	Introduction	137
11.2	Characterization	138
11.2.1	Uniqueness Criterion	138
11.3	Transfer	140
11.3.1	GCD, LCM in UFD	140
11.3.2	Content	141

11.3.3	The Transfer Theorem	142
11.4	Irreducibility of the Cyclotomic Polynomial Over \mathbf{Q}	142
11.5	Additional exercises	145
III	Metrics on Real and Complex Vector Spaces	147
12	Euclidean Spaces	149
12.1	Perspective	149
12.2	Basics on Euclidean Geometry	150
12.2.1	Examples	150
12.2.2	Euclidean Norm	150
12.2.3	Dual of an Euclidean space, Orthogonal	151
12.2.4	Cross product in an oriented Euclidean space	153
12.2.5	Orthogonalization	154
12.2.6	Gram matrices	156
12.2.7	Minimization of distance	157
12.3	Geodesic distance on the Euclidean sphere	158
12.4	Adjoint morphism	160
12.5	Comparison	161
12.6	Real Normal Endomorphisms	161
12.6.1	Reduction of Normal Real endomorphisms	161
12.6.2	Reduction of Orthogonal Morphisms	163
12.6.3	Application to Real Self-adjoint endomorphisms	165
12.6.4	Ellipsoid	167
12.6.5	Application to Real Skew-adjoint Endomorphisms	168
12.7	Additional Exercises	169
13	Euclidean Geometry	173
13.1	Perspective	173
13.2	Topological Properties of the Orthogonal Group	174
13.3	Study of \mathcal{S}_n^{++}	174
13.4	Loewner Ellipsoid	176
13.5	Compact subgroups of $\mathrm{GL}_n(\mathbf{R})$	177
13.6	Polar Decomposition.	178
13.7	Algebraic Properties of $\mathrm{O}_n(\mathbf{R})$	179
13.8	Euclidean Similitude	181
13.9	Additional Exercises	182

14 Quaternion Algebra and Euclidean Geometry in dimension ≤ 4	185
14.1 Perspective	185
14.2 Construction	185
14.3 Imaginary Quaternions	187
14.3.1 Quaternions and $SO_3(\mathbf{R})$	188
14.3.2 Cross product and Rodrigues formula, tbd	189
14.3.3 Quaternions and $SU_2(\mathbf{R})$	189
14.3.4 Quaternions and $SO_4(\mathbf{R})$	189
14.4 Spin	190
14.5 Additionnal Exercises	190
15 Finite Groups of Euclidean Isometries in Dimension ≤ 3	191
15.1 Perspective	191
15.2 Subgroups of $O_2(\mathbf{R})$	191
15.3 Subgroups of $SO_3(\mathbf{R})$	193
15.3.1 Normal Subgroups	196
15.4 Appendix: Local Extrema	197
16 Convexity in Euclidean Spaces	199
16.1 Perspective	199
16.2 Generalities	199
16.2.1 Definitions	199
16.2.2 Convex hull	200
16.2.3 Topology of a convex set	202
16.3 The Farkas Lemma	204
16.4 Projection to a Closed Convex Set	206
16.5 Krein-Milman Theorem	209
16.6 Polar Dual of a Convex Body	210
16.7 Additional Exercises	212
17 Complex Hermitian Spaces	215
17.1 Perspective	215
17.2 Basics on Hermitian Geometry	216
17.2.1 Examples	216
17.2.2 Hermitian Norm	217
17.2.3 Dual of an Hermitian space, Orthogonal	217
17.2.4 Orthogonalization	218
17.2.5 Gram matrice, Minimization of distances	218
17.3 Adjoint morphism	219

17.4 Complex Normal endomorphisms 220

 17.4.1 Reduction of Complex Normal endomorphisms 220

 17.4.2 Reduction of Unitarian endomorphisms 221

 17.4.3 Reduction of Hermitian endomorphisms 222

17.5 Topological Properties of the Unitary Group 224

 17.5.1 Study of \mathcal{H}_n^{++} 224

17.6 Compact subgroups of $GL_n(\mathbf{C})$ 224

 17.6.1 Complex Polar Decomposition. 225

17.7 Additional Exercises 225

IV Geometries 227

18 A primer in Projective Geometry 229

18.1 Perspective 230

18.2 Introduction 230

18.3 Topology of real or complex projective space 231

18.4 Algebraic and Geometric descriptions 232

 18.4.1 Homogeneous coordinates 232

 18.4.2 Affine charts 233

 18.4.3 Lifting of affine isomorphisms, tbd 235

18.5 The Fundamental theorem of Projective Geometry 235

 18.5.1 Statement of the main theorem 235

 18.5.2 The dimension 2 case 236

 18.5.3 The general case 238

 18.5.4 The dimension 1 case 240

18.6 Reminder on Affine Geometry 241

 18.6.1 Universal vector envelop of an affine space 242

19 Sesquilinear Forms and Projective Geometry 245

19.1 Perspective 245

19.2 Introduction 246

 19.2.1 Notations and Reminders 246

19.3 Sesquilinear Forms 246

 19.3.1 Non-degenerate Forms 247

19.4 (Left) Orthogonality 248

 19.4.1 Adjoint 249

20 ε-Hermitian Forms	251
20.1 Perspective	251
20.2 Introduction	251
20.3 Definitions	252
20.4 Orthogonality	253
20.5 Alternating Forms	255
20.5.1 Classification	255
20.5.2 Pfaffian	256
20.6 Supplementary exercises	257
21 Quadratic Forms	259
21.1 Perspective	259
21.2 Polar Form	260
21.3 Orthogonal Bases	262
21.4 Quadratic Spaces	264
21.5 Anisotropic spaces	265
21.6 Invariants of Quadratic Forms	265
21.7 Isotropy and Index	266
21.8 Classification over an algebraically closed field	269
21.9 Classification over \mathbf{R}	269
21.10 Conics and Quadrics in \mathbf{R}^2 and \mathbf{R}^3 , Ellipsoid	270
21.11 Classification over finite fields.	271
21.12 Witt's Extension Theorem	272
21.13 Appendix: Quadratic Pencils	273
21.14 Supplementary Exercises	275
22 The general orthogonal group	277
22.1 Perspective	277
22.2 Definition	278
22.3 The case of dimension 2	278
22.4 Orthogonal Symmetries	279
22.5 Orthogonal Similitude	280
22.6 Generators of the orthogonal group	281
23 Automorphisms of classical groups	285
23.0.1 Perspective	285
23.1 Automorphisms of $O_n(\mathbf{R})$	286
23.1.1 Involutions	286
23.1.2 The main theorem	288

<i>CONTENTS</i>	11
23.2 Automorphisms of $GL(V)$	290
23.2.1 Involutions	290
23.2.2 Pairs of extremal involutions	291
23.2.3 Proof of the theorem	295
23.2.4 Automorphisms of \mathbf{k}^*	297
23.2.5 Normal subgroups of $GL(V)$	298
23.2.6 Additional exercises	299
V Supplements	301
24 Bilinear forms without symmetries	303
24.1 Perspective	303
24.2 Introduction	303
24.3 Existence of a Decomposition	304
24.4 The Typical Bilinear Space	306
24.5 Uniqueness	307
24.6 Classification: Algebraically Closed Case	308
25 Index et bibliography	313

Chapter 1

Introduction

In 1872, Felix Klein posed the following question. "Given a multiplicity and a group, to study the beings from the point of view of properties that are not altered by the transformations of the group... this can also be expressed as follows: given a multiplicity and a transformation group; develop the theory of invariants relative to this group" ([23]).



Felix Klein

In these notes on vector, quadratic, and Hermitian geometry, we illustrate this visionary viewpoint by classifying geometric objects via invariants under various group actions (invariant factors, similarity invariants, discriminant, index, signature...).

We strive to do so in a *concrete* manner, i.e., with methods that lead to algorithms. It is indeed better to know how to construct an object than to simply know of its existence. The aim of the course, however, is not to provide optimized programs in terms of efficiency (that's another subject, and interesting at that!), but to explore the *how-to*. One quickly encounters the numerical flaws of typical pivot algorithms. It is not, however, about giving formally constructivist methods ([6]) but about providing as much as possible existence theorems that can explicitly lead to the construction of the object in question, for example, through a computer.

The material of this book is more or less classical, only the perspective being somehow more original. The titles of the (few) chapters whose content is less classical are followed by an asterisk *. For the convenience of the reader, we have included some wellknown general results in the chapters beginning by "Reminder of...".

We strongly advise the reader to implement the various algorithms on a machine: this will allow them to verify that they have thoroughly understood the proofs. On our part, we have used the SAGEMATH

program, based on Python.

I extend my warm thanks to Peter Haïssinki who kindly provided his beautiful notes on the quadratic part, notes on which I relied heavily for a first version of the text, and to Olivier Debarre for his examples of endomorphism reduction.

Photo credits: ChronoMaths, Flickr user Duncan, Patrick Fradin, Marcel Gotlib, UQAM, Wikipedia.

1.1 Conventions



Unless expressly stated otherwise, the rings are assumed to be commutative and with an identity, generally denoted R . They are assumed, unless explicitly stated otherwise, to be non-zero, i.e., $1 \neq 0$. Their multiplicative group of units is denoted R^\times .

This grants them the following property: Every ring admits a proper maximal ideal for inclusion, a result we shall consider as an axiom (in this generality, this is equivalent to the axiom of choice).

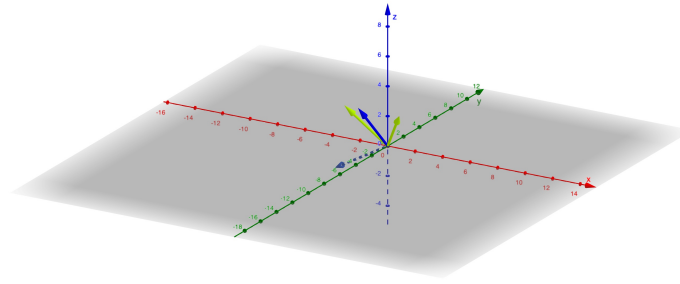
Otherwise, the reader will easily demonstrate this by applying Zorn's Lemma to the set of proper ideals of R (1.3). In practice, it can often be dispensed with if one really insists. Naturally, it will only be used for existence theorems: it has no algorithmic value. Zorn's Lemma also allows us to demonstrate, essentially formally, that, just as \mathbf{Q} is contained in \mathbf{C} , any field \mathbf{k} is contained in an algebraically closed field Ω .

It will be used without further specification. The key to this result is the elementary fact that every polynomial with coefficients in \mathbf{k} has a root in some possibly larger field K . The existence of Ω then formally follows from the existence of maximal ideals in any non-zero rings. However, readers who dislike the axiom of choice will check that the existence of the aforementioned fields K suffices for us and that the existence of Ω is just a convenience of language, in fact.

1.2 Prerequisites

No other knowledge of linear algebra is assumed beyond the basics of dimension theory, the relationship between matrices and endomorphisms, and the elementary properties of the determinant (notion of characteristic polynomial and eigenvalue included). The reader is assumed to be familiar with the Gauss elimination method. Readers who have studied the theory in the context of real or complex vector spaces will make an effort to accept (or verify) that nothing changes on an arbitrary field.

In general, it is recalled that line and column operations on rectangular matrices with coefficients in a ring R are obtained by multiplication on the right or left by transvections $T_{i,j}(r) = \text{Id} + rE_{i,j}$, $i \neq j$ (where $E_{i,j}$ is the standard square matrix with all coefficients zero except the one at row i and column j , which is 1), line or column permutations by permutation matrices M_σ (defined by $M_{i,j} = \delta_{i,\sigma(j)}$ for

Transvection $T_{1,2}(2)$

every permutation $\sigma \in S_n$ ¹, these matrices being invertible (of determinant ± 1). The multiplication of a principal pivot by a scalar r is achieved by product with a elementary dilatation $D(r) = \text{Id} + (r - 1)E_{1,1}$, which is invertible as long as r is. Geometrically, both transvections and dilatations add to a given vector $\sum x_j e_j$ the vector of constant direction e_i with "algebraic length" a constant multiple of x_j .

From a general point of view, the reader is assumed to be familiar with the general definitions of rings, ideals. . . . More specifically, besides the notion of a field, the notion of a principal ring (integral with all ideals generated by one element), at least in the case of \mathbf{Z} and $\mathbf{k}[T]$, is assumed known. To make reading easier, a proof of the main results will be given in the chapter on factorial rings (11). For the most part, we will use two things: Bézout's identity and the fact that a principal ring is factorial (11.2.1.6) (existence and uniqueness, apart from order, of decomposition into irreducible factors), which allows us to relate the notion of GCD both to the decomposition into irreducible factors and to Bézout's identity. For convenience of the reader, we recall the notion of quotient (3.6).

1.3 Complement: Zorn's Lemma and application

Let E be a (partially) ordered set. We can think, for example, of the set of subsets of a given set ordered by inclusion. But there are many other examples.

Definition 1.3.0.1. *We say that E is inductive if every non-empty totally ordered part has an upper bound in E .*

Example(s) 1.3.0.2. *\mathbf{R} equipped with the usual order relation is not inductive. Similarly, the set of intervals $[0, x]$, $x \in \mathbf{R}$ ordered by inclusion is not inductive. On the other hand, the set of subsets of a set ordered by inclusion is inductive.*

¹where $\delta_{i,j}$ is the Kronecker symbol equal to 1 if $i = j$ and 0 if not.



Max Zorn

Lemma 1.3.0.3 (Zorn's lemma). *Every non-empty inductive set has a maximal element.*

This lemma can be seen as an axiom of set theory, in fact equivalent to the axiom of choice: if (E_i) is a non-empty family of sets, then $\prod E_i$ is non-empty. We will consider it as such.

Corollary 1.3.0.4. *Every non-zero ring has a maximal ideal. More generally, every proper ideal of a ring is contained in a maximal ideal.*

Proof. Let E be the family of proper ideals of A containing a given proper ideal J (for instance $J = \{0\}$ because our rings are nonzero). Because J is proper, E is non-empty. Obviously, E is inductive: the union of a totally ordered family of proper ideals is still a proper ideal, which is an upper bound. Zorn's lemma finishes the job. \square

Part I

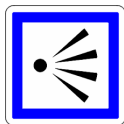
Reduction of endomorphisms

Chapter 2

Warm-up: review on basic linear algebra



2.1 Perspective



The purpose of this introductory chapter is to prove the main theorems of Euclidean and general linear geometry in the real plane E . Our motivation is twice. First to refresh general linear algebra knowledge in this elementary context. Second, more fundamentally, to emphasize that almost all problems of linear algebras appear in dimension ≤ 2 . We'll see in many occasions that the general case follows from this small dimension study. In fact this simple observation is quite deep as the reader will see in the next coming years, for instance if he has to look at the theory of Lie or algebraic groups where the role of the 2 by 2 matrices of SL_2 is crucial.

2.2 Euclidean plane

We start with a "physical" perspective, namely we assume that our real plane E ($n = \dim(E) = 2$) has a metric, meaning a scalar product

$$\begin{cases} E \times E & \rightarrow & \mathbf{R} \\ (v, w) & \mapsto & \langle v, w \rangle \end{cases}$$

Recall that this means that this map is linear in each variable and positive definite (or > 0 for short): $q(v) = \langle v, v \rangle > 0$ unless $v = 0$.

Definition 2.2.0.1. *A Euclidean space is a real finite-dimensional vector space equipped with a scalar product. An isometry of Euclidean spaces is a linear isomorphism preserving the scalar products. An isometric endomorphism of positive determinant is called a rotation.*

Of course the typical examples are $E = \mathbf{C}$ with

$$\langle z, z' \rangle = \operatorname{Re}(\bar{z}z')$$

or \mathbf{R}^2 endowed with the standard scalar product

$$\langle (v_1, v_2), (w_1, w_2) \rangle = v_1w_1 + v_2w_2,$$

both being canonically isomorphic.

The set of isometries (resp. rotations) is a subgroup $O_2(E)$ of $GL_2(E)$ (resp. $SO_2(E)$ of $SL_2(E)$)¹.

2.2.1 Euclidean Norm

Proposition 2.2.1.1 (Cauchy-Schwartz). *Let $v, w \in E$ and let us write $\|v\| = \sqrt{\langle v, v \rangle}$.*

1. *One has $\langle v, w \rangle \leq \|v\|\|w\|$ with equality if and only if v, w are positively colinear.*
2. *One has $|\langle v, w \rangle| \leq \|v\|\|w\|$ with equality if and only if v, w are colinear.*

Proof. We may assume v and w are non-zero. The Cauchy-Schwartz inequality (1) is nothing but the inequality

$$2 - 2\langle v/\|v\|, w/\|w\| \rangle = q(v/\|v\| - w/\|w\|) \geq 0$$

with equality if and only if $v/\|v\| - w/\|w\| = 0$, namely if v, w are positively colinear. We get (2) from (1) changing w in $-w$. □

¹As usual, we'll simply write $O_2(\mathbf{R})$ (resp. $SO_2(\mathbf{R})$) for $O_2(E)$ (resp. for $SO_2(E)$) when E is the standard Euclidean plane \mathbf{R}^2

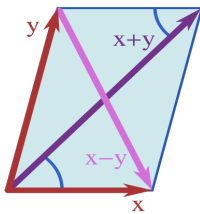
Theorem 2.2.1.2. *The mapping $v \mapsto \|v\|$ is a norm called the Euclidean norm.*

Proof. We define, for $v \in E$, $\|v\| = \langle v, v \rangle$. As q is positive definite, to show that $\|\cdot\|$ is a norm, it suffices to verify the triangle inequality

$$\begin{aligned} (\|v\| + \|w\|)^2 - \|v + w\|^2 &= \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 - \|v\|^2 - 2\langle v, w \rangle - \|w\|^2 \\ &= 2\|v\|\|w\| - 2\langle v, w \rangle \\ \text{(by Cauchy-Schwartz)} &\geq 0 \end{aligned}$$

□

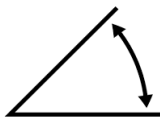
One immediately checks the important property of the Euclidean norm: the median equality



$$\text{For any } x, y \in E, \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

2.2.2 Non oriented angle of pair of vectors or lines

By Cauchy-Schwartz inequality, the absolute value of the scalar product of two unit vectors is ≤ 1 therefore can define the angle $\widehat{(v, w)}$ between two nonzero vectors v, w by the formula



$$\widehat{(v, w)} = \arccos \left\langle \frac{v}{\|v\|}, \frac{w}{\|w\|} \right\rangle$$

thought as an element of $\mathbf{R}/2\pi\mathbf{Z}$ defined **up to sign**.

Thanks to trigonometry formulae, we obtain the usual formula from elementary geometry (the Chasles formula)

$$\widehat{(v_1, v_2)} + \widehat{(v_2, v_3)} = \widehat{(v_1, v_3)}.$$

Of course, the parity of the arccos function and the homogeneity of the scalar product ensures that the non oriented angle of two non zero vector neither depends on their order or on any nonzero multiple of them. This allows to define the (non oriented) angle of two lines ℓ_1, ℓ_2 by the non oriented angle of any vector basis of them, no matter the order of the lines.

Remark(s) 2.2.2.1. *Rather than "angle" we should have said "measure of the angle" in an Euclidean plane (see 2.2.5.6).*

2.2.3 Orthogonality in oriented Euclidean planes

If ℓ is a line (dimension $d = 1$), its orthogonal ℓ^\perp has equation $\langle \cdot, v \rangle = 0$ for any chosen basis v of ℓ and therefore has dimension $\dim(\ell^\perp) = n - d = 1$ (see 12.2.3 for the general case).

Remark(s) 2.2.3.1. *Let us recall that two bases of some finite dimensional vector space define the same orientation if the determinant of the base change matrix is > 0 . An orientation is then defined by a basis defined up to the action of the group of matrix of positive determinant $\mathrm{GL}_+(\mathbf{R})$. These bases are said positively oriented or direct.*

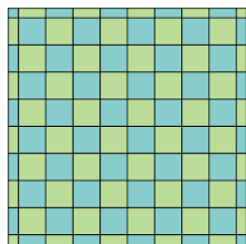
For instance, if we change the order of a basis of the plane, we change the orientation of the plane. Therefore, given a normed vector v of an oriented Euclidean plane, there exists a unique positive orthonormal basis of the plane (v, w) .

Notice that $\mathrm{GL}_+(\mathbf{R})$ is connected (13.6.0.2). It follows that orientation is the only way to assign a continuous sign to any basis of \mathbf{E} .

Because a line has obviously only two opposite normed vectors, we get just like in high school

Proposition 2.2.3.2. *Let \mathbf{E} be an oriented Euclidean plane. For any normed vector $v \in \mathbf{E}$, there exists a unique normed vector v^\perp such that (v, v^\perp) is a positively oriented orthonormal basis.*

In the standard Euclidean plane \mathbf{R}^2 with the usual orientation defined by the canonical basis, we have explicitly for $v = (a, b)$, $a^2 + b^2 = 1$ the usual formula $v^\perp = (-b, a)$.



We indeed have defined an algorithm, which will be heavily generalized: if we start with an arbitrary basis (v_1, v_2) of \mathbf{E} , there exists a unique orthonormal basis $(e_1 = v_1/\|v_1\|, e_2 = e_1^\perp)$ such that $e_1 \in \mathbf{R}v_1$ and $(e_2, v_2) > 0$: this is the Gram-Schmidt process in the plane (see 12.2.5.1 in general).

The following statement is well-known and useful.

Proposition 2.2.3.3. *1. A morphism of Euclidean spaces (of any dimension) is an isometry (resp. a rotation) if and only if it maps an orthonormal (resp. direct orthonormal) basis to an orthonormal (resp. direct orthonormal) basis.*

2. An endomorphism f of an Euclidean space (of any dimension) is an isometry if and only if its matrix M with respect to (any) orthonormal basis satisfies ${}^tMM = \mathrm{Id}$

3. The determinant of an isometry is ± 1 . The determinant of a rotation is $+1$.

Proof. We assume the existence of orthonormal basis for granted in general (see 12.2.5.1). (1) is a direct consequence of the bilinearity of the scalar product.

(2) If (e_i) is our orthonormal basis, one has f isometry if and only if

$$(\text{Id})_{i,j} = \delta_{i,j} = \langle f(e_i), f(e_j) \rangle = \langle \sum_a m_{a,i} e_a, \sum_b m_{b,j} e_b \rangle = \sum_a m_{a,i} m_{a,j} = ({}^t\text{MM})_{i,j}$$

proving (2).

(3) Follows from (2) and the multiplicativity of the determinant. □

We get the well-known formula

$$\text{SO}_n(\mathbf{R}) = \{M \mid {}^t\text{MM} = \text{Id} \text{ and } \det(M) = 1\}$$

Because the base change morphism between two orthonormal bases is an isometry, we get

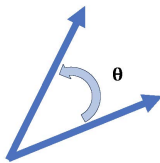
Corollary 2.2.3.4. *Two Euclidean planes are (non canonically) isomorphic.*

2.2.4 Oriented angles of vectors

Let E be an oriented Euclidean plane. Using the above results, we can define the oriented angle of two non zero vectors v, w as follows. If v, w are normed, one has a unique writing $w = av + bv^\perp$ with $a^2 + b^2 = 1$. Therefore, there exists a unique $\widehat{(v, w)} \in \mathbf{R}/2\pi\mathbf{Z}$ such that

$$(a, b) = (\cos(\widehat{(v, w)}), \sin(\widehat{(v, w)}))$$

Because $\langle w, v \rangle = a$, one has $\widehat{(v, w)} = |\widehat{(v, w)}|$.



In the general case, one defines $\widehat{(\frac{v}{\|v\|}, \frac{w}{\|w\|})} \in \mathbf{R}/2\pi\mathbf{Z}$.

Remark(s) 2.2.4.1. *By construction, if θ is the oriented angle between two normed vectors v, w , the base change matrix from (v, v^\perp) to (w, w^\perp) is $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$. The addition formulas for the trigonometric functions \sin, \cos give the important formula*

$$R_\theta \circ R_{\theta'} = R_{\theta+\theta'}$$

Of course, we again obtain the usual formula of elementary geometry like the Chasles formula

$$\widehat{(v_1, v_2)} + \widehat{(v_2, v_3)} = \widehat{(v_1, v_3)}.$$

2.2.5 Isometries

Let E be an oriented Euclidean plane.

Proposition 2.2.5.1. *Let v, w be two normed vectors and $\theta = \widehat{(v, w)}$.*

1. *There exists a unique rotation ρ_θ mapping v to w whose matrix in any direct orthonormal basis is R_θ .*
2. *One has*

$$\cos(\widehat{(v, w)}) = \langle w, \rho(w) \rangle = \cos(\theta) = \frac{\text{tr}(\rho_{v,w})}{2}.$$

Proof. (1) The base change morphism from (v, v^\perp) to (w, w^\perp) is definitely a positive isometry, that is a rotation ρ giving the existence. Conversely any isometry mapping v to w maps v^\perp to $\pm w^\perp$ and therefore to w^\perp if it is positive giving the uniqueness. The matrix of ρ_θ in (v, v^\perp) is R_θ (cf. (2.2.4.1)). If $\mathcal{B} = (v_1, v_2)$ is another direct orthonormal basis, the base change matrix from (v, v^\perp) to \mathcal{B} is R_α (2.2.4.1). Therefore

$$\text{Mat}(\mathcal{B}, \rho) = R_\alpha^{-1} \circ R_\theta \circ R_\alpha = R(-\alpha + \theta + \alpha) = R_\theta$$

proving (1).

Let us chose any orientation on E . By (2.2.3.2), one can assume $v = e_1$ is the first vector of an orthonormal basis (e_1, e_2) . Because w is a unit vector, it can be written as $w = \cos(\theta)e_1 + \sin(\theta)e_2$ for a uniquely defined $\theta \in \mathbf{R}/2\pi\mathbf{Z}$. But $w, w' = -\sin(\theta)e_1 + \cos(\theta)e_2$ is the unique direct orthonormal basis with first vector w . Therefore the endomorphism ρ mapping (e_1, e_2) to (w, w') is the unique relevant positive isometry.

(2) follows directly from the proof of (1). □

To specify the structure of isometries, let us choose a direct orthonormal basis \mathcal{B} of E . We will identify any endomorphism f with its matrix in \mathcal{B} .

Corollary 2.2.5.2. 1. *The map $\theta \mapsto \rho_\theta$ defines an isomorphism*

$$\mathbf{R}/2\pi\mathbf{Z} \simeq \text{SO}(E)$$

2. *ρ_θ is complex diagonalizable with complex eigenvalues are $\exp(\pm i\theta)$.*

3. ρ_θ is real diagonalizable if and only if $\theta \equiv 0 \pmod{2\pi}$ or $\theta \equiv \pi \pmod{2\pi}$ that is to say it is equal $\rho_\theta = \pm \text{Id}$.

4. The matrices negatives isometries are orthogonal symmetries.

Proof. Only the last point has not be proven yet. Let $\mathcal{B} = (e_1, e_2)$ be a direct orthonormal basis and $S_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ be the matrix of the orthogonal symmetry along the (second) diagonal $\mathbf{R}(e_1 + e_2)$. Then, for any negative isometry, the product of S_0 by its matrix S is some rotation $S_0S = R_\theta$. We get

$$R = S_0R_\theta = \begin{pmatrix} \sin(\theta) & \cos(\theta) \\ \cos(\theta) & -\sin(\theta) \end{pmatrix}$$

whose square is Id by direct calculation. □

From this, one recover any elementary facts about plane isometries known for the highschool time (see 12.6.2.3 in the general case).

Remark(s) 2.2.5.3. *If one prefers the identification $E \sim \mathbf{C}$ with its orthogonal basis $(1, \mathbf{i})$, the corresponding statement is that rotations are as usual of the form $\theta \mapsto \exp(\mathbf{i}\theta)z$ and symmetries of the form $\theta \mapsto \exp(\mathbf{i}\theta)\bar{z}$.*

Exercise(s) 2.2.5.4. *Show that the application which associates to an an orthogonal symmetry its invariant vector line is a bijection from the set of symmetries onto the set of vector lines. Show that the compound of two symmetries associated with two lines making a (non-oriented) angle θ is a rotation whose (non-oriented) angle is 2θ .*

Exercise(s) 2.2.5.5. *Determine the real and complexe eigenvalues and th corresponding eigenspaces of any planar isometry. When are they diagonalizable over \mathbf{R} ? Over \mathbf{C} ?*

Remark(s) 2.2.5.6. *We could have defined an oriented angle in a non oriented plane as the former rotation itself. The value of the angle would then have been in $\text{SO}_2(\mathbf{R})$. The link between the our definition is that the choice of an orientation define a canonical isomorphism $\text{SO}_2(E) \simeq \mathbf{R}/2\pi\mathbf{Z}$, recovering our notion of angle which could be in this context be defined as the measure of the angle. But the usual modern point of view is to see an angle as we did, and therefore we have to choose an orientation of the plane.*

2.2.6 Symmetric real matrices

We know (2.2.5.1) that the matrices of a negative isometries in an orthonormal basis are of the form $\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$, in particular are symmetric. Like all symetries, they are diagonalizable with spectrum $\{\pm 1\}$. But, we have more. The eigenspaces are orthogonal. Indeed, if we identify E with \mathbf{C} thanks to \mathcal{B} , our symmetry is nothing but $z \mapsto \exp(\mathbf{i}\theta)\bar{z}$ whose (real) $+1$ -eigenspace is the line $\mathbf{R}\exp(\mathbf{i}\theta/2)$ and (real) -1 -eigenspace is the orthogonal line $\mathbf{iR}\exp(\mathbf{i}\theta/2)$. We recover the well known fact that orthogonal symmetries are orthogonally diagonalizable. This fact is general.

Proposition 2.2.6.1. *Symmetric matrices of $M_2(\mathbf{R})$ are exactly orthogonally diagonalizable matrices (with respect to the standard Euclidean structure of \mathbf{R}^2).*

Proof. We identify E with the standard Euclidean plan \mathbf{R}^2 with its standard orthogonal basis \mathcal{B} . If $X, Y \in \mathbf{R}^2$ and $M \in M_2(\mathbf{R})$, we have $\langle X, Y \rangle = {}^tXY$ and therefore

$$\langle MX, Y \rangle = {}^t(MX)Y = {}^tX{}^tMY = \langle X, {}^tMY \rangle.$$

The characteristic polynomial of $M = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ is $\chi_M(T) = T^2 - (a+d)T + (ad - b^2)$ with discriminant $\Delta = (a+d)^2 - 4(ad - b^2) = (a-d)^2 + 4b^2 \geq 0$. Therefore, it is split over \mathbf{R} with distinct roots unless $b = 0$ and $a = d$, i.e. $M = a\text{Id}$.

If $\Delta = 0$, then M is scalar and the canonical orthonormal basis of \mathbf{R}^2 and therefore orthogonally *diagonal*. Assume $\Delta > 0$ and let $x, y \in \mathbf{R}$ the distinct roots of χ_M . If X, Y are normed eigenvector of our real symmetric matrix M relatively x, y , one gets

$$x\langle X, Y \rangle = \langle MX, Y \rangle = \langle X, MY \rangle = y\langle X, Y \rangle$$

hence $\langle X, Y \rangle = 0$. Therefore, after the orthonormal base change $\mathcal{B} \rightarrow (X, Y)$, the matrix becomes $\text{diag}(x, y)$. □

2.3 General linear maps of the plane



In this section E denotes a rank real plane without any Euclidean structure. We will explain the reduction theory in this simple but non trivial case due to the fact that the scalar field \mathbf{R} is not algebraically closed (compare with the general results of 5.5.0.2, 5.6.0.1 and 5.9).

Let $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in M_2(\mathbf{R})$.

2.3.1 Minimal polynomial

A direct computation shows that $\chi_M(T) = T^2 - (a + d)T + (ad - bc)$ annihilates M : this is the Cayley-Hamilton theorem in dimension 2. Because $\mathbf{R}[T]$ is a principal ideal domain, the ideal of real polynomials annihilating M is generated by a unique monic polynomial μ_M . Because $\chi_M(M) = 0$, one has $\mu_M | \chi_M$ and therefore

- either $\mu_M = \chi_M$
- either χ_M is of degree 1 and M is the scalar matrix $\frac{\text{tr}(M)}{2} \text{Id}$.

Definition 2.3.1.1. *If M is non scalar, we define the similarity invariants P_2, P_1 of M by $P_1 = \chi_M = \mu_M$ and $P_2 = 1$. If M is scalar, we define $P_1 = P_2 = \mu_M$.*

2.3.2 Cyclic vectors

Assume that M is not a scalar matrix. Then M has at most two eigenlines (because $\deg(\chi_M) = 2$). Let $X \in \mathbf{R}^2$ not belonging to these lines (a real plane is never the union of two lines!). Then X and MX are certainly independent vectors, and are therefore a basis of the plane. Writing M in this basis, remembering the equation $\chi_M(M).X = 0$, we get that M is similar to $C(\chi) = \begin{pmatrix} 0 & -\det(M) \\ 1 & \text{tr}(M) \end{pmatrix}$. Because a matrix is scalar if and only if $\deg(\mu_M) = 1$, we therefore get the plane version of the Frobenius theorem 5.9.

Theorem 2.3.2.1 (Jordan-Frobenius in the plane). *Let M be real matrix.*

1. *One has $P_2 | P_1$ and $P_2 P_1 = \chi_M$.*
2. *Two matrices are similar if and only if they have the same similarity invariants.*
3. *If M is not scalar, it is similar to the "companion" matrix $C(\chi)$ of $P_1 = \chi_M = \mu_M$.*
4. *M is nilpotent if and only if it is similar to the standard matrix $J = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.*

In a certain extent, the rest of the book is dedicated to generalize these results in any dimension.

2.4 Reminder on Gauss elimination method

a version of Gauss elimination not using dilatations nor permutation matrices as far as possible. Let R be a ring and $p, q \geq 1$ two integers. We say that two matrices of $M_{p,q}(R)$ with $p, q \geq 1$ are t -equivalent if they differ by a series of left and right by multiplications by transvections (that we call t -operations).

Proposition 2.4.0.1. *Let $A \in M_{p,q}(\mathbf{k}) - \{0\}$.*

1. *There exists $\delta \in \mathbf{k}^*$ such that A is t -equivalent to $\text{diag}(0_{p-r, q-r}, \delta, \text{Id}_r)$ with $r = \text{rank}(A) - 1$.*
2. *Any square matrix $A \in M_n(\mathbf{k})$ is t -equivalent to $\text{diag}(0_{n-r}, \det(A), \text{Id}_r)$ $r = \text{rank}(A) - 1$.*
3. *$\text{SL}_n(\mathbf{k})$ is generated by transvections.*

Proof. Proof of (1). Induction on $p + q \geq 2$, the case $p + q = 2$ being trivial we assume now $p > 1$ or $q > 1$. If both the last column and line are zero, one applies the induction to the (necessarily non zero) remaining $M_{p-1, q-1}(\mathbf{k})$ matrix.

If there exists $i < p$ or $j < q$ such that $a_{i,q} \neq 0$ or $a_{p,j} \neq 0$, by a t -operation $L_p \mapsto L_p - a_{p,q}/a_{i,q}L_i$ or $C_q \mapsto C_q - a_{p,q}/a_{p,j}C_j$, we can assume $a_{p,q} = 1$. Then, again using t -operations $C_j \mapsto C_j - a_{p,j}C_q$ and $L_i \mapsto L_i - a_{i,q}C_q$, one can now assume that the only non zero coefficient of the last line and column is $a_{p,q} = 1$ and we finish by induction on the remaining $M_{p-1, q-1}(\mathbf{k})$ matrix. .

If $a_{p,q}$ is the only non zero coefficient of the last line and column, we put a non zero term using $L_{p-1} \mapsto L_{p-1} + L_p$ if $p > 1$ and $C_{q-1} \mapsto C_{q-1} + C_q$ else.

(2) and (3) are direct consequences of (1). □

2.4.1 Review of Transvections

Let V be an n -dimensional vector space with $n \geq 2$, $\mathbf{P}V$ its set of lines (dimension 1 sub vector-spaces), $\mathbf{P}V^*$ its set of hyperplanes (dimension $n - 1$ sub vector spaces)². If $f \in \text{Hom}_{\mathbf{k}}(V/D, D)$ we denote by $\tilde{f} \in \text{End}_{\mathbf{k}}(V)$ the linear map $\tilde{x} \mapsto x + f(x \bmod D)$. The set vector space of V of dimension 1 is

Proposition 2.4.1.1. *Let $\tau \in \text{End}_{\mathbf{k}}(V)$. The following properties are equivalent.*

1. *$H(\tau) = \text{Ker}(\tau - \text{Id})$ is a hyperplane of V containing $D(\tau) = \text{Im}(\tau - \text{Id})$, which is a line in V .*
2. *There exist $\varphi \in V^*$ and $v \in V$, both nonzero, such that $\tau(x) = x + \varphi(x)v$ with $\varphi(v) = 0$.*
3. *There exists a (unique) $f \in \text{Hom}_{\mathbf{k}}(V/D(\tau), D(\tau))$ such that $\tau = \tilde{f}$.*
4. *The restriction to the affine hyperplane defined by the equation $\varphi(x) = 1$ is a translation by the vector v .*
5. *The natural morphism $\text{Hom}(V/D, D) \rightarrow \text{GL}(V)$*

²At this stage, this is just a notation; cf. chapter 18 for further insights

6. The matrices of τ are similar to $\text{Id}_n + E_{1,2} = \begin{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & \text{Id}_{n-2} \end{pmatrix}$.

We say that τ is a transvection of V of type $(D(\tau), H(\tau)) \in \mathbf{P}V \times \mathbf{P}V^*$. If φ, v are as above, let us define $\tau_\lambda(x) = x + \lambda\varphi(x)v$, $\lambda \in \mathbf{k}$. Under these conditions, we have:

- $H(\tau) = \text{Ker}(\varphi), D(\tau) = \langle v \rangle$,
- Transvections of type $(\langle v \rangle, \langle \varphi \rangle)$ are given by τ_λ , $\lambda \in \mathbf{k}^*$, and $\lambda \mapsto \tau_\lambda$ is an injective group morphism $(\mathbf{k}, +) \rightarrow (\text{SL}(V), \times)$,
- ${}^t\tau$ is a transvection of V^* of type $(H(\tau), D(\tau)) \in \mathbf{P}V^* \times \mathbf{P}V$.

Proof. TBD □

Recall that the derived subgroup $D(G)$ of a group G is the subgroup generated by the commutators $[g, h] = ghg^{-1}h^{-1}$, $g, h \in G$. It is normal and $G/D(G)$ is the largest abelian quotient of G .

Corollary 2.4.1.2. *One has*

1. $D(\text{GL}(V)) = \text{SL}(V)$ except if $n = 2$ and $\text{Card}(\mathbf{k}) = 2$.
2. $D(\text{SL}(V)) = \text{SL}(V)$ except if $n = 2$ and $\text{Card}(\mathbf{k}) = 2, 8$.

A group G with $D(G) = G$ is called perfect.

Proof. Proof of (1). Because the derived group is normal and all transvections are conjugate in $\text{GL}(V)$, it is enough to show that in our case one transvection is a commutator. If $n \geq 3$ and any characteristic, one computes $[\text{Id} + E_{2,1}, \text{Id} + E_{1,3}] = \text{Id} + E_{2,3}$. If $n = 2$, let us choose $\lambda \neq 0, 1$. Then, $[\text{diag}(\lambda, 1), T_{1,2}(\lambda) = T_{1,2}(\lambda - 1)$ which is a transvection.

Proof of (2). If $n \geq 3$, two transvections $\tau' = g\tau g^{-1}$ are certainly conjugate not only under $\text{GL}(V)$ [Because one can change g by a dilation of ration $\det(g)^{-1}$ commuting with τ]. We leave the $n = 2$ case in exercise (adapt the GL argument with a general diagonal matrix in SL_2). □

2.4.2 Normal subgroups of $\text{GL}(V)$

We will explain the so-called Iwasawa to study normal subgroups of perfect groups G , or equivalently we will give a criterium of simplicity of $G/Z(G)$ where $Z(G)$ is the centrum of G .

Definition 2.4.2.1. Let G be a group acting on a set X , and $B \subseteq X$.

1. We say that B is a G -block and if for all $g \in G$, the sets gB and B are either equal or disjoint. Blocks reduced to a point or to the whole X are called trivial.
2. We say G acts primitively on X if:
 - (a) The action of G on X is transitive;
 - (b) the only G -blocks are trivial.³
3. We say G acts 2-transitively on X if for all $x_1, x_2, y_1, y_2 \in X$, $x_1 \neq x_2$, $y_1 \neq y_2$, there exists $g \in G$ such that $g \cdot x_1 = y_1$ and $g \cdot x_2 = y_2$.

Lemma 2.4.2.2. Let G be a group acting 2-transitively on a set E . Then the action is primitive. For instance, $SL(V)$ and $GL(V)$ act 2-transitively on $\mathbf{P}V$ if $\dim(V) \geq 2$.

Proof. Let B be a subset of X having at least two elements and such that $B \neq X$. Let us show that there exists $g \in G$ such that $gB \neq B$ and $gB \cap B \neq \emptyset$ and therefore that B is not a G -block.

Let $a \neq b \in B$ and $c \in X \setminus B$. By 2-transitivity, there exists $g \in G$ such that $ga = a$ and $gb = c$. We have $a \in gB \cap B$, hence $gB \cap B \neq \emptyset$, and $c \in gB$, $c \notin B$, hence $gB \neq B$. \square

Proposition 2.4.2.3 (Iwasawa criterium). Let G be a group acting faithfully and primitively on a set X . We assume that there exists a family $K_x \subset G_x, x \in X$ such that

1. Each K_x is abelian.
2. For any $g \in G$, $G = \langle gK_xg^{-1} \rangle$.
3. $\cup_{x \in X} K_x$ generates G .

Then any normal subgroup acting non trivially on X contains $D(G)$.

Proof. We start with the direct part of the previous footnote.

Lemma 2.4.2.4. The stabilizer G_x of any primitive action is a maximal subgroup of G .

³Or equivalently (Exercice if the stabilizer G_x of a point $x \in X$ is a maximal subgroup of G).

Proof. Let $G_x \subset H \subset G$ and $B = \{hx, h \in H\}$. I claim that B is a block. If not, assume $B \cap g(B) \neq \emptyset$. There exists $h, h' \in H$ such that $hx = gh'x$ hence $h^{-1}gh' \in G_x \subset H$. Therefore, $g \in H$ and $g(B) \subset B$ proving $B = \{x\}$ and $B = X$ by primitivity assumption. In the first case, $H = G_x$ and we are done. In the second case, H acts transitively on X . Therefore, for any $g \in G$ there exists $h \in H$ such that $gx = hx$ hence $gh^{-1} \in G_x \subset H$ showing $g \in H$. \square

Let N be a normal subgroup and let $x \in X$. Since N is normal, NG_x is a subgroup of G containing G_x and is therefore equal to G_x or G by maximality.

If $NG_x = G_x$, we have $N \subseteq G_x$, and therefore for all

$$g \in G, gNg^{-1} \subset gG_xg^{-1} = G_{gx}.$$

By normality of N , we get $N = N \cap gNg^{-1} \subset G_x \cap G_{gx}$, hence N acts trivially on X and therefore $N = \{1\}$ because G hence N acts faithfully on X : we are done in this case.

Assume now $NG_x = G$. One has $Nx = NG_x x = Gx = X$ because G acts transitively and therefore N acts transitively on X . Let $y = nx, n \in N$ be any point of X and $\kappa \in K_y = nK_x n^{-1}$ which can therefore be written $\kappa = nkn^{-1}$ with $(n, k) \in N \times K_x$. We have

$$\kappa = nkn^{-1} = nkn^{-1}k^{-1}k \stackrel{N \triangleleft G}{\in} NK_x$$

proving $K_y \subset NK_x$ for any $y \in X$ hence $G = NK_x$. We deduce that the morphism $k \mapsto k \bmod N$ is a surjection from the abelian group K_x to G/N commutative hence $N \subset D(G)$. \square

Corollary 2.4.2.5. *If $\dim(V) \geq 2$, any normal nontrivial normal subgroup of $GL(V)$ (or $SL(V)$) contains $SL(V)$ unless \mathbf{k} is a field with 2 (or 8) elements.*

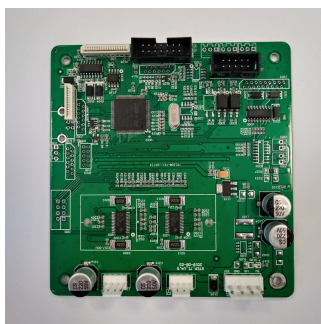
Proof. Take $X = \mathbf{P}(V)$ and $T_D \xrightarrow{\sim} \text{Hom}(V/D, D)$ be the group of transvections of line D (cf. 7.7.1.1) and apply Iwasawa criterium and 2.4.1.2. \square

2.4.3 Supplementary exercises

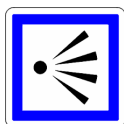
Exercise(s) 2.4.3.1. *Let G act primitively and faithfully on a set X . Assume that for some $x \in X$, the G_x contains an abelian normal subgroup whose conjugate subgroups generate G . Then $D(G) \subset G$ [Adapt the proof of Iwasawa criterium].*

Chapter 3

Generalities on modules



3.1 Perspective



This chapter introduces the language of modules and diagrams in as light a manner as possible. It is suggested that the reader first browse through it focusing on solving the exercises, then later familiarize themselves with its use in the following chapters in a concrete manner.

Thus, it will only be consulted afterward if absolutely necessary: the idea is that all the formal constructions of vector spaces or abelian groups apply *mutatis mutandis* to this general framework by accepting scalars valued in a ring rather than in a field (or integers for abelian groups).

As will be seen here and throughout the text, the diagrammatic perspective (see 3.3) once familiar is extremely valuable, unifying, and simplifying. Paradoxically, this effort in abstraction, besides opening the doors to modern and deep mathematics, often makes them very concrete, even computable and algorithmic.

This will be particularly illustrated in the chapters 4, 5, 6, 8, and 9 dedicated to the study of the linear group and the similarity classes of square matrices. Unlike the usual methods of linear algebra that largely depend on the study of eigenvalues of endomorphisms, we will focus on polynomials and their

action on endomorphisms. While annihilating polynomials play a special role, their roots are not actually important for deciding whether two endomorphisms are similar, for example. The advantage is generally... we do not know how to compute the roots of polynomials. Worse, the constructions of linear algebra are often discontinuous in the coefficients of matrices and thus poorly support the numerical approximation of these roots. Of course, the notion of eigenvalue remains essential as will be seen repeatedly. But its often useless when one cannot compute the roots of the polynomial characteristic or, worse, when the characteristic polynomial is not split.

3.2 Vocabulary and first examples

We know that a vector space over a field \mathbf{k} is an abelian group M equipped with an external law $\mathbf{k} \times M \rightarrow M$ verifying for all $a, a' \in \mathbf{k}$ and $m, m' \in M$ (on the left say) the four usual compatibilities.

1. $a(m + m') = am + am'$
2. $(a + a')m = am + a'm$
3. $1m = m$
4. $a(a'm) = (aa')m$

The notion of a module is obtained exactly in the same way, by allowing the field \mathbf{k} to be a ring R (for us commutative unitary):

Definition 3.2.0.1. *A module M over a unitary ring R is an abelian group equipped with a law $R \times M \rightarrow M$ verifying the previous compatibility properties.*

Example(s) 3.2.0.2. *By definition, modules over fields are vector spaces. Let's provide more interesting examples.*

1. *\mathbf{Z} -modules are identified with abelian groups through external multiplication*


$$n.m = \text{sign}(n) \sum_{i=0}^{|n|} m, \quad n \in \mathbf{Z}, m \in M.$$

2. *If V is a \mathbf{k} -vector space, the set of formal polynomials¹ with coefficients in V is naturally a $k[T]$ -module.*
3. *If R is integral and M a module, the set M_{tors} of elements of M annihilated by a nonzero element of M is a submodule called torsion module .*

4. In general, if M is an arbitrary \mathbf{R} -module, we denote $\text{Ann}_M(r) = \text{Ker}(r : M \rightarrow M)$ and $M[r] = \cup_{n>0} \text{Ker}(r^n : M \rightarrow M)$, which is indeed a submodule as a union of increasing submodules.
5. The set $C_c(\mathbb{T}, \mathbf{R})$ of continuous functions with compact support from a topological space \mathbb{T} to \mathbf{R} is a module over the ring of continuous functions from \mathbb{T} to \mathbf{R} . If \mathbb{T} is a non-compact metric space, $C_c(\mathbb{T}, \mathbf{R})$ is an ideal but not a ring (*exercice*). This ideal is not finitely generated for example if $\mathbb{T} = \mathbf{R}^n$ (*exercice*).
6. Let $M_i, i \in I$ be a family of modules. As in linear algebra, the abelian group product $\prod M_i$ has a natural module structure: it is the unique structure such that all projections $\pi_j : \prod M_i \rightarrow M_j$ are linear. In other terms, $a.(m_i) = (am_i)$ (cf. 3.5.1.1).
7. With the previous notation, the subset $\oplus M_i$ of $\prod M_i$ consisting of almost null families is a submodule called the direct sum of M_i . The (finitely supported) family (m_i) is often denoted $\sum m_i$. If I is furthermore finite, then $\oplus M_i = \prod M_i$.
8. The formula $(\sum_j \lambda_{i,j} T^j)_i = \sum_j (\lambda_{i,j})_i T^j$ allows us to identify $(\mathbf{k}[\mathbb{T}])^n$ and $(\mathbf{k}^n)[\mathbb{T}]$ which we will do henceforth.


We summarize in the following table how the formal constructions of linear algebras adapt to modules. To lighten the notation, the Greek letters $\lambda, \mu \dots$ denote elements of a ring \mathbf{R} while the elements of the modules are Latin letters $x, m, n \dots$ for elements of the modules. The statements are implicitly universally quantified. Thus we write $\lambda(\mu x) = (\lambda\mu)x$ for $\forall \lambda, \mu \in \mathbf{R}$ and $\forall x \in M$, we have $\lambda(\mu x) = (\lambda\mu)x$.

¹That is, sums $\sum_{i \geq 0} v_i T^i$ with $v_i = 0$ if i is large enough.

 Generalities		
Property/Definition	Vector space	Module
Scalars \mathbb{R}	$\mathbb{R} = \text{field}$	$\mathbb{R} = \text{ring}$
Addition	$(M, +)$ abelian group	
External multiplication	$\lambda(\mu x) = (\lambda\mu)x$ and $1x = x$	
Distributivity	$\lambda(x + y) = \lambda x + \lambda y$, $(\lambda + \mu)x = \lambda x + \mu x$	
Linear combination	$\sum_{\text{finite}} \lambda_i x_i$	
Subspace N	N stable by linear combinations	
Generated subspace $\langle x_i \rangle$	$\langle x_i \rangle = \{\text{linear combinations of } x_i\}$	
Sum of subspaces N_i	$+N_i = \{\text{linear combinations of } x_i \in N_i\}$	
Product ² of N_i	$\prod N_i = \{(x_i), x_i \in N_i\}$	
Direct sum ² of N_i	$\oplus N_i = \{(x_i) \in \prod N_i \mid \text{Card}\{i \mid x_i \neq 0\} < \infty\}$	
$\mathbb{R}^{(I)}, \mathbb{R}^n$	$\mathbb{R}^{(I)} = \oplus_I \mathbb{R}$, $\mathbb{R}^n = \oplus_{i=1}^n \mathbb{R} = \prod_{i=1}^n \mathbb{R}$	

The notion of a linear application is translated into that of module morphisms as in the following table, the notion of kernel, image and quotient³ being the same as in linear algebra.

Example(s) 3.2.0.3. \mathbb{Z} -module morphisms are morphisms of abelian groups. See 3.2.4 for the case of V_a .

 Generalities		
Property/Definition	Vector space	Module
Morphism $f \in \text{Hom}_{\mathbb{R}}(M, M')$	morphisms of groups $f(\lambda x) = \lambda f(x)$	
Isomorphism	Bijective morphism	
$\text{Hom}_{\mathbb{R}}(\mathbb{R}^n, M)$	$\text{Hom}_{\mathbb{R}}(\mathbb{R}^n, M) = M^n$	
Matrices	$\text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m) = M_{m,n}(\mathbb{R})$	

²See 3.5.1.

³See 3.2.1.

Specifically, we have

Lemma 3.2.0.4. *If M, N are two R -modules, the set of morphisms $\text{Hom}_R(M, N)$ is naturally a module.*

If $M = R^n$, the natural application

$$\begin{cases} \text{Hom}_R(R^n, N) & \rightarrow & N^n \\ f & \mapsto & (f(\delta_{i,j}))_j \end{cases}$$

is an isomorphism. In particular, $\text{Hom}_R(R^n, R^m) = M_{m,n}(R)$.

3.2.1 Quotient, cokernel

The problem we are tackling is as follows. Let $f : M \rightarrow N$ be a morphism of R -modules. The injectivity of f is characterized by the nullity of the kernel $\text{Ker}(f)$ of f . Can we find a module whose nullity measures the surjectivity?

We define a relation on N by the condition

$$n \sim n' \text{ if and only if } \exists m \text{ such that } n - n' = f(m).$$

This is an equivalence relation thanks to the linearity of f for the law $+$. The equivalence class of $n \in N$ is

$$\bar{n} = \{n + f(m), m \in M\} = n + f(M)$$

We denote $\text{Coker}(f)$ the set of equivalence classes of \sim . Thus, as a set,

$$\text{Coker}(f) = \{n + f(M), n \in N\}$$

and the application $\pi : N \rightarrow \text{Coker}(f)$ defined by $n \mapsto \pi(n) = \bar{n}$ is surjective. The following statement is also as immediate as it is important.

Proposition 3.2.1.1. *There exists a unique R -module structure on $\text{Coker}(f)$ such that π is a morphism. It is characterized by $\bar{n} + \bar{n}' = \overline{n + n'}$ and $\lambda \bar{n} = \overline{\lambda n}$; its neutral is $\bar{0}$ simply noted 0 . Moreover, f is surjective if and only if $\text{Coker}(f) = \{0\}$.*

Thus, we have resolved our problem. A particular, fundamental case is when f is injective. In this case, f induces an isomorphism of M onto its image $f(M)$ which is thus a submodule N' of N .

Definition 3.2.1.2. *Let N' be a submodule of N and denote j the inclusion of N' in N . We say that $\text{Coker}(j)$ is the quotient of N by N' and we denote it N/N' .*

It is important to characterize the cokernel, up to canonical isomorphism, by its properties rather than by its construction. This is what is explained in 3.5.2.1.

Remark(s) 3.2.1.3. *In general, we are interested in modules up to isomorphism. Thus, we will identify two modules between which exists a canonical isomorphism, that is, one that depends on no choice. The reader is, for example, used in linear algebra to identify a finite-dimensional vector space with its bidual (cf. 7.4.0.1), a Euclidean space with its dual (cf. more generally 19.3.1), a square matrix of dimension 1 with its unique coefficient (its trace actually)... Similarly, as in linear algebra, we will most often identify an injective morphism $j : M \rightarrow N$ with the submodule image $j(M)$ because j defines a canonical isomorphism $M \simeq j(M)$ and we simply say (but somewhat abusively) that M is a submodule of N . We will see other examples.*

The following result is formal but important (compare with 3.5)

Proposition 3.2.1.4. *If $f \in \text{Hom}_{\mathbb{R}}(M, N)$, then f induces a canonical isomorphism $\bar{f} : M/\text{Ker}(f) \simeq \text{Im}(f)$.*






Proof. We define

$$\bar{f}(\bar{m}) = \bar{f}(m + \text{Ker}(f)) = f(m + \text{Ker}(f)) = f(m) + f(\text{Ker}(f)) = f(m) \in \text{Im}(f).$$

Thus, \bar{f} is well defined and linear. It is surjective. If \bar{m} is in the kernel, $\bar{f}(\bar{m}) = f(m) = 0$ and therefore $m \in \text{Ker}(f)$ so $\bar{m} = 0$. \square

3.2.2 Properties to handle with caution

While the definitions of free families, generating families, or bases do not change just like that of supplementary, **most of the existence theorems become false in the case of modules** as summarized in the table below. This often comes from *torsion* phenomena: it happens, frequently as we will see, that the equation $am = 0$ does not entail a or m being zero. We will return to this.

 Bases, dimension, supplementary		
Property/Definition	Vector space	Module
Free family $(x_i)_{i \in I}$	$\sum \lambda_i x_i = 0 \Rightarrow \lambda_i \equiv 0$ or $\mathbb{R}^{(I)} \xrightarrow{\lambda_i \mapsto \sum \lambda_i x_i} M$ injective	
	$x \neq 0$ is free	$x \neq 0$ is rarely free
x torsion = x non free	$x = 0$	$\exists \lambda \neq 0 \mid \lambda x = 0$
Generating family $(x_i)_{i \in I}$	$\langle x_i \rangle = M$ or $\mathbb{R}^{(I)} \xrightarrow{\lambda_i \mapsto \sum \lambda_i x_i} M$ surjective	
Base $(x_i)_{i \in I}$	(x_i) free and generating or $\mathbb{R}^{(I)} \xrightarrow{\lambda_i \mapsto \sum \lambda_i x_i} M$ bijective	
	vector spaces have bases	modules rarely have bases
Free module M	$M \simeq \mathbb{R}^{(I)}$ i.e. M admits a base	
Bases of free module	all bases have the same cardinal ⁴	
	vector spaces are free	modules are rarely free
Complement S of N in M	$M = N \oplus S$	
	vector subspaces have supplementary	submodules rarely have supplementaries ⁵

As in linear algebra, giving a linear application from a free module to any module is equivalent to giving the images of a base. Similarly, linear applications between free modules equipped with a base are identified with matrices with coefficients in \mathbb{R} of the appropriate size (cf. 3.2.0.4 above).

Example(s) 3.2.2.1. 1. Multiplication makes \mathbb{R} a (free) module over itself (with base 1) and its submodules are the ideals of \mathbb{R} .

2. $\mathbb{R}_{<n}[\mathbb{T}]$ is a free \mathbb{R} -module with base $X^i, i < n$ therefore of rank n for $n \in \bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$.

3. Multiplication by the elements of \mathbb{R} makes $M_{n,m}(\mathbb{R})$ a free module with base the standard matrices $(E_{i,j})_{1 \leq i \leq n, 1 \leq j \leq m}$.

⁴See 3.8.0.4

⁵When this is the case, we say that N is a direct factor.

4. The module R^m is free with base (canonical) $(e_j = (\delta_{i,j}, i = 1, \dots, m))_{1 \leq j \leq m}$ (cf. 3.5.1.1).
5. If $(e_i)_{1 \leq i \leq n}$ is a basis of the \mathbf{k} -vector space V , the e_i seen as constant polynomials of $V[\mathbf{T}]$ form a basis for $V[\mathbf{T}]$, a module which we will thus identify with $\mathbf{k}[\mathbf{T}]^n$ through this means (*exercice*).

3.2.3 Cyclic modules

We know that the subgroups of a cyclic group are cyclic and that the subgroups of $\mathbf{Z}/n\mathbf{Z}$ are of the form $n/d\mathbf{Z}/n\mathbf{Z}$ with $d|n$. Replacing \mathbf{Z} with a principal ring, we obtain

Lemma 3.2.3.1 (Cyclic modules). *Let R be a principal ring and $M = Rm$ a cyclic (or monogenic) module and let (r) be a generator of the ideal $\text{Ann}_R(m)$.*

- We have $M \simeq R/(r)$.

Let N be a submodule of M and ρ' a generator of the ideal $[N : M] = \{x \in R \mid xM \subset N\}$. We have

- $\rho' | r = \rho\rho'$ and $R/(\rho) \xrightarrow{x \mapsto x\rho' m} N$ is an isomorphism of R -modules.
- r, ρ, ρ' are well defined up to a unit. In particular, the submodules of M are finite in number as soon as r is non-zero.

Proof. As m is a generator of M , the homothety of ratio m on M is surjective. As its kernel is the ideal $\text{Ann}_R(m) = (r)$ we have $M \simeq R/(r)$ according to 3.2.1.4.

The morphism

$$\begin{cases} [N : M] & \rightarrow & N \\ x & \mapsto & xm \end{cases}$$

is surjective because m generates M and its kernel is precisely $\text{Ann}_R(m) = (r) \subset [N : M] = (\rho')$ so that $[N : M]/\text{Ann}_R(m) \simeq N$ according to 3.2.1.4. As $r \in (\rho')$, we have indeed $\rho' | r = \rho\rho'$ so that multiplication by m induces an isomorphism $(\rho')/(\rho\rho') \simeq N$. But then, multiplication by ρ' in turn induces an isomorphism $R/(\rho) \simeq (\rho')/(\rho\rho')$ whence the second point. The third follows from the fact that, up to a unit, the number of divisors of r is $\prod n_i$ where n_i is the exponent of an irreducible factor p_i in a decomposition into products of distinct irreducibles of r (cf. 11).

□

⁶*i.e.* almost null sequences.

3.2.4 The $\mathbf{k}[T]$ -module V_a



If $R = \mathbf{k}[T]$ and M is an R -module, multiplication by the elements of \mathbf{k} seen as constant polynomials makes M a \mathbf{k} -vector space. Furthermore, multiplication by T defines $a \in \text{End}_{\mathbf{k}}(M)$: the homothety of ratio T . Conversely, if V is a \mathbf{k} -vector space and $a \in \text{End}_{\mathbf{k}}(V)$, we define a R -module structure V_a on V by the formula $T.v = a(v)$ and by linearity

$$P(T).v = P(a)(v) \forall P \in R = \mathbf{k}[T], v \in V_a = V$$

These two constructions are inverses of each other:

*The $\mathbf{k}[T]$ -modules are identified with the pairs $(V, a), a \in \text{End}_{\mathbf{k}}(V)$.
Submodules of V_a are then identified with subspaces of V stable by a (*exercice*).*

From the perspective of morphisms, the identification works as follows. If $N = W_b$ is a second module associated with an endomorphism $b \in \text{End}_{\mathbf{k}}(W)$, a morphism $f \in \text{Hom}_R(M, N) = \text{Hom}_{\mathbf{k}[T]}(V_a, V_b)$ is defined by $f \in \text{Hom}_{\mathbf{k}}(V, W)$ such that

$$f \circ a(m) = f(Tm) = Tf(m) = b \circ f(m) \forall m \in M$$

i.e.

(i) $\text{Hom}_{\mathbf{k}[T]}(V_a, V_b) = \{f \in \text{Hom}_{\mathbf{k}}(V, W) \text{ such that } b \circ f = f \circ a\}$

Corollary 3.2.4.1. *If $f \in \text{Isom}_{\mathbf{k}[T]}(V_a, V_b)$ if and only if $a = f^{-1} \circ b \circ f$ so that V_a and V_b are isomorphic if and only if a and b are similar.*

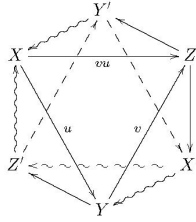
Remark(s) 3.2.4.2. *Following the general principle of formal transposition, the reader will have guessed that $\text{Hom}_R(M, N)$ denotes the space of R -linear applications from M to N , ditto for $\text{End}_R(M), \dots$. When the context is clear, the mention of the ring in the index will be omitted.*

In particular, when $a = b$, we have

(ii) $\text{End}_{\mathbf{k}[T]}(V_a) = \text{Com}(a)$

where $\text{Com}(a)$ is the commutant of a , the set of endomorphisms of V that commute with a .

3.3 Exact sequences and diagrams



3.3.1 Exact sequences

If $f \in \text{Hom}(M, N)$ a morphism of modules; we have a canonical sequence of morphisms

$$\text{Ker}(f) \xrightarrow{\iota} M \xrightarrow{f} N \xrightarrow{\pi} \text{Coker}(f).$$

We notice that the composed of two successive morphisms $d \circ \delta$ (namely $f \circ \iota$ and $\pi \circ f$) are null, which is equivalent to the inclusions $\text{Im}(\delta) \subset \text{Ker}(d)$. But we have better: these inclusions are equalities! This leads to the following definition

Definition 3.3.1.1. Let $d_i \in \text{Hom}(M_i, M_{i+1})$ morphisms, noted as a «sequence»:

$$\cdots M_{i-1} \xrightarrow{d_{i-1}} M_i \xrightarrow{d_i} M_{i+1} \cdots$$

- We say that the sequence is a complex (at i) if $d_i \circ d_{i-1} = 0$ ie $\text{Im}(d_{i-1}) \subset \text{Ker}(d_i)$.
- We say that the sequence is exact (at i) if in addition $\text{Im}(d_{i-1}) \supset \text{Ker}(d_i)$ ie $\text{Ker}(d_i) = \text{Im}(d_{i-1})$.

An exact sequence is therefore a particular complex.

Exercise(s) 3.3.1.2. Let $f \in \text{Hom}(M, N)$.

- Show that $0 \rightarrow M \xrightarrow{f} N$ is exact if and only if f is injective. What is the analogue for surjectivity?
- Show that the sequence $0 \rightarrow K \rightarrow M \xrightarrow{f} N$ is exact if and only if K can be identified (canonically) with the kernel of f . Compare with 3.4.0.2 infra.
- Show that the product or direct sum of exact sequences is still exact.

3.3.2 A fundamental exact sequence

Example(s) 3.3.2.1. Let $d \in R$. Then, the sequence

$$R \xrightarrow{r \mapsto dr} R \xrightarrow{r \mapsto r \pmod{d}} R/(d) \rightarrow 0$$

is exact. More generally, for $(d_i) \in R^\nu$, the «diagonal» sequence

$$R^\nu \xrightarrow{(r_i) \mapsto (d_i r_i)} R^\nu \xrightarrow{(r_i) \mapsto (r_i \pmod{d_i})} \prod_{i=1}^{\nu} R/(d_i) \rightarrow 0$$

is exact (for example as a product of exact sequences).

Generalizing the previous example to the case of matrices $D \in M_{n,m}(R)$ «diagonal» in the sense that its coefficients $d_{i,j}$ are zero if $i \neq j$. Thus, we have a block decomposition (possibly empty)

$$D = \begin{pmatrix} \text{diag}(d_i)_{\nu,\nu} & 0_{\nu,m-\nu} \\ 0_{n-\nu,\nu} & 0_{n-\nu,m-\nu} \end{pmatrix}$$

with $\nu = \min(m, n)$ and $d_i = d_{i,i}$, $i = 1, \dots, \nu$ (and where we note that $0_{n-\nu,m-\nu}$ is the matrix ... empty !). We have two exact sequences: the first

$$R^\nu \xrightarrow{(r_i) \mapsto (d_i r_i)} R^\nu \xrightarrow{(r_i) \mapsto (r_i \pmod{d_i})} \prod_{i=1}^{\nu} R/(d_i) \rightarrow 0$$

according to the previous example, the second

$$R^{m-\nu} \xrightarrow{0_{n-\nu,m-\nu}} R^{n-\nu} \xrightarrow{Id_{n-\nu}} R^{n-\nu} \rightarrow 0$$

because the first arrow is... null !

The sum of these two sequences remains exact: we deduce the important lemma

Lemma 3.3.2.2. The sequence

$$R^m \xrightarrow{D} R^n = R^\nu \times R^{n-\nu} \xrightarrow{((r_i), r') \mapsto ((r_i \pmod{d_i}), r')} \prod_{i=1}^{\nu} R/(d_i) \times R^{n-\nu} \rightarrow 0$$

is exact.

3.3.3 Commutative diagrams

We want to see properties of morphisms in terms of diagrams. For example, to say that $f, g \in \text{Hom}_k(V, W)$ are equivalent endomorphisms in the sense of linear algebra is to say there exist endomorphisms p, q of

W, V such that $p \circ f = g \circ q$ with p, q isomorphisms. The first condition $p \circ f = g \circ q$ (resp. both conditions) is then translated by saying that the diagram

$$\begin{array}{ccc} V & \xrightarrow{p} & V \\ \downarrow g & & \downarrow f \\ W & \xrightarrow{q} & W \end{array} \quad \text{resp.} \quad \begin{array}{ccccccc} 0 & \longrightarrow & V & \xrightarrow{p} & V & \longrightarrow & 0 \\ & & \downarrow g & & \downarrow f & & \\ 0 & \longrightarrow & W & \xrightarrow{q} & W & \longrightarrow & 0 \end{array}$$

is *commutative* with exact lines⁷ (this last condition being empty for the first diagram). A general formal definition (which we encourage the reader not to read!) might be

Definition 3.3.3.1. Let $G = (S, A)$ be a directed graph with vertices S and edges A .

- A *diagram* is the data for each vertex $\Sigma \in S$ of a module M_Σ and for each edge $a : \Sigma_{>} \rightarrow \Sigma_{<}$ of A of a morphism $f_a : M_{\Sigma_{>}} \rightarrow M_{\Sigma_{<}}$.
- The diagram is said to be *commutative* if for every couple of vertices Σ, Σ' , the composed of the f_a associated with an oriented path from Σ to Σ' depends only on the vertices and not on the chosen path.

In practice, we will only deal with diagrams composed of squares or triangles for which the definition of commutativity will be obvious.

3.4 Functoriality and diagram chasing

Although very simple, the following functoriality statements are crucial. This is a very convenient form to formulate the universal properties of kernels and cokernels (cf. §3.5).

Proposition 3.4.0.1 (Functoriality I). Assume we have a commutative diagram of R -modules where the top horizontal line is exact and the bottom line is a complex.

$$\begin{array}{ccccccc} M_1 & \xrightarrow{\mu_1} & M_2 & \longrightarrow & M_3 & \longrightarrow & 0 \\ \downarrow & & \downarrow & & & & \\ N_1 & \xrightarrow{\nu_1} & N_2 & \longrightarrow & N_3 & \longrightarrow & 0 \end{array}$$

Then there exists a unique morphism

$$f_3 : M_3 \rightarrow N_3$$

making the completed diagram commutative

$$\begin{array}{ccccccc} M_1 & \xrightarrow{\mu_1} & M_2 & \longrightarrow & M_3 & \longrightarrow & 0 \\ \downarrow & & \downarrow & & \downarrow f_3 & & \\ N_1 & \xrightarrow{\nu_1} & N_2 & \longrightarrow & N_3 & \longrightarrow & 0 \end{array}$$

⁷By convention, the lines of a diagram are horizontal, the columns vertical.

If in addition, the lower complex line is an exact sequence and the two arrows $M_i \rightarrow N_i$, $i = 1, 2$ are isomorphisms, then f_3 is an isomorphism..

Proof. We focus on the existence and uniqueness of the commutative diagram

$$\begin{array}{ccccccc} M_1 & \xrightarrow{\mu_1} & M_2 & \xrightarrow{\mu_2} & M_3 & \longrightarrow & 0 \\ \downarrow f_1 & & \downarrow f_2 & & \downarrow f_3 & & \\ N_1 & \xrightarrow{\nu_1} & N_2 & \xrightarrow{\nu_2} & N_3 & & \end{array}$$

If there are two arrows f_3 and f'_3 that work, we have $f_3 \circ \mu_2 = \nu_2 \circ f_2 = f'_3 \circ \mu_2$ so f_3 and f'_3 coincide on $\mu_2(M_2) = M_3$ and therefore are equal, hence the uniqueness.

For existence, let $m_3 \in M_3$ and consider m_2 one antecedent by μ_2 . If m_2 is not unique, it is defined modulo $\text{Ker}(\mu_2) = \text{Im}(\mu_1)$. By linearity, the image $\nu_2 \circ f_2(m_2)$ is well defined modulo $\nu_2 \circ f_2 \circ \mu_1(M_1)$. But by commutativity of the left square, we have $\nu_2 \circ f_2 \circ \mu_1 = \nu_2 \circ \nu_1 \circ f_1 = 0$ because $\nu_2 \circ \nu_1 = 0$ by hypothesis. Thus, $\nu_2 \circ f_2(m_2)$ is well defined, *i.e.* depends only on m_3 . Then set $f_3(m_3) = \nu_2 \circ f_2(m_2)$ which is checked to work.

For the second part, we can easily verify by hand that the bijectivity of f_1, f_2 implies that of f_3 (**exercice**). Let's give a «categorical»proof, which has the advantage of generalizing to other contexts. Under the bijectivity assumptions of f_1, f_2 , we want to prove that f_3 admits a left inverse g_3 and a right inverse d_3 . From $g_3 \circ f_3 = \text{Id}_{M_3}$ we then obtain by composing on the right by d_3 the equality $g_3 = d_3$ and thus that f_3 is invertible.

Let's show the existence of g_3 . Call g_1, g_2 the inverses of f_1, f_2 . As $f_2 \circ \mu_1 = \nu_1 \circ f_1$, by composing on the left by g_2 and on the right by g_1 we have $\nu_2 \circ g_1 = g_2 \circ \nu_1$ so we have a commutative diagram with exact lines

$$\begin{array}{ccccccc} M_1 & \xrightarrow{\mu_1} & M_2 & \xrightarrow{\mu_2} & M_3 & \longrightarrow & 0 \\ \downarrow f_1 & & \downarrow f_2 & & \downarrow f_3 & & \\ N_1 & \xrightarrow{\nu_1} & N_2 & \xrightarrow{\nu_2} & N_3 & \longrightarrow & 0 \\ \downarrow g_1 & & \downarrow g_2 & & \downarrow g_3 & & \\ M_1 & \xrightarrow{\nu_1} & M_2 & \xrightarrow{\nu_2} & M_3 & \longrightarrow & 0 \end{array}$$

that we can complete uniquely in a commutative diagram with exact lines according to the first point

$$\begin{array}{ccccccc} M_1 & \xrightarrow{\mu_1} & M_2 & \xrightarrow{\mu_2} & M_3 & \longrightarrow & 0 \\ \downarrow f_1 & & \downarrow f_2 & & \downarrow f_3 & & \\ N_1 & \xrightarrow{\nu_1} & N_2 & \xrightarrow{\nu_2} & N_3 & \longrightarrow & 0 \\ \downarrow g_1 & & \downarrow g_2 & & \downarrow g_3 & & \\ M_1 & \xrightarrow{\nu_1} & M_2 & \xrightarrow{\nu_2} & M_3 & \longrightarrow & 0 \end{array}$$

But by looking at the outer square, taking into account $g_1 \circ f_1 = \text{Id}_{M_1}$ and $g_2 \circ f_2 = \text{Id}_{M_2}$, we have a commutative diagram with exact lines

$$\begin{array}{ccccccc}
 M_1 & \xrightarrow{\nu_1} & M_2 & \xrightarrow{\nu_2} & M_3 & \longrightarrow & 0 \\
 \downarrow \text{Id} & & \downarrow \text{Id} & & \downarrow g_3 \circ f_3 & & \\
 M_1 & \xrightarrow{\nu_1} & M_2 & \xrightarrow{\nu_2} & M_3 & \longrightarrow & 0
 \end{array}$$

But we also have a commutative diagram

$$\begin{array}{ccccccc}
 M_1 & \xrightarrow{\nu_1} & M_2 & \xrightarrow{\nu_2} & M_3 & \longrightarrow & 0 \\
 \downarrow \text{Id} & & \downarrow \text{Id} & & \downarrow \text{Id} & & \\
 M_1 & \xrightarrow{\nu_1} & M_2 & \xrightarrow{\nu_2} & M_3 & \longrightarrow & 0
 \end{array}$$

which, thanks to the uniqueness in the first point, gives $g_3 \circ f_3 = \text{Id}_{M_3}$. By exchanging the roles of M, N , we construct the right inverse of f_3 . □

We obtain exactly the same statement by «reversing the direction of the arrows»⁸

Proposition 3.4.0.2 (Functoriality II). *Suppose we have a commutative diagram of R -modules where the bottom horizontal line is exact and the top line is a complex.*

$$\begin{array}{ccccccc}
 0 & \longrightarrow & M_1 & \longrightarrow & M_2 & \xrightarrow{\mu_2} & M_3 \\
 & & & & \downarrow & & \downarrow \\
 0 & \longrightarrow & N_1 & \longrightarrow & N_2 & \xrightarrow{\nu_2} & N_3
 \end{array}$$

Then there exists a unique morphism

$$\iota_1 : M_1 \rightarrow N_1$$

making the completed diagram commutative

$$\begin{array}{ccccccc}
 0 & \longrightarrow & M_1 & \longrightarrow & M_2 & \xrightarrow{\mu_2} & M_3 \\
 & & \downarrow \iota_1 & & \downarrow & & \downarrow \\
 0 & \longrightarrow & N_1 & \longrightarrow & N_2 & \xrightarrow{\nu_2} & N_3
 \end{array}$$

If in addition, the lower complex line is an exact sequence and the two arrows $M_i \rightarrow N_i$, $i = 2, 3$ are isomorphisms, then ι_3 is an isomorphism.

A sometimes useful generalization is the famous (and formal) five lemma

Exercise(s) 3.4.0.3. Consider a commutative diagram of modules with exact lines

⁸an injection $0 \rightarrow M \rightarrow N$ being thus replaced by a surjection $M \rightarrow N \rightarrow 0$ and vice versa! This is a general phenomenon: any formal statement involving commutative diagrams, complexes, and exact sequences gives rise to an analogous statement by reversing the direction of the arrows. We can give a precise sense to this statement valid in any «abelian category». We will content ourselves, and it is quite sufficient, to see this as a meta-principle.

$$\begin{array}{ccccccccc}
 M_1 & \longrightarrow & M_2 & \longrightarrow & M_3 & \longrightarrow & M_4 & \longrightarrow & M_5 \\
 \downarrow f_1 & & \downarrow f_2 & & \downarrow f_3 & & \downarrow f_4 & & \downarrow f_5 \\
 N_1 & \longrightarrow & N_2 & \longrightarrow & N_3 & \longrightarrow & N_4 & \longrightarrow & N_5
 \end{array}$$

- If f_2, f_4 injective and f_1 surjective, then f_3 injective.
- If f_2, f_4 surjective and f_5 injective, then f_3 bijective.

We use it most often in the following weakened form: Consider a commutative diagram of modules with exact lines

$$\begin{array}{ccccccccc}
 0 & \longrightarrow & M_2 & \longrightarrow & M_3 & \longrightarrow & M_4 & \longrightarrow & 0 \\
 & & \downarrow f_2 & & \downarrow f_3 & & \downarrow f_4 & & \\
 0 & \longrightarrow & N_2 & \longrightarrow & N_3 & \longrightarrow & N_4 & \longrightarrow & 0
 \end{array}$$

If f_2, f_4 bijective f_3 bijective.

By construction of the cokernel, we therefore have a canonical exact sequence

$$(0) \quad M_1 \xrightarrow{\mu_1} M_2 \rightarrow \text{Coker}(\mu_1) \rightarrow 0$$

We then have the important characterization of the cokernel (compare with exercise 3.3.1.2.)

Proposition 3.4.0.4. *Show that the sequence $M_1 \xrightarrow{\mu_1} M_2 \xrightarrow{\mu_2} M_3 \rightarrow 0$ is exact if and only if M_3 can be identified (canonically) with the cokernel of μ_1 .*

Proof. Just apply the functoriality 3.4.0.1 to the commutative diagram with exact lines

$$\begin{array}{ccccccccc}
 M_1 & \xrightarrow{\mu_1} & M_2 & \longrightarrow & \text{Coker}(\mu_1) & \longrightarrow & 0 \\
 \downarrow \text{Id} & & \downarrow \text{Id} & & & & \\
 M_1 & \xrightarrow{\mu_1} & M_2 & \xrightarrow{\mu_2} & M_3 & \longrightarrow & 0
 \end{array}$$

□

Exercise(s) 3.4.0.5. *State and prove the result obtained by reversing the direction of the arrows.*

3.5 Universal properties

The question posed is to characterize the various modules M in question by the «calculation» of

$$h(T) = \text{Hom}(T, M) \text{ or } h^\vee(T) = \text{Hom}(M, T)$$

for T an arbitrary «test module». Thus, T is seen as a variable and h, h^\vee as a function of T whose values are sets. One should say functor: the composition with $f \in \text{Hom}_R(M, N)$ defines an application (linear) $h_f(T) : h_M(T) \rightarrow h_N(T)$ (resp. $h_f^\vee : h^\vee(N) \rightarrow h_M^\vee(T)$) which is compatible with composition⁹ The correct general framework to formulate what follows is that of the Yoneda lemma in categories, but we will stay in the framework of modules for the examples that interest us to avoid unnecessary formalism.

3.5.1 Sum and product

Let $M_i, i \in I$ be a family of modules. We denote $M_i \xrightarrow{\varphi_i} \oplus M_i$ the canonical injections and $\prod M_i \xrightarrow{\pi_i} M_i$ the canonical projections. If T is a test module we have two tautological applications

$$\underline{h}^\vee(T) : \begin{cases} \text{Hom}_R(\oplus M_i, T) & \rightarrow & \prod \text{Hom}(M_i, T) \\ f & \mapsto & (\varphi_i \circ f) \end{cases}$$

and

$$\underline{h}(T) : \begin{cases} \text{Hom}_R(T, \prod M_i) & \rightarrow & \prod \text{Hom}(T, M_i) \\ g & \mapsto & (g \circ \pi_i) \end{cases}$$

Lemma 3.5.1.1 (Universal properties of sum and product). *The applications $\underline{h}(T)$ and $\underline{h}^\vee(T)$ are bijective.*

The proof is immediate and left as an **exercice**. In the case of the direct sum, the meaning of the lemma is that giving a morphism $f : \oplus M_i \rightarrow T$ is equivalent to giving a collection of morphisms $f_i : M_i \rightarrow T$ (thanks to the formula $f(\sum m_i) = \sum f_i(m_i)$ which is well defined because the sum is actually finite).

3.5.2 Kernel and cokernel



Let $f : M \rightarrow N$ be a morphism of modules. By construction, we have two exact sequences

$$0 \rightarrow \text{Ker}(f) \xrightarrow{j} M \rightarrow N$$

and

$$M \rightarrow N \xrightarrow{p} \text{Coker}(f) \rightarrow 0$$

that characterize kernel and cokernel (3.3.1.2 and 3.4.0.4).

⁹The reader will recognize the usual notion of «restriction» of a morphism for $h_f(T)$ and dually of «transpose» for $h^\vee(f)$.

If T is a test module we have two tautological applications

$$h^\vee(T) : \begin{cases} \text{Hom}(\text{Coker}(f), T) & \rightarrow & \text{Hom}_0(N, T) = \{\psi \in \text{Hom}(N, T) \mid \psi \circ f = 0\} \\ \varphi & \mapsto & \varphi \circ p \end{cases}$$

and

$$h(T) : \begin{cases} \text{Hom}(T, \text{Ker}(f)) & \rightarrow & \text{Hom}_0(T, M) = \{\psi \in \text{Hom}(T, M) \mid f \circ \psi = 0\} \\ \varphi & \mapsto & j \circ \varphi \end{cases}$$

Lemma 3.5.2.1 (Universal properties of kernel and cokernel). *The applications $h(T)$ and $h^\vee(T)$ are bijective.*

Proof. Let's prove, for example, the universal property of the cokernel ie construct the inverse of $h^\vee(T)$. Observing that we have an exact sequence $0 \rightarrow T \xrightarrow{\text{Id}} T \rightarrow 0$. Let then $\psi \in \text{Hom}_0(N, T)$. The condition $\psi \circ f = 0$ precisely ensures the commutativity of the diagram

$$\begin{array}{ccccccc} M & \xrightarrow{f} & N & \xrightarrow{p} & \text{Coker}(f) & \longrightarrow & 0 \\ \downarrow & & \downarrow \psi & & \downarrow & & \\ 0 & \longrightarrow & T & \xrightarrow{\text{Id}} & T & \longrightarrow & 0 \end{array}$$

so that 3.4.0.1 ensures the existence of a unique φ making the diagram

$$\begin{array}{ccccccc} M & \xrightarrow{f} & N & \xrightarrow{p} & \text{Coker}(f) & \longrightarrow & 0 \\ \downarrow & & \downarrow \psi & & \downarrow \varphi & & \\ 0 & \longrightarrow & T & \xrightarrow{\text{Id}} & T & \longrightarrow & 0 \end{array}$$

commute. We verify that the application $\psi \mapsto \varphi$ is the inverse of $h^\vee(T)$. □

The meaning of the lemma is that providing a morphism φ from the cokernel to T is equivalent to providing a morphism ψ from N to T such that the composition $\psi \circ f$ is zero, or ψ factors through the quotient (or passes to the quotient) in φ if and only if $\psi \circ f = 0$ (and the analogous for the kernel by reversing the directions of the arrows). From a diagrammatic perspective, we often summarize by keeping only the informal meaning of the statement:

If $\psi \circ f = 0$ then
$$\begin{array}{ccccc} & & & & T \\ & & & & \uparrow \exists! \varphi \\ M & \xrightarrow{f} & N & \longrightarrow & \text{Coker}(f) \end{array}$$

Another way of expressing this, in terms of the functors h and h^\vee , is that the sequences of module morphisms they define

$$0 \rightarrow \text{Hom}(\text{Coker}(f), T) \rightarrow \text{Hom}(N, T) \rightarrow \text{Hom}(M, T)$$

and

$$0 \rightarrow \text{Hom}(T, \text{Ker}(f)) \rightarrow \text{Hom}(T, M) \rightarrow \text{Hom}(T, N)$$

are exact.

3.6 Quotient rings

Let R be a ring. Recall that an ideal I of R is an additive subgroup of R such that

$$\forall r \in R, rI \subset I.$$

By 3.2.1, there exists a unique group structure on R/I making the projection $\pi : R \rightarrow R/I$ a morphism.

The main (simple but important) result goes as follows:

Proposition 3.6.0.1. *There exists a unique group structure on R/I making the projection $\pi : R \rightarrow R/I$ a morphism whose kernel is I . One has the following universal property (cf. 3.5.2.1) : for any ring T , the natural sequence*

$$0 \rightarrow \text{Hom}_{ring}(R/I, T) \rightarrow \text{Hom}_{ring}(R, T) \rightarrow \text{Hom}_{\mathbf{Z}}(I, T)$$

is exact. Moreover, if $f \in \text{Hom}(R, R')$, then f induces a canonical isomorphism of rings $\bar{f} : R/\text{Ker}(f) \simeq \text{Im}(f)$ (cf. 3.2.1.4).

In a diagrammatic way, the main point summarizes as

$$\text{If } \psi(I) = 0 \text{ then} \quad \begin{array}{ccccc} & & & & T \\ & & & \nearrow \psi & \uparrow \exists! \varphi \\ I \subset & R & \longrightarrow & R/I & \end{array}$$

Proof. The proof goes straightforward as in the module case except for the fact that π is multiplicative which follows from the computation

$$\pi(r_1)\pi(r_2) = (r_1 + I)(r_2 + I) + I = r_1r_2 + r_1I + r_2 + I^2 + I = r_1r_2 + I$$

because $r_1I + r_2 + I^2 \subset I$ (recall that if I, J are ideals, IJ denotes the ideal generated by all products ij where $i \in I, j \in J$). □

3.7 A variant of the Chinese remainder theorem

«When General Han Ting arranges his soldiers in threes, there remain two soldiers, when he arranges them in fives, there remain three, and when he arranges them in sevens, there remain two. How many soldiers does Han Ting's army consist of? », Sun Zi, around the 4th century.



Terracotta Army
Mausoleum of Emperor Qin

The key result for us that follows is the following, a slightly generalized version from $\mathbf{k}[T]$ (resp. \mathbf{Z}) to a principal ring of the famous kernel lemma in usual linear algebra (resp. the usual Chinese remainder theorem on integers).

Let M be a module over a principal ring R . For p irreducible, we define $M[p] = \cup_{n>0} \text{Ker}(p^n : M \rightarrow M)$, the p -primary component of M . This is a submodule, as a union of increasing submodules. We assume here that there exists $r \in R$ that annihilates M . Recall the definition of the annihilator $\text{Ann}_M(r) = \text{Ker}(r : M \rightarrow M)$.

Proposition 3.7.0.1 (Chinese Remainder Theorem for modules or Primary decomposition). *Let $r = \prod r_i \in R$. We assume $\text{GCD}(r_i, r_j) = 1$ if $i \neq j$. Let M be a module annihilated by r .*

1. *There exist $u_i \in R$ (independent of M) such that $\sum u_i r/r_i = 1$.*
2. *Then, $M = \oplus \text{Ann}_M(r_i)$ and the projection p_i onto $\text{Ann}_M(r_i)$ parallel to $\oplus_{j \neq i} \text{Ann}_M(r_j)$ is the homothety of ratio $u_i r/r_i \in R$.*
3. *The p_i form an orthogonal family of projectors of M i.e. $\sum p_i = \text{Id}$ and $p_i p_j = \delta_{i,j} p_i$.*
4. *Suppose further $r_i = p_i^{n_i}$ with p_i irreducible. Then $\text{Ann}_M(p^{n_i}) := \text{Ker}(p_i^{n_i} : M \rightarrow M) = M[p_i]$ and: $M = \oplus \text{Ker}(p_i^{n_i} : M \rightarrow M)$.*

Proof. The r/r_i are coprime overall so that the first point is the identity of Bézout.

For the second point, let's first prove that the sum of the $\text{Ann}_M(r_i)$ is direct. Suppose therefore $\sum m_i = 0$ with $m_i \in \text{Ann}_M(r_i)$. For every j , we rewrite $m_j = -\sum_{i \neq j} m_i$. We deduce that the ideal I_j annihilator of m_j contains r_j (left-hand side of the equality) and $\prod_{i \neq j} p_i = r/r_i$ (right-hand side) and hence their GCD by Bézout. Since r_j and r/r_j are coprime, $\text{GCD}(r_i, r/r_i) = 1 \in I_j$ and $1m_j = m_j = 0$ for every j .

Then let $m \in M$. We have $r = \sum u_i r / r_i m$ and $r_i(u_i r / r_i m) = u_i r m = 0$ therefore $m_i \in \text{Ann}_M(r_i)$. The orthogonality is obvious since $1 = \sum u_j r / r_j$ and each $u_j r / r_j$ is divisible by r_i if $j \neq i$.

The third point is a particular case of the second. ■ □

In particular, the projections $M \rightarrow M[p]$ are «functorial» in the following sense: let $f \in \text{Hom}_R(M, N)$; if r as in the proposition annihilates both N and M , we have a *commutative* diagram where the vertical arrows are the projections (thus the homotheties of ratio $u_i r / r_i$)

$$\begin{array}{ccc}
 N & \xrightarrow{f} & M \\
 \downarrow u_i r / r_i & \circlearrowleft & \downarrow u_i r / r_i \\
 N[p_i] & \xrightarrow{f} & M[p_i]
 \end{array}$$

Example(s) 3.7.0.2 ("Kernel lemma"). If $R = k[T]$ with $P = \prod P_i \in$ with P_i pairwise coprime and $M = V_a$ (3.2.4), then $\text{Ann}_M(P) = \text{Ker}(P(a))$ and we recover the usual kernel lemma¹⁰ $\text{Ker}(P(a)) = \oplus \text{Ker}(P_i(a))$.

Exercise(s) 3.7.0.3. Let N be a submodule of M . Show the equality $N[p] = N \cap M[p]$.

If M is the $R = k[T]$ -module V_a (3.2.4) with χ_a split, show that $M[P] = \text{Ker}(a - \lambda \text{Id})^v$ if $P = T - \lambda$ with $\chi_a(\lambda) = 0$ and $M[P] = 0$ otherwise. In other words, the primary components of V_a are its characteristic spaces. What do we recover as a statement about stable spaces of an endomorphism of a vector space?

Remark(s) 3.7.0.4. If R is Euclidean, the calculation of the u_i is algorithmic. Most of the results that we will demonstrate for $k[T]$ algorithmically transpose *mutatis mutandis* to Euclidean rings. They remain true in the principal framework, but without a general algorithm (we then use explicitly or not decompositions into irreducible factors to find Bézout pairs, notably in the *infra* generalization of the Gaussian pivot). This difference is in fact profound: it is a window to K-theory. For a generalization of the Chinese remainder theorem to the non-principal case, see 3.8.0.12.

Example(s) 3.7.0.5. The ring $\mathbf{Z}[\sqrt{-19}]$ is principal but not Euclidean; if one prefers geometry, the same is true of the ring of functions on the circle of radius $\sqrt{-1}$, the ring $\mathbf{R}[x, y]/(x^2 + y^2 + 1)$ (see 4.4) for references.

See also the exercise 3.8.0.12.

Proposition 3.7.0.6 (Chinese Remainder Theorem for rings or Primary decomposition). *Let R be a principal ideal ring and $r = \prod r_i \in R$ with $\text{GCD}(r_i, r_j) = 1$ if $i \neq j$.*

1. *There exist $u_i \in R$ such that $\sum u_i r/r_i = 1$.*
2. *Define $e_i = (u_i r/r_i \pmod r) \in R/(r)$. We have $e_i e_j = \delta_{i,j} e_i$ and the projection*

$$\begin{cases} R/(r) & \rightarrow & \prod R/(r_i) \\ x & \mapsto & (x \pmod{r_i})_i \end{cases}$$

is a ring isomorphism whose inverse is $\varphi : (x_i) \mapsto \sum x_i e_i$. In other words, each projection onto $R/(r_i)$ alongside $R/(r_j), j \neq i$ is defined by multiplication by the idempotent e_i .

3. *Furthermore, the kernel $\text{Ker}(R/(r) \xrightarrow{r_i} R/(r)) \simeq R/(r_i)$ of the multiplication by r_i is $e_i R/(r)$.*

Sketch. For (1), observe that the irreducible factors of $\text{GCD}(r/r_i)$ are the factors r_i of r . However, since $r/r_i = \prod_{j \neq i} r_j$ and r_i is coprime to all $r_j, j \neq i$, it is coprime to r/r_i . Thus, $\text{GCD}(r/r_i) = 1$ and the first point follows from Bézout's identity.

Since $r_i | (r/r_j)$ for $j \neq i$, we have $r = r_i (r/r_i) | (r/r_j) (r/r_i)$ and therefore $e_j e_i = 0$ if $i \neq j$. But since $\sum e_i = 1$ by projection onto $R/(r)$ from the previous Bézout's identity, multiplying by e_i gives us the missing equation $e_i^2 = e_i$. The rest follows immediately. □

3.8 Additional Exercises

Exercise(s) 3.8.0.1. 1. *Show that an abelian group is finite if and only if the associated \mathbf{Z} -module is of finite type and torsion.*

2. *Show that if V_a corresponds to (V, a) (refer to 3.2.4), then V is finite-dimensional if and only if V_a is of finite type and torsion.*

Exercise(s) 3.8.0.2. *Let R be a commutative ring, M, N two R -modules, and M' a submodule of M . Denote π the canonical surjection $\pi : M \rightarrow M/M'$.*

1. *What are the submodules of the R -module R ? What can be said in this case about the quotient?*
2. *Construct from π a bijection between the set of sub- R -modules of M containing M' and the set of sub- R -modules of M/M' .*

Let $f : M \rightarrow N$ be a morphism of R -modules (i.e., an R -linear application).

3. *Show that $\text{Ker } f$ and $\text{Im } f$ are R -modules, as well as $\text{Coker } f = N/\text{Im } f$. Show that there is an isomorphism of R -modules*

$$M/\text{Ker } f \xrightarrow{\sim} \text{Im } f.$$

4. Consider the application $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ associated with the matrix $A = (a_{i,j})$, with

$$a_{i,j} = \begin{cases} 0 & \text{if } i \neq j \\ d_i & \text{if } i = j. \end{cases}$$

Give the structure as an \mathbf{R} -module of $\text{Coker } f$.

Exercise(s) 3.8.0.3. Consider an exact sequence of modules $0 \rightarrow M_1 \xrightarrow{f_1} M_2 \xrightarrow{f_2} M_3 \rightarrow 0$. It is said that $\sigma \in \text{Hom}_{\mathbf{R}}(M_3, M_2)$ is a section of f_2 if $f_2 \circ \sigma = \text{Id}_{M_3}$. When such a section exists, the sequence is said to be split.

1. Assuming such a section exists, show that the application $(m_1, m_3) \mapsto f_1(m_1) + \sigma(m_3)$ defines an isomorphism $M_1 \oplus M_3 \simeq M_2$. Deduce that $M_1 \simeq f_1(M_1)$ then admits a supplement.
2. Conversely, assume that $M_1 \simeq f_1(M_1)$ admits a supplement S . Show that f_2 defines an isomorphism $S \simeq M_3$.
3. Show that a submodule N of M is a direct factor if and only if the exact sequence $0 \rightarrow N \rightarrow M \rightarrow M/N \rightarrow 0$ is split. In this case, show that every supplement of N is isomorphic to M/N .
4. Show that if $n > 1$, the canonical exact sequence $0 \rightarrow \mathbf{Z} \rightarrow \mathbf{Z} \rightarrow \mathbf{Z}/n\mathbf{Z} \rightarrow 0$ is not split.
5. Let $\pi : \mathbf{R}^{n+m} \rightarrow \mathbf{R}^m$ be the projection onto the last m coordinates. Show that there is an exact sequence $0 \rightarrow \mathbf{R}^n \rightarrow \mathbf{R}^{n+m} \xrightarrow{\pi} \mathbf{R}^m \rightarrow 0$ and that this sequence is split.
6. Suppose there are three square matrices A, B, C with coefficients in \mathbf{R} of size $n, n+m, m$ making the diagram commutative

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathbf{R}^n & \longrightarrow & \mathbf{R}^{n+m} & \longrightarrow & \mathbf{R}^n \longrightarrow 0 \\ & & \downarrow A & & \downarrow B & & \downarrow C \\ 0 & \longrightarrow & \mathbf{R}^n & \longrightarrow & \mathbf{R}^{n+m} & \longrightarrow & \mathbf{R}^n \longrightarrow 0 \end{array}$$

Show that B is block triangular and identify the diagonal blocks. State and prove a reciprocal.

Exercise(s) 3.8.0.4. Let M be an \mathbf{R} -module.

1. Show that a proper ideal I of \mathbf{R} is maximal if and only if \mathbf{R}/I is a field.
2. Show that M is of finite type if and only if there exists a surjective \mathbf{R} -linear mapping $\mathbf{R}^n \rightarrow M$ for some $n \in \mathbf{N}$.
3. Show that if $f \in \text{Hom}_{\mathbf{R}}(\mathbf{R}^m, \mathbf{R}^n) = M_{n,m}(\mathbf{R})$ is surjective then $m \geq n$.
Hint: Consider a maximal ideal I of \mathbf{R} and see that after reduction modulo I , the application f remains surjective modulo I .
4. Show that if f is an isomorphism, then $n = m$.

5. Show that a free module of finite type L has a finite basis and that all its bases have the same cardinality: the rank of L .
6. Show that the rank of L is the minimal cardinal of a finite generating family.

Exercise(s) 3.8.0.5. Suppose V is a \mathbf{R} -vector space of dimension 2 and D is a line in V .

Assuming that the line D is given in **parametric form**, that is, a direction vector v of D is given, i.e., a vector $v \in V$ such that $D = \mathbf{R} \cdot v$.

1. Define the linear application $\varphi : t \in \mathbf{R} \mapsto t \cdot v \in V$, show that the following sequence of \mathbf{R} -vector spaces is exact:

$$\{0\} \longrightarrow \mathbf{R} \xrightarrow{\varphi} V.$$

2. What is the image of the \mathbf{R} -linear morphism φ ?

Assume now that the line D is given in **implicit form**, i.e., an equation of the line D is given, that is, a linear form $f \in V^*$ such that $D = \text{Ker}(f)$.

1. Show that the sequence of \mathbf{R} -vector spaces following is exact:

$$V \xrightarrow{f} \mathbf{R} \longrightarrow 0.$$

2. Complete this sequence into a short exact sequence:

$$\{0\} \longrightarrow \mathbf{R} \xrightarrow{\varphi} V \xrightarrow{f} \mathbf{R} \longrightarrow \{0\}.$$

3. Generalize the exercise to any field, a vector space V of arbitrary (finite) dimension, and to arbitrary subspaces.

Exercise(s) 3.8.0.6. Let \mathbf{k} be a field and R a ring.

- Show that the invertibles of $\mathbf{k}[T]$ are the non-zero constant polynomials from \mathbf{k}^* .
- Show that a matrix from $M_n(R)$ is invertible if and only if its determinant is an invertible of R^\times .
Deduce that $M \in M_n(\mathbf{k}[T])$ is invertible if and only if $\det(M) \in \mathbf{k}^*$.

Exercise(s) 3.8.0.7. Recall Zorn's Lemma. Let I be a non-empty ordered set assumed to be inductive (every totally ordered subset has a maximal element). Zorn's Lemma assures that I has a maximal element. Show that Zorn's Lemma implies the existence of maximal ideals (i.e. proper maximal ideals).

Exercise(s) 3.8.0.8 (Snake Lemma). Consider a commutative diagram of modules with exact rows:

$$\begin{array}{ccccccc} A & \xrightarrow{i} & B & \xrightarrow{p} & C & \longrightarrow & 0 \\ f \downarrow & & g \downarrow & & h \downarrow & & \\ 0 \longrightarrow & A' & \xrightarrow{i'} & B' & \xrightarrow{p'} & C' & \end{array}$$

1. Show that i sends $\text{Ker } f$ into $\text{Ker } g$ and p sends $\text{Ker } g$ into $\text{Ker } h$.

2. Show that i' induces a morphism $\text{Coker } f \rightarrow \text{Coker } g$ and that p induces a morphism $\text{Coker } g \rightarrow \text{Coker } h$.

3. Show that there exists a unique morphism $\delta : \text{Ker } h \rightarrow \text{Coker } f$ such that the following sequence is exact:

$$\text{Ker } f \longrightarrow \text{Ker } g \longrightarrow \text{Ker } h \xrightarrow{\delta} \text{Coker } f \longrightarrow \text{Coker } g \longrightarrow \text{Coker } h.$$

Show that if i is injective and p is surjective, then the following sequence is exact:

$$0 \longrightarrow \text{Ker } f \longrightarrow \text{Ker } g \longrightarrow \text{Ker } h \xrightarrow{\delta} \text{Coker } f \longrightarrow \text{Coker } g \longrightarrow \text{Coker } h \longrightarrow 0.$$

4. (Bonus) Retrieve the Five Lemma from the Snake Lemma.

Exercise(s) 3.8.0.9. We will show that if the ring R is not assumed to be commutative, then it may occur that the R -modules R^n , $n \geq 1$ are all isomorphic. To this end, we fix a real vector space V equipped with a countable base $(e_k)_{k \in \mathbf{N}}$ and we denote R the ring of linear applications on V (equipped with composition), identified as «infinite matrices» of $\mathbf{R}^{\mathbf{N} \times \mathbf{N}}$. Define two linear applications T and T' on V by the following relations for $n \in \mathbf{N}$:

$$\begin{cases} T(e_{2n}) = e_n, \\ T(e_{2n+1}) = 0, \end{cases} \quad \text{and} \quad \begin{cases} T'(e_{2n}) = 0, \\ T'(e_{2n+1}) = e_n. \end{cases}$$

Write the «matrices» of T and T' . Given $n \in \mathbf{N}^*$, we consider R^n as an R -module for scalar multiplication:

$$R \times R^n \rightarrow R^n, \quad \left(r, \begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{pmatrix} \right) \mapsto \begin{pmatrix} r \circ T_1 \\ r \circ T_2 \\ \vdots \\ r \circ T_n \end{pmatrix}.$$

1. Provide a one-element base for the R -module R^1 .

2. Show that (T, T') is also a base for the R -module R^1 .

3. Show that R^1 and R^2 are isomorphic as R -modules then that R^n is isomorphic to R for every $n \in \mathbf{N}^*$.

Exercise(s) 3.8.0.10. Let $d \geq 1$ be a natural number, R a principal ring and $M = R^d$. Let N be a submodule of M . We aim to prove by induction on d that N is isomorphic to R^δ with $\delta \leq d$. Assume $d \geq 1$ and the theorem proven for submodules of $R^{d'}$ if $d' < d$.

1. Let $\underline{\nu} = (\nu_1, \dots, \nu_d) \in R^d - \{0\}$ and i such that $\nu_i \neq 0$. The map $\pi_i : (x_1, \dots, x_d) \mapsto x_i$ induces an exact sequence

$$(iii) \quad 0 \rightarrow K \rightarrow N \xrightarrow{\pi_i} C \rightarrow 0$$

where C is a nontrivial submodule of A and $K \subset R^{d-1}$.

2. Show that there exist $d' < d$ and an exact sequence

$$0 \rightarrow \mathbf{R}^{d'} \xrightarrow{j} \mathbf{N} \xrightarrow{\pi} \mathbf{R} \rightarrow 0.$$

3. Show that there exists a section $\sigma = \mathbf{A} \rightarrow \mathbf{N}$ of π , i.e., satisfying $\pi \circ \sigma = \text{Id}_{\mathbf{A}}$.

4. Show that the map $\begin{cases} \mathbf{R}^{d'} \oplus \mathbf{R} & \rightarrow & \mathbf{N} \\ (x, y) & \mapsto & j(x) + \sigma(y) \end{cases}$ is an isomorphism.

5. Conclude.

Exercise(s) 3.8.0.11. TBD

Exercise(s) 3.8.0.12. Let $I_i, 1 \leq i \leq n$ be a finite number of ideals in a ring \mathbf{R} . Assume $I_i + I_j = \mathbf{R}$. Prove by induction on n the following generalization of the Chinese Remainder Theorem. We have:

1. $\sum I_i = \mathbf{R}$.
2. The natural projection $\mathbf{R} \rightarrow \prod \mathbf{R}/I_i$ is surjective.
3. Its kernel $I_1 \cap \dots \cap I_n$ is the product ideal $I_1 \dots I_n$ generated by products of n elements in I_1, \dots, I_n respectively.

Exercise(s) 3.8.0.13 (Resultant). Let \mathbf{R} be a ring and $P, Q \in \mathbf{R}[\mathbf{T}]$ be two polynomials of degrees $p, q > 0$. Let $\text{Res}(P, Q)$ denote the resultant of P and Q , defined as the determinant, in canonical bases (cf. 3.2.4), of the linear map between free modules of rank $p + q$

$$\rho(P, Q) : \begin{cases} \mathbf{R}_{<q}[\mathbf{T}] \times \mathbf{R}_{<p}[\mathbf{T}] & \rightarrow & \mathbf{R}_{<p+q}[\mathbf{T}] \\ (A, B) & \mapsto & AP + BQ \end{cases}$$

1. Calculate $\text{Res}(P, Q)$ if P has degree 1.
2. By considering the comatrix of $\rho(P, Q)$, show that there exist $A, B \in \mathbf{R}[\mathbf{T}]$ of degrees q, p respectively such that $AP + BQ = \text{Res}(P, Q)$. Hence deduce that if P, Q have a common root in \mathbf{R} , then $\text{Res}(P, Q) = 0$.
3. If P, Q are also monic, show that $\rho(P, Q)$ is the matrix of the multiplication $\mu : \mathbf{R}[\mathbf{T}]/(Q) \times \mathbf{R}[\mathbf{T}] \rightarrow \mathbf{R}[\mathbf{T}]/(PQ)$ in canonical bases (of monomial classes \mathbf{T}^i).
4. Still assuming P, Q are monic, show that there is a commutative diagram with exact rows

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathbf{R}[\mathbf{T}]/(PQ) & \xrightarrow{(\mathbf{T}-r)} & \mathbf{R}[\mathbf{T}]/((\mathbf{T}-r)PQ) & \xrightarrow{\text{ev}_r} & \mathbf{R} \longrightarrow 0 \\ & & \uparrow \rho(P, Q) & & \uparrow \rho((\mathbf{T}-r)P, Q) & & \uparrow Q(r) \\ 0 & \longrightarrow & \mathbf{R}[\mathbf{T}]/(Q) \times \mathbf{R}[\mathbf{T}]/(P) & \xrightarrow{(1, (\mathbf{T}-r))} & \mathbf{R}[\mathbf{T}]/(Q) \times \mathbf{R}[\mathbf{T}]/((\mathbf{T}-r)P) & \xrightarrow{\text{ev}_Q(r)} & \mathbf{R} \longrightarrow 0 \end{array}$$

where $\text{ev}(A) = A(r)$ and $\text{ev}_Q(A, B) = A(r)$. Hence deduce that $\rho((\mathbf{T}-r)P, Q)$ is block triangular with diagonal $\text{diag}(\rho(P, Q), Q(r))$, and then that $\text{Res}((\mathbf{T}-r)P, Q) = Q(r) \text{Res}(P, Q)$.

5. If Q is monic, show that $\text{Res}(\prod(\mathbf{T}-r_i), Q) = \prod Q(r_i)$. What happens if Q is not assumed to be monic?

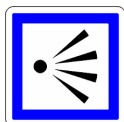
6. If $R = \mathbf{k}$ is a field, show that $\deg(\text{PGCD}(P, Q)) > 0$ if and only if there exist nonzero $A, B \in \mathbf{k}[T]$ of degree $< q$ and $< p$ respectively such that $AP = BQ$. Deduce that P, Q are coprime if and only if their resultant $\text{Res}(P, Q) \neq 0$.

Chapter 4

Equivalence Classes in $M_{p,q}(\mathbf{k}[T])$.



4.1 Perspective



We present the theory from a perspective as concrete and algorithmic as possible by generalizing classical Gaussian pivot techniques on matrices with coefficients in a field \mathbf{k} to the case where the coefficients belong to a polynomial rings $\mathbf{k}[T]$ (or a Euclidean ring).

As mentioned at the end of the chapter, there are good reasons to consider the pivot with values in rings R : the underlying presence of a new hidden group, the algebraic K-theory group $SK_1(R)$.

In the rest of this chapter, A will denote a polynomial rectangular matrix in $M_{p,q}(\mathbf{k}[T])$.

4.2 Introduction

The reader who has taken a basic course in group theory with the classification of finite abelian groups will recognize in this section a simple adaptation of what has been seen for matrices with integer coefficients (see exercices 10.4.0.3 and 10.4.0.4). It will therefore be a simple "reminder". For others, let's embark on the discovery.

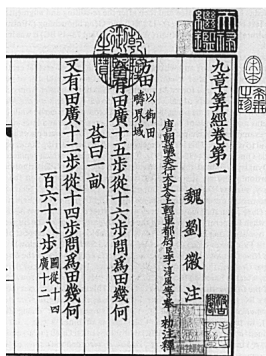
Recall that two matrices A, B in $M_{p,q}(R)$ are equivalent (denoted $A \sim B$) if and only if there exist $Q \in GL_q(R), P \in GL_p(R)$ such that $A = QBP^{-1}$. This indeed defines an equivalence relation. Because

elementary operations of matrices can be defined using left or right multiplication by suitable invertible matrices, we know that any equivalence class is invariant under elementary transformation.

We will essentially use this notion of equivalence of matrices only in the case of $R = \mathbf{k}[T]$. We will exhibit in each equivalence class of $M_{p,q/\sim}(\mathbf{k}[T])$ a canonical representative (4.3.2.3). The reader will generalize the statements of this section to any ring equipped with a Euclidean division by simply substituting $\mathbf{k}[T]$ with such a ring.

Remark(s) 4.2.0.1. *In the case $R = \mathbf{k}$, we know that two matrices with coefficients in \mathbf{k} are equivalent if and only if they have the same rank (an immediate application of the Gaussian elimination method, for example, or of the incomplete basis theorem, as preferred).*

4.3 Elementary Divisors



The nine chapters



Karl Friedrich Gauss

The elimination method was rediscovered by Gauss and Jordan in the 19th century. But it was known to the Chinese at least in the 1st century BCE ([12]).

4.3.1 Existence

Proposition 4.3.1.1. *There exists a family of monic polynomials $\underline{P} = (P_r | \dots | P_2 | P_1)$ such that A is equivalent to the diagonal matrix¹*

$$\Delta(\underline{P}) = \begin{pmatrix} \text{diag}(P_r, \dots, P_1) & 0_{r,q-r} \\ 0_{p-r,r} & 0_{p-r,q-r} \end{pmatrix}$$

Proof. We freely use elementary operations on matrices because they leave the equivalence class invariant. We can assume A is non-zero. We proceed by induction on $p + q \geq 2$. If $p + q = 2$, there is nothing to prove. Suppose the statement is proven for $p + q \leq n$ and let A be non-zero with $p + q = n + 1$.

Let d be the minimal degree of the non-zero coefficients of all matrices in the equivalence class of A . We can assume that this degree is attained for a coefficient of A .

- By permuting rows and/or columns, we can assume this coefficient is $a_{1,1}$.
- Using the suitable elementary dilatation $D(\delta)$ with δ the coefficient of the leading term of the polynomial $a_{1,1}$ (1.2), one can assume that $a_{1,1}$ is monic.
- $a_{1,1}$ necessarily divides all $a_{1,l}$ and $a_{l,1}$ for $l > 1$ (we can replace these coefficients by their remainder from the Euclidean division by $a_{1,1}$, which is zero due to the minimality of d). Using the same argument, we can assume $a_{1,l} = a_{l,1} = 0$ for $l > 1$.
- $a_{1,1}$ necessarily divides every coefficient $a_{i,j}$ with $i, j > 1$. Indeed, we can use elementary row operations to place $a_{i,j}$ on the first row. Then, by elementary column operations, $C_j \mapsto C_j - \alpha C_1$ with α being the quotient of the division of $a_{i,j}$ by $a_{1,1}$, we can reduce the degree to be less than d , resulting in a zero remainder due to the minimality of d .
- Thus, $a_{1,1} = P_r$ is the greatest common divisor of the coefficients, and A can be written in blocks as

$$P_r \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$$

with $B \in M_{p-1, q-1}(\mathbf{k}[T])$. We conclude by induction.

□

Remark(s) 4.3.1.2. Note that r is clearly the rank of A viewed as a matrix with coefficients in the field of fractions of $\mathbf{k}[T]$. Thus, it depends only on A .

This proof can be easily made algorithmic (5.14 or, alternatively, see [22] or [31]). We strongly encourage the reader to implement it themselves using a computer (for example, using the open-source software based on Python, SageMath²). This will be an excellent programming exercise.

4.3.2 What Uniqueness?

Let's recall that for any integer subsets $I \subset [1, \dots, p]$ and $J \subset [1, \dots, q]$ of the same cardinality n , the minor $A_{I,J}$ of A the square matrix $(a_{i,j})_{i \in I, j \in J}$.

We define for $n \geq 1$

$$\delta_n(A) = \text{GCD}(\wedge^n(A))$$

where $\wedge^n A$ is the ideal generated by all minors of order n of A . For instance, if a square matrix A is triangular and invertible, we have $\delta_i(A) = 1$ for all i .

Lemma 4.3.2.1. *If*

$$\Delta(\underline{P}) = \begin{pmatrix} \text{diag}(P_r, \dots, P_1) & 0_{r,q-r} \\ 0_{p-r,r} & 0_{p-r,q-r} \end{pmatrix} P_r | \dots | P_2 | P_1 \text{ monic}$$

then

$$\delta_n(A) = P_r \cdots P_{r-n+1}$$

with the convention here that $P_n = 0$ if $n \leq 0$.

Proof. All minors $\Delta_{I,J}$ of $\Delta = \Delta(\underline{P})$ are triangular with at least one zero diagonal element if $I \neq J$. If $I = (i_1 > \dots > i_n)$, we have $\det(\Delta_{I,I}) = P_{i_n} \cdots P_{i_1}$ if $n \leq r$ and is zero otherwise. If $n \leq r$, we have $i_j \leq r + 1 - j$ so that $P_r \cdots P_{r-n+1} | P_{i_n} \cdots P_{i_1}$ because of the decreasing property of P_i for divisibility. \square

Lemma 4.3.2.2. *Let $A, B \in M_{p,q}(\mathbf{k}[T])$. If A and B are equivalent, then*

$$\delta_n(A) = \delta_n(B) \text{ for all } n \geq 0.$$

Proof. Since the determinant of a matrix is equal to that of its transpose, we have $\delta_n(A) = \delta_n({}^t A)$ for all n . It follows that it suffices to show that for any matrix $P \in M_{q,r}(\mathbf{k}[T])$ (whether invertible or not) we have

$$\wedge^n(AP) \subset \wedge^n(A).$$

The learned reader will invoke the general Binet-Cauchy formula

$$\det((AP)_{I,J}) = \sum_{K | \text{Card}(K)=n} \det(A_{I,K}) \det(P_{K,J})$$

for computing minors of a product of arbitrary matrices. But we don't need that precision. We can proceed as follows. Each column of AP is a linear combination of columns of A . The multilinearity of the determinant then ensures that the minor $(AP)_{I,J}$ is a linear combination of determinants of matrices extracted of size n where the columns are columns of A (possibly equal) and the rows are indexed by I . If two columns are equal, the determinant is zero (the determinant is alternating). Otherwise, the set of columns in question is indexed by a set K of cardinality n and the determinant in question is of the form $A_{I,K}$ which implies that $\det(AP)_{I,J}$ is a linear combination of $\det(A_{I,K})$ with $\text{Card}(K) = n$, and therefore is indeed in $\wedge^n(A)$. \blacksquare

\square

From the previous calculation in the diagonal case (4.3.2.1) we obtain

Theorem 4.3.2.3 (Elementary Divisors of a Polynomial Matrix). *Let $A \in M_{p,q}(\mathbf{k}[T])$ be a polynomial matrix.*

- *There exists a unique sequence of monic polynomials $\underline{P} = (P_r, \dots, P_1)$ associated with A such that for all n we have $\delta_n(A) = P_r \cdots P_{r-n+1}$. These are called the elementary divisors of A .*
- *Two matrices in $M_{p,q}(\mathbf{k}[T])$ are equivalent if and only if they have the same elementary divisors.*
- *The sequence of elementary divisors of $\Delta(\underline{P})$ (4.3.2.1) is \underline{P} .*
- *If \underline{P} is the sequence of elementary divisors of A , then $A \sim \Delta(\underline{P})$. This sequence can be computed algorithmically using Gaussian pivot (4.3.1.1).*

Example(s) 4.3.2.4. *Let $A \in M_n(\mathbf{k}[T])$ be a matrix such that $a_{i,j} = 0$ if $i > j + 1$ and $a_{i+1,i} = 1$:*

$$\begin{pmatrix} * & * & * & \dots & * \\ 1 & * & * & \dots & * \\ 0 & 1 & * & * & \dots \\ & & \dots & & \\ 0 & \dots & 0 & 1 & * \end{pmatrix}$$

Then, the elementary divisors of A are $(1, \dots, 1, \det(A))$. Indeed, the $(n-1)$ minor A_{1^c, n^c} is upper triangular with diagonal entries equal to 1 showing $\delta_i(A) = 1$ for $i < n$ as already observed

Exercise(s) 4.3.2.5. *Let $P, Q \in \mathbf{k}[T]$ be monic polynomials and $A = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}$. Compute $\delta_1(A)$ and $\delta_2(A)$ and deduce that the similarity invariants of A are $\text{GCD}(P, Q), \text{lcm}(P, Q)$. Retrieve this result using the pivot.*

From this, deduce another algorithm than the pivot to compute the similarity invariants of a diagonal matrix in $\mathbf{k}[T]$. [If $Q_i, i \in I$ are its diagonal coefficients, consider $\text{GCD}(Q_{i_1} \cdots Q_{i_r})$ when $\{i_1, \dots, i_r\}$ runs through the r -element subsets of I].

Remark(s) 4.3.2.6. *The reader will easily adapt the previous theorem to the case of equivalence for matrices with coefficients in a Euclidean ring (equipped with Euclidean division). To do this, one simply needs to accept uniqueness of elementary divisors up to multiplication by units. The statement of uniqueness generalizes without change. If the existence (4.3.1.1) remains true in a principal ideal domain, its*

proof by pivot no longer works (exercise infra). Moreover, there exist principal ideal domains that are not Euclidean, such as $\mathbf{Z}[\frac{1+\sqrt{-19}}{2}]$ ([26]) or $\mathbf{R}[x,y]/(x^2 + y^2 + 1)$ ([4]). In the principal ideal case, one must therefore add one additional permitted operation (4.3.2.7). I emphasize that in general one must. This algorithmic difference is a window into algebraic K-theory (4.4).

Exercise(s) 4.3.2.7. Let R be a principal ideal ring, and consider elements of R such that $au - bv = 1$. We define Bézout operations on matrices with coefficients in R as left or right multiplications by block diagonal invertible matrices of the form:

$$\begin{pmatrix} \begin{pmatrix} a & v \\ b & u \end{pmatrix} & 0_{2,n} \\ 0_{n,2} & \text{Id}_n \end{pmatrix}$$

Generalize the proof of 4.3.1.1 by allowing Bézout operations in addition to elementary row operations.

4.3.3 Equivalence Classes in $M_{p,q}(\mathbf{k}[T])$

So we have solved our initial problem (4). Indeed, if A belongs to an equivalence class of $M_{p,q}(\mathbf{k}[T])/\sim$, its elementary divisors $\underline{P}(A) = \underline{P} = P_i$ are well defined and depend only on the class $(A \bmod \sim)$. Theorem 4.3.2.3 ensures that the quotient $M_{p,q}(\mathbf{k}[T])/\sim$ is identified with the set of sequences \underline{P} of decreasing monic polynomials of length $r \leq \min(p, q)$, and the quotient map is identified with $A \mapsto \underline{P}(A)$.

4.4 Supplementary Section: Insight into K-Theory



This section is cultural and can therefore be skipped at the first glance. It aims to introduce an important idea in mathematics: how to measure the obstruction to a result being true. Here, the question is how to measure the potential impossibility of *diagonalizing* matrices by means of Gaussian elimination in a ring R .

The precise question one naturally addresses is then: is the group $GL_n(R)$ generated by the elementary matrices of transvections of pivot type (1.2)? We will consider the matrices of permutation and dilatations (because they can be easily handled through the determinant function below).

The first step is to move away from n : for this, we view $GL_n(R)$ as the subgroup of $GL_{n+1}(R)$ consisting of block diagonal matrices of the form $\text{diag}(M, 1)$, where $M \in GL_n(R)$. This allows us to consider their infinite union $GL(R)$, seen as the set of matrices of infinite size, containing all finite-sized linear groups. We then define $E(A)$ as the subgroup of $GL(A)$ generated by all transvections with determinant 1 that we can reach by pivot (even if we allow enlarging the matrices).

The first result is both simple and remarkable, especially in the proof provided by [24].

Lemma 4.4.0.1 (Whitehead). *For any ring R , the group $E(R)$ is the derived group $[GL(R), GL(R)]$ generated by the commutators $[A, B] = ABA^{-1}B^{-1}$ of matrices in $GL(R)$.*

In particular, $E(A)$ is a normal subgroup, and the quotient $K_1(R) = GL(R)/[GL(R), GL(R)]$ is a commutative group, as it is the abelianization of $GL(R)$! This is the group of algebraic K-theory of degree 1. As the determinant of any commutator is 1, the determinant map passes to the quotient (3.6) to define the special group of algebraic K-theory of degree 1:

$$SK_1(R) = \text{Ker}(GL(R) \xrightarrow{\det} R^\times).$$

This group avoids considering dilations and permutation matrices, which do not play a crucial role in pivoting. The inclusion $R^\times = GL_1(R) \hookrightarrow GL(R)$ followed by the quotient projection $GL(R) \rightarrow K_1(R)$ allows us to define a map:

$$R^\times \times SK_1(R) \rightarrow K_1(R),$$

which is visibly an isomorphism.

The group $SK_1(R)$ is evidently the obstruction to the pivot algorithm (infinite) being able to diagonalize matrices. And our results prove that if R is Euclidean, $SK_1(R) = 0$. It is noteworthy that this obstruction is very sudden. For example, in the case of the non-Euclidean principal ring $R = \mathbf{Z}[\frac{1+\sqrt{-19}}{2}]$, we have $SK_1(R) = \{1\}$ (this follows from a general deep theorem about so-called Dedekind rings, [3]). In other words, this is not an example where the pivot with elementary matrices is insufficient, at least when allowing to increase the size of matrices. Finding a principal R such that $SK_1(R)$ is non-trivial is difficult. An example is given in [17]: take the subring of $\mathbf{Z}(T)$ generated by $\mathbf{Z}[T]$ and the $(T^m - 1)^{-1}$ for $m \geq 1$. This is a principal ring (!) whose SK_1 is even infinite.

4.5 Additional Exercises

Exercise(s) 4.5.0.1. *Let*

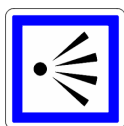
R be a Euclidean ring. Show that $SL_n(R)$ is generated by transvections.

Chapter 5

Similarity classes of $M_n(\mathbf{k})$



5.1 Point of view



In our study of linear algebra, we propose an "algorithmic" perspective on the reduction of endomorphisms (in finite dimensions) over an arbitrary field. We systematically adopt the dictionary between $\mathbf{k}[T]$ -modules and endomorphisms (cf. 3.2.4). Precisely, we try to do computable methods as far as possible, even we'll not try to find very practically efficient algorithms.

One motivation is the simplicity of the theory when we accept the language of modules, and, above all, the fact that the usual reduction theory uses, more or less explicitly, the roots of the characteristic polynomial, although we generally do not know how to calculate... As strange as it may seem, the module perspective makes the theory algorithmic and transparent, freeing us from the knowledge of these roots. Of course, eigenvalues still play an important underlying role, theoretically and often practically. We will clarify that, at least partially (cf. 6.2 and 8.3).

5.2 Introduction



René Descartes

Following Descartes' teachings, we will classify endomorphisms completely (5.9.0.2) and reduce them to two very simple classes: the well-named semi-simple endomorphisms and the nilpotent one (6.3.2.1). Rules of Descartes' method^a:

^aR. Descartes, *Discourse on the Method* (1637), Gallimard (2009).

1. *Not to accept anything as true that I did not clearly know to be so.*
2. *To divide each of the difficulties I examined into as many parts as possible and as might be required for their best resolution.*
3. *To conduct my thoughts in an orderly manner, beginning with the simplest and easiest-to-know objects in order to ascend little by little, as if by steps, to the knowledge of the most complex.*
4. *To make everywhere such complete enumerations, and such general reviews, that I might be assured of omitting nothing. This is the rule of enumeration. To make a complete review of objects, which involves prudence and circumspection.*

We will freely use the usual properties of principal rings (Bézout's identity, factoriality...) and the fact that Euclidean rings are principal. The key examples for us are \mathbf{Z} and $\mathbf{k}[T]$. The reader may, if necessary, refer (without circular reasoning) to chapter 11 for the factoriality of principal rings.

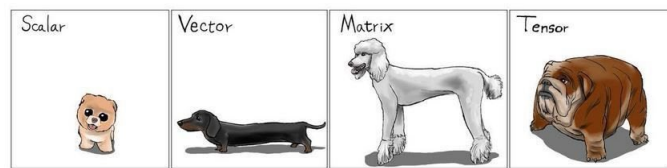
A useful result is a generalization of Euclidean division when the divisor has a leading coefficient that is a unit: this is the result *infra* whose proof is a simple rereading of the usual proof (exercise).

Lemma 5.2.0.1 (Generalized Euclidean division). *Let A, B be two polynomials with coefficients in a unital ring \mathcal{R} **not necessarily commutative** with B nonzero. Suppose that the leading coefficient of B is right (resp. left) invertible. Then there exist $Q_r, R_r \in \mathcal{R}[T]$ (resp. Q_l, R_l) such that $A = BQ_r + R_r$ with $\deg(Q_r) < \deg(R_r)$ right (resp. $A = Q_l B + R_l$ with $\deg(Q_l) < \deg(R_l)$ left) Euclidean division. If, moreover, \mathcal{R} is left (resp. right) integral, then there is uniqueness on the left (resp. right).*



Although it will be necessary to deal with nilpotent matrices to understand matrix reduction, one should not lose sight of the fact that these matrices are actually pathological. In the complex case, for example, a randomly drawn matrix has almost surely distinct eigenvalues and is therefore almost surely non nilpotent! One reason is provided as a warm-up exercise (5.15.0.3). However, mathematics naturally provides many "improbable" matrices.

5.2.1 Notations



In this chapter, we refer to (see 3):

- V as a finite-dimensional space of dimension n over an arbitrary field \mathbf{k} .
- $V[\mathbf{T}]$ as the $\mathbf{k}[\mathbf{T}]$ -module of polynomials with coefficients in V .
- $V_a, a \in \text{End}_{\mathbf{k}}(V)$ as the $\mathbf{k}[\mathbf{T}]$ -module $V = V_a$ characterized (3.2.4) by

$$\mathbf{T}v = a(v) \text{ for all } v \in V = V_a$$

and more generally, $P(\mathbf{T})v = P(a)(v)$.

- We will denote by $A, B \dots$ the matrices of $a, b \dots$ after choosing a basis for V .
- We denote by χ_a (resp. μ_a) the characteristic (resp. minimal)¹ polynomial of a .
- $\pi_a \in \text{Hom}_{\mathbf{k}[\mathbf{T}]}(V[\mathbf{T}], V_a)$ defined by

$$\pi_a\left(\sum v_i \mathbf{T}^i\right) = \sum a^i(v_i)$$

the canonical surjection extending the identity of V seen as the set of constant polynomials of $V[\mathbf{T}]$.

- $\mathbf{k}_\lambda = \mathbf{k}[\mathbf{T}]/(\mathbf{T} - \lambda)$ the module V_a with $V = \mathbf{k}$ and a being the multiplication by $\lambda \in \mathbf{k}$. We'll denote by 1_Λ the unit $1 \in \mathbf{k}_\lambda$.
- $\tilde{a} \in \text{End}_{\mathbf{k}[\mathbf{T}]}(V[\mathbf{T}])$ is the unique $\mathbf{k}[\mathbf{T}]$ -linear extension of $a \in \text{End}_{\mathbf{k}}(V)$ to $V[\mathbf{T}]$ characterized by

$$\tilde{a}(v\mathbf{T}^i) = a(v)\mathbf{T}^i.$$

¹It is not required at this stage to know even the existence of the minimal polynomial μ_a of a (see 5.6.0.2).

- $C(P)$ the companion matrix of the monic polynomial $P = T^n + \sum_{i=0}^{n-1} a_i T^i$

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in M_n(\mathbf{k}).$$

Thus, $C(P)$ is the empty matrix if $P = 1$,

- $J_n = C(T^n)$ the standard Jordan block

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \in M_n(\mathbf{k})$$

of size n .

Once a basis for V has been chosen, V will be identified without further precision as \mathbf{k}^n and $V[T]$ as $\mathbf{k}[T]^n$ (cf. 3.2.4 (2)), so that a and \tilde{a} have the same matrix.

5.3 Strategy

Recall that two square matrices A, B from $M_n(\mathbf{k})$ are similar (in $\mathbf{k}!$) (denoted $A \approx B$) if and only if there exists $P \in GL_n(\mathbf{k})$ such that $A = PBP^{-1}$. This indeed defines... an equivalence relation (to distinguish it from the previous equivalence of polynomial matrices, we denote by $\overline{\overline{A}}$ the similarity class of a square matrix $A \in M_n(\mathbf{k})$).

We will exhibit in each equivalence class a canonical representative. To do this, we will show for every square matrix $A \in M_n(\mathbf{k})$ a canonical representative in its similarity class in two steps, similar to what we did for polynomial matrices.

1. Note that $P(T \text{Id} - A)P^{-1} = T \text{Id} - PAP^{-1}$ such that if $A \approx B$, then $T \text{Id} - A \sim T \text{Id} - B$. The miracle is that the converse is true. We will therefore demonstrate that the mapping

$$\begin{cases} M_n(\mathbf{k})_{/\approx} & \hookrightarrow M_n(\mathbf{k}[T])_{/\sim} \\ \overline{\overline{A}} & \mapsto \overline{\overline{T \text{Id} - A}} \end{cases}$$

is injective: this is the corollary 5.4.0.2.

2. Similarly to the above, we will exhibit a canonical representative $C(\underline{P})$ in any similarity class $M_n(\mathbf{k})_{/\approx}$ using companion matrices associated with the elementary divisors \underline{P} of $T \text{Id} - A$, this sequence depending only on $\overline{\overline{T \text{Id} - A}}$ (4.3.2.3): this is the Frobenius decomposition theorem 5.9.0.2.

5.4 Invariance by equivalence of $\text{TId} - A$ of the module V_a and applications

Let $a \in \text{End}_{\mathbf{k}}(V)$ and V_a be the associated $\mathbf{k}[\text{T}]$ -module (3.2.4).

Lemma 5.4.0.1. *The sequence*

$$(i) \quad 0 \rightarrow V[\text{T}] \xrightarrow{\text{TId} - \tilde{a}} V[\text{T}] \xrightarrow{\pi_a} V_a \rightarrow 0$$

is exact.

Proof. Let $v \in V$. The image of the constant polynomial $v \in V[\text{T}]$ by π_a is v . Therefore π_a is onto.

We then have

$$\pi_a \circ (\text{TId} - \tilde{a})(v) = \text{T}\pi_a(v) - a(v) = a(v) - a(v) = 0$$

hence $\pi_a \circ (\text{TId} - \tilde{a}) = 0$ since V generates $V[\text{T}]$ and therefore $\text{Im}(\text{TId} - \tilde{a}) \subset \text{Ker}(\pi_a)$.

Conversely, let $v(\text{T}) = \sum_{i \geq 0} \text{T}^i v_i \in \text{Ker}(\pi_a)$, i.e.

$$v_0 + \sum_{i \geq 1} a^i(v_i) = 0.$$

Thus, we have

$$v(\text{T}) = \sum_{i \geq 1} (\text{T}^i \text{Id} - \tilde{a}^i)(v_i).$$

But since TId and \tilde{a} commute, we have (geometric series sum)

$$\text{T}^i \text{Id} - \tilde{a}^i = (\text{TId} - \tilde{a}) \circ \left(\sum_{j=0}^{i-1} \text{T}^j \tilde{a}^{i-1-j} \right)$$

and thus $v(\text{T}) \in \text{Im}(\text{TId} - \tilde{a})$. Hence the exactness in the middle. The exactness on the left, being easy and unnecessary for us, is left as an **useful exercise**. \square

In other terms, we have

$$(ii) \quad \text{Coker}(\text{TId} - \tilde{a}) = V_a$$

or if one is purist $\text{Coker}(\text{TId} - \tilde{a}) = \pi_a$ (3.4.0.1).

Choosing a basis of V identifies V to \mathbf{k}^n and $V[\text{T}]$ to $\mathbf{k}[\text{T}]^n$ (3.2.4). The matrix of \tilde{a} is then A , the matrix of a in the chosen base. The aforementioned exact sequence (i) is thus identified to

$$(iii) \quad 0 \rightarrow (\mathbf{k}[\text{T}])^n \xrightarrow{\text{TId} - A} (\mathbf{k}[\text{T}])^n \xrightarrow{\pi_A} V_A = (\mathbf{k}^n)_A \rightarrow 0$$

with $\pi_A(\sum X_i \text{T}^i) = \sum A^i X_i$ and $\text{T}.X = AX$ for every $X_i, X \in \mathbf{k}^n$. We deduce the important result

Corollary 5.4.0.2. *Let $A, B \in M_n(\mathbf{k})$ be the matrices of $a, b \in \text{End}_{\mathbf{k}}(V)$ in a base. The following propositions are equivalent.*

1. A and B are similar in $M_n(\mathbf{k})$.
2. $T \text{Id} - A$ and $T \text{Id} - B$ are equivalent in $M_n(\mathbf{k}[T])$.
3. The $\mathbf{k}[T]$ -modules V_a and V_b are isomorphic.

Moreover, if $T \text{Id} - A \simeq \Delta \in M_n(\mathbf{k}[T])$, then $V_a \simeq \text{Coker}(\Delta : \mathbf{k}[T]^n \rightarrow \mathbf{k}[T]^n)$.

Proof. $1 \Rightarrow 2$. If $P \in \text{GL}_n(\mathbf{k})$ satisfies $PAP^{-1} = B$, then $P(T \text{Id} - A)P^{-1} = T \text{Id} - B$ and thus $T \text{Id} - A \sim T \text{Id} - B$.

$2 \Rightarrow 3$. There exist $P(T), Q(T) \in \text{GL}_n(\mathbf{k}[T])$ such that $P(T)(T \text{Id} - A)Q(T)^{-1} = T \text{Id} - B$. In this case, $P(T), Q(T)$ define, according to the foregoing, a commutative diagram with exact rows and isomorphism columns

$$\begin{array}{ccccc} \mathbf{k}[T]^n & \xrightarrow{T \text{Id} - A} & \mathbf{k}[T]^n & \xrightarrow{\pi_A} & V_A \longrightarrow 0 \\ Q(T) \downarrow & & \downarrow P(T) & & \\ \mathbf{k}[T]^n & \xrightarrow{T \text{Id} - B} & \mathbf{k}[T]^n & \xrightarrow{\pi_B} & V_B \longrightarrow 0 \end{array}$$

and thus by functoriality of the cokernel (3.4.0.1) a unique $\mathbf{k}[T]$ -linear isomorphism

$$\iota : V_A \rightarrow V_B,$$

i.e. (3.2.4.1) an invertible matrix

$$S : V_A = \mathbf{k}^n \rightarrow \mathbf{k}^n = V_B$$

satisfying $SA = BS$ (since $\iota(T.v) = \iota(a(v)) = T.\iota(v) = b(\iota(v))$).

$2 \Rightarrow 3$. It has already been noted (3.2.4.1) that the existence of the isomorphism $\iota : V_A = \mathbf{k}^n \rightarrow \mathbf{k}^n = V_B$ defines $S \in \text{GL}_n(\mathbf{k})$ such that $SA = BS$. \square

The equivalence of the first two points is rewritten as $A \approx B$ if and only if $T \text{Id} - A \sim T \text{Id} - B$, representing the sought injectivity in (3.2.4.)

5.5 Similarity Invariants of $a \in \text{End}_{\mathbf{k}}(V)$

Following corollary 5.4.0.2, it is reasonable to propose the following definition.

Definition 5.5.0.1. *The elementary divisors of $T \text{Id} - A, A \in M_n(\mathbf{k})$ are called the similarity invariants of A .*

Corollary 5.4.0.2 is then rewritten

Theorem 5.5.0.2 (Similarity Invariants). *Let $a, b \in \text{End}_{\mathbf{k}}(V)$. The following propositions are equivalent.*

1. a and b have the same similarity invariants \underline{P} .
2. a and b are similar in $\text{End}_{\mathbf{k}}(V)$.
3. $T\text{Id} - A$ and $T\text{Id} - B$ are equivalent to $\Delta(\underline{P})$
4. The $\mathbf{k}[T]$ -modules V_a and V_b are isomorphic.
5. The $\mathbf{k}[T]$ -modules V_a and $\text{Coker}(\Delta(\underline{P}))$ are isomorphic.



Corollary 5.5.0.3. *Let $A, B \in M_n(\mathbf{k})$ and K be a overfield of \mathbf{k} . Then,*

- A and B are similar over K if and only if they are similar over \mathbf{k} .
- A and tA are similar.

Proof. The first point follows, for example, from the fact that the similarity invariants of A are computed by Gaussian elimination on $T\text{Id} - A$, an algorithm independent of the overfield in which it is computed. The second results from two observations: $A \sim B$ implies ${}^tA \sim {}^tB$ (write the definition of equivalence) and $A \sim \Delta(\underline{P}) = {}^t\Delta(\underline{P})$ because Δ is a diagonal square matrix. \square

Remark(s) 5.5.0.4. *Another way to say the same thing is the following: let $(A, B), (A', B')$ be square matrices with A invertible. Then, there exist P, Q invertible such that $PAQ = A'$ and $PBQ = B'$ if and only if A' is likewise and $A^{-1}B$ and $A'^{-1}B'$ have the same similarity invariants. Indeed, the direct part is obvious because in this case A' is invertible $(A'^{-1}PA)(A^{-1}B + T)Q = (A'^{-1}B' + T)$. Conversely, if $A^{-1}B$*

and $A'^{-1}B'$ have the same similarity invariants, there exists an invertible Π such that $\Pi^{-1}A^{-1}B\Pi = A'^{-1}B'$ and we set $P = A'\Pi^{-1}A^{-1}$, $Q = \Pi$.

Before moving to the second point announced in (3.2.4), the Frobenius decomposition 5.9.0.2, let us provide some specific properties related to the similarity invariants A linked to the very specific nature of the polynomial matrix $T\text{Id} - A$ and draw some delightful corollaries.

5.6 Calculation of V_a and Applications

Let A be a matrix of a in some basis. Since $V_a = \text{Coker}(T\text{Id} - A)$, it depends up to isomorphism only on the equivalence class of $T\text{Id} - A$ and thus on its elementary divisors \underline{P} , which by definition are the similarity invariants of a . Since $T.\text{Id} - A \sim \Delta(\underline{P})$, it is sufficient to compute $V_a = \text{Coker}(\Delta(\underline{P}))$ in this diagonal case.

Proposition 5.6.0.1. *Let $a \in \text{End}_{\mathbf{k}}(V)$ and $\underline{P} = (P_r | \cdots | P_1)$ be its similarity invariants.*

1. We have $r = n$ and $\prod_{i=1}^n P_i = \chi_a(T)$.
2. We have $\Delta(\underline{P}) = \text{diag}(P_1, \dots, P_n)$.
3. We have $P_1 | \chi_a | P_1^n$ so that χ_a and P_1 have the same irreducible factors (and hence the same roots in any extension of \mathbf{k}).
4. The $\mathbf{k}[T]$ -module V_a is isomorphic to $\bigoplus_{i=1}^n \mathbf{k}[T]/(P_i)$.
5. $P(a) = 0$ if and only if $P_1 | P$. In other words, μ_a is the minimal polynomial of a .
6. $\chi_a(a) = 0$ (Cayley-Hamilton).

Proof. Let $A \in M_n(\mathbf{k})$ be the matrix of a in a basis V such that $T\text{Id} - A \sim \Delta(\underline{P})$ (4.3.2.3). Hence, there exist $P(T), Q(T)$ in $GL_n(\mathbf{k}[T])$ such that

$$P(T) \begin{pmatrix} \text{diag}(P_1, \dots, P_r) & 0_{r, n-r} \\ 0_{n-r, r} & 0_{n-r, n-r} \end{pmatrix} Q(T) = T\text{Id} - A.$$

Since the determinants of $P(T)$ and $Q(T)$ are nonzero scalars and both the P_i and the characteristic polynomial χ_A are monic, by taking the determinant of the preceding identity $r = n$ and $\chi_A(T) = P_1 \cdots P_n$ hence (1) and (2).

Because P_1 is a multiple of each P_i , by taking the product, we find that P_1^n is a multiple of χ_a , thus $P_1 | \chi_a | P_1^n$ thanks to (1), hence (3).

According to (3.3.2.2), the sequence

$$(\mathbf{k}[T])^n \xrightarrow{\Delta} (\mathbf{k}[T])^n \rightarrow \bigoplus_{i=1}^n \mathbf{k}[T]/(P_i) \rightarrow 0$$

is exact and is identified by functoriality of the cokernel (cf. 5.5.0.2)

$$\text{Coker}(\Delta) = \bigoplus_{i=1}^n \mathbf{k}[T]/(\Delta_{i,i}) = V_a$$

hence (4).

Since P_1 is a multiple of each P_i , it annihilates all the $\mathbf{k}[T]/(P_i)$. Conversely, if $P \in \mathbf{k}[T]$ annihilates $\bigoplus_{i=1}^n \mathbf{k}[T]/(P_i)$, it is a multiple of each P_i and therefore (equivalently) of P_1 . Because V_a is isomorphic to $\bigoplus_{i=1}^n \mathbf{k}[T]/(P_i)$ as $\mathbf{k}[T]$ -modules, this means exactly $P(a) = 0$ if and only if $P_1|P$. In other words, P_1 is indeed the minimal μ_a of a hence points (5) and (6). \square

Remark(s) 5.6.0.2. • Notice that the above proposition 5.6.0.1 proves the very existence of μ_a without any previous knowledge. By construction, it is the unique monic polynomial of least degree annihilating a .

- As we will see later (for example 5.8.0.1), the last P_i are often equal to 1. They contribute by the zero module to V_a .

5.7 Diagonalization

We keep the notations of 5.6.0.1.

Corollary 5.7.0.1. *The endomorphism a is diagonalizable if and only if its minimal polynomial $P_1 = \mu_a$ is split over \mathbf{k} with simple roots. In particular, the restriction of a diagonalizable to a stable subspace is diagonalizable.*

Proof. Suppose P_1 , and hence all P_i (which divide it), are split with simple roots. Then, according to (5.6.0.1) and the Chinese Remainder Theorem 3.7.0.1, we have an isomorphism

$$\iota^{-1} : V_a \simeq \bigoplus_i \mathbf{k}[T]/P_i \simeq \bigoplus_i \bigoplus_{\lambda|P_i(\lambda)=0} \mathbf{k}[T]/(T - \lambda) = \bigoplus_i \bigoplus \mathbf{k}_\lambda.$$

This isomorphism identifies a with the multiplication of T on the right hand side. Because $T \cdot 1_\lambda = \lambda 1_\lambda$ by the very definition of \mathbf{k}_λ (5.2.1), the various $\iota(1_\lambda)$ define a diagonalization basis of a . The converse is clear.

If W is stable by a , the restriction a_W of a to W is annihilated by μ_a which is thus a multiple of its minimal polynomial μ_{a_W} . Thus, μ_{a_W} has simple roots like μ_a . \square

Remark(s) 5.7.0.2. This criterion is often used in the following equivalent form: an endomorphism a is diagonalizable if and only if it admits an annihilating polynomial split over \mathbf{k} with simple roots. For example, any complex matrix satisfying $A^N = \text{Id}$ is diagonalizable. It's however no longer true that a finite order element of $\text{GL}(\Omega)$ is diagonalizable if Ω is of positive characteristic (*find an example*).

Matrices of a family of diagonal matrices commute pairwise. It is remarkable and important that the converse is true. The main point is the following easy but important lemma.

Lemma 5.7.0.3. If $a, b \in \text{End}_{\mathbf{k}}(V)$ commute, then any eigenspace of a is b -stable.

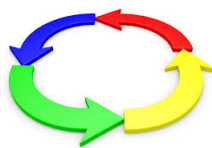
Proof. Let $v \in \text{Ker}(a - \lambda \text{Id})$. One has $a(b(v)) = b(a(v)) = b(\lambda v) = \lambda b(v)$ proving $b(v) \in \text{Ker}(a - \lambda \text{Id})$. \square

Corollary 5.7.0.4. Let (a_i) be an arbitrary family of diagonalizable endomorphisms of V . Then, if $f_i \circ f_j = f_j \circ f_i$ for all i, j , there exists a common diagonalization basis for all the f_i .

Proof. We use induction on $n = \dim(V) \geq 0$. We may assume that $n > 0$ and that the statement is true in dimension $< n$. If all the f_i are homotheties $\lambda_i \text{Id}$, any base is suitable. Otherwise, let i such that f_i is not a homothety. Then, f_i has at least two distinct eigenvalues so that all its eigenspaces $E_i(\lambda)$ are of dimension $< n$. But they are stable by all the f_j and their restrictions $f_j(\lambda)$ to each $E_i(\lambda)$ are diagonalizable for all j . For each λ , we then choose a common diagonalization base for the $f_j(\lambda)$ and the union of these bases suits. \square

We now move to the second point announced in (5.3). Thus, we are looking for a canonical representative $C(\underline{P})$ in every similarity class $\overline{\overline{A}}$ just as we found the representative $\Delta(\underline{P})$ in $\overline{\overline{\text{TId} - A}}$. The difficulty is that $\Delta(\underline{P})$ is not of the form $\text{T} \cdot \text{Id} - A'$. But this is not a problem, as we will see. Let's start with the case where only one of the similarity invariants is of degree > 0 .

5.8 Cyclic Endomorphisms



Let V be of dimension n and let $P = T^n + \sum_{i=0}^{n-1} a_i T^i \in \mathbf{k}[T]$.

Proposition 5.8.0.1. *Let $a \in \text{End}_{\mathbf{k}}(V)$. The following statements are equivalent:*

1. *The matrix of A in a suitable basis is the companion matrix $C(P)$.*
2. *$\mu_a = \chi_a = P$.*
3. *The similarity invariants are $1, \dots, 1, P$.*
4. *V_a and $\mathbf{k}[T]/(P)$ are isomorphic $\mathbf{k}[T]$ -modules.*
5. *V_a is cyclic as ($\mathbf{k}[T]$ -module) and $\mu_a = P$.*

Proof. $1 \Rightarrow 2$. If $e_i, 0 \leq i \leq n-1$, is the basis in question, then $e_i = u^i(e_0)$ and

$$u^n(e_0) = -\sum_{i < n} a_i e_i = \sum_{i < n} a_i u^i(e_0).$$

Thus, $V_a = \mathbf{k}[T].e_0$ and $P(T).e_0 = 0$, hence $\mu_a | P$. Since $u^i(e_0), i < n$, are free, there does not exist a monic polynomial Q of degree $< n$ such that $Q(a) = 0$ because otherwise $Q(a)(e_0) = 0$ would be a linear relation, and thus $\mu_a = P$ and $\mu_a = \chi_a$ due to degree reasons since $\mu_a | P$.

$2 \Rightarrow 3$. Saying $\mu_a = P$ implies $P_1 = \chi_a$ and thus $P_i = 1$ for $i > 1$ as per point 1 of proposition 5.6.0.1.

$3 \Rightarrow 4$ as per point 4 of proposition 5.6.0.1.

$4 \Rightarrow 5$ is tautological: we have already seen $P = \mu_a$ in the proof of point 5 of 5.6.0.1. If ι is then an isomorphism of $\mathbf{k}[T]/(P)$ onto V_a , the element $\iota(1 \pmod{P})$ generates V_a .

$5 \Rightarrow 1$. Let e_0 be a generator of V_a . The kernel of the unique $\mathbf{k}[T]$ surjective morphism $\mathbf{k}[T] \rightarrow V_a$ that sends 1 to e_0 is the annihilator of e_0 in V_a , and thus contains $\mu_a = P$. But if its monic generator were Q of degree $< n$, then $Q(T)e_0 = 0$ but also $Q(T)V_a = (0)$ since $V_a = \mathbf{k}[T].e_0$. Therefore, $\mu_a = P$ would divide Q , which is impossible since $\deg(P) = n$. \square

Remark(s) 5.8.0.2. *Observe that we already knew the similarity invariants (4.3.2.4). Notice also that any cyclic submodule $\mathbf{k}[T].v$ of V_a is a quotient of $\mathbf{k}[T]$ and therefore of $\mathbf{k}[T]/(\mu_a)$ because $\mu_a(T).v = \mu_a(a)(v) = 0$. In particular, $\dim \mathbf{k}[T].v \leq \deg(\mu_a)$. A vector v such that $\mathbf{k}[T].v$ has the maximal dimension $\deg(\mu_a)$ is called cyclic. In other words, v is cyclic if $\mathbf{k}[T].v$ is a cyclic subspace of maximal dimension.*

5.9 Frobenius Decomposition

Definition 5.9.0.1. *Let $\underline{P} = (P_n, \dots, P_1)$ be a sequence of monic polynomials. We define the generalized companion matrix $C(\underline{P})$ of size $n = \sum \deg(P_i)$ by $C(\underline{P}) = \text{diag}(C(P_i))$ (5.2.1).*



We can now conclude with the second point announced in (5.3). We are thus looking for a canonical representative $C(\underline{P})$ in each similarity class \overline{A} .

Ferdinand Georg Frobenius

Note that $C(1)$ is... the empty matrix, as any matrix of endomorphism of $\mathbf{k}[T]/(1) = (0)$!

Theorem 5.9.0.2 (Frobenius Reduction). *Let $\underline{P}=(P_n|\cdots|P_1)$, $i = 1, \dots, n$ be monic polynomials of $\mathbf{k}[T]$ and $A \in M_n(\mathbf{k})$.*

1. *The family of similarity invariants of $C(\underline{P})$ is \underline{P} .*
2. *If \underline{P} is the family of similarity invariants of A , then A is similar to $C(\underline{P})$.*

Proof. Let r be the highest index i such that $d_i = \deg(P_i) > 0$. According to the characterization of cyclic endomorphisms (5.8.0.1), for every $i \leq r$, the matrix $T \text{Id} - C(P_i)$ is equivalent to $\text{diag}(1, \dots, 1, P_i)$ (with 1 repeated $d_i - 1$ times) and thus can be written as $Q'_i \text{diag}(1, \dots, 1, P_i) Q_i^{-1}$ with $Q_i, Q'_i \in \text{GL}_{\deg(P_i)}(\mathbf{k})$, while $C(P_i)$ is empty for $i > r$. Thus with $Q = \text{diag}(Q_i), Q' = \text{diag}(Q'_i), i \leq r$

$$T \text{Id} - A = Q \text{diag}_i \left(\text{diag}(1, \dots, 1, P_i) \right) Q'^{-1} \sim \text{diag}(P_1, \dots, P_r, 1, \dots, 1)$$

with 1 repeated $\sum_{i \leq r} (d_i - 1) = n - r$ times thus $\text{diag}(P_1, \dots, P_r, 1, \dots, 1) = \text{diag}(P_1, \dots, P_n)$. By uniqueness of elementary divisors, (1) follows.

For (2), from (1) it follows that A and $C(\underline{P})$ have the same similarity invariants, therefore are similar. \square

Exercise(s) 5.9.0.3. *Let $\alpha, \beta \in \mathbf{k}$ and $a \in \text{End}_{\mathbf{k}}(V)$. Compute the similarity invariants of $\alpha a + \beta \text{Id}$ based on α, β , and the invariants of a .*

Exercise(s) 5.9.0.4. *Compute the similarity invariants of 2 by 2 matrix. In the real case, are these invariants continuous with respect of the matrix coefficients ?*

5.9.1 Equivalent Formulation

With the previous notations, we have $V_a = \oplus V_i$ where $V_i \simeq \mathbf{k}[T]/(P_i)$ as $\mathbf{k}[T]$ -modules. In particular, the antecedent $v_i \in V_i$ of $1 \in \mathbf{k}[T]/(P_i)$ generates V_i : it is a cyclic vector of the restriction of a to V_i . In other words, the unitary generator μ_{a, v_i} of the ideal of polynomials P such that $P(a)(v_i) = 0$ is of maximal degree, namely the degree of the minimal polynomial of $a|_{V_i}$. Alternatively, equivalently, $\mu_{a, v_i} = \mu_{a|_{V_i}}$. We can rewrite the Frobenius theorem by stating that there exists a decomposition $V_a = \oplus V_i$ where each V_i is cyclic of minimal polynomial P_i with the usual divisibility condition. Moreover, given such a decomposition, the P_i are the similarity invariants.

5.10 Summary

Collating what we have proved, we have the following results.

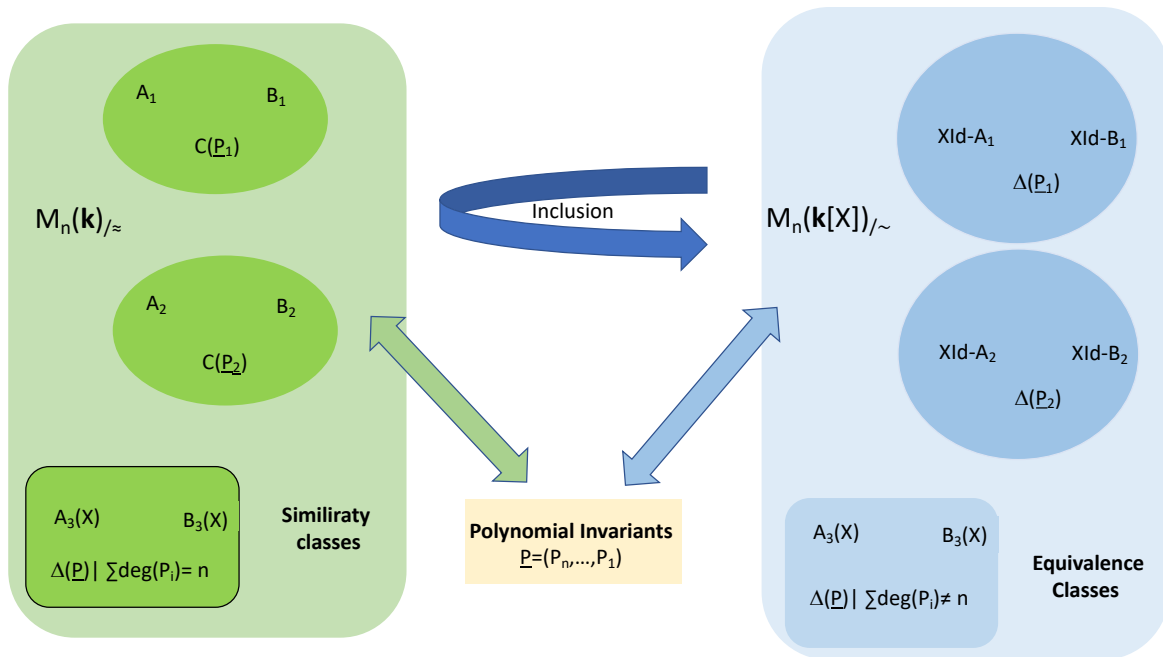
Let $A, B \in M_n(\mathbf{k})$ and $\underline{P} = (P_n | \dots | P_1)$ a family of monic polynomials.

- A and B are similar if and only if they have the same similarity invariants.
- The family of similarity invariants of $C(\underline{P})$ is \underline{P} .

If \underline{P} is the family of similarity invariants of A, we have:

- A and $C(\underline{P})$ are similar.
- $V_A \simeq \oplus \mathbf{k}[T]/(P_i)$ where A also denotes the endomorphism of $V = \mathbf{k}^n$ associated.
- $T \text{Id} - A$ is equivalent to $\text{diag}(P_1, \dots, P_n)$.
- \underline{P} is calculated by Gauss elimination by "diagonalizing" $T \text{Id} - A$ in $M_n(\mathbf{k}[T])$.
- We have $\chi_A = P_1 \dots P_n$ and $P_1 = \mu_A$.
- The similarity invariants of $C(P)$ are $(1, \dots, 1, P)$.

The proof strategy is illustrated by the following diagram.



5.11 Application: Commutant

It is then easy to study the commutant (see ii)

$$\text{Com}(a) = \text{End}_{\mathbf{k}[T]}(V_a) \simeq \text{End}_{\mathbf{k}[T]}(\oplus \mathbf{k}[T]/(P_i)).$$

for example, to calculate its dimension.

Proposition 5.11.0.1. *The dimension of $\text{Com}(a)$ is $\sum (2i - 1) \deg(P_i)$. In particular, $\dim \text{Com}(a) \geq n$ with equality if and only if a is cyclic.*

Proof. We have

$$\text{End}_{\mathbf{k}[\mathbf{T}]}(\oplus \mathbf{k}[\mathbf{T}]/(P_i)) = \oplus_{i,j} \text{Hom}_{\mathbf{k}[\mathbf{T}]}(\mathbf{k}[\mathbf{T}]/(P_i), \mathbf{k}[\mathbf{T}]/(P_j))$$

Since $\mathbf{k}[\mathbf{T}]/(P_i)$ is cyclic generated by the class of 1, an element of

$$\text{Hom}_{\mathbf{k}[\mathbf{T}]}(\mathbf{k}[\mathbf{T}]/(P_i), \mathbf{k}[\mathbf{T}]/(P_j))$$

is determined by its image $(P \bmod P_j)$ where P satisfies

$$(*) \quad P_i P \equiv 0 \pmod{P_j}$$

(universal property of the quotient 3.6.0.1). If $i \leq j$, we have $P_j | P_i$, and this condition is automatically satisfied so that

$$\text{Hom}_{\mathbf{k}[\mathbf{T}]}(\mathbf{k}[\mathbf{T}]/(P_i), \mathbf{k}[\mathbf{T}]/(P_j)) \simeq \mathbf{k}[\mathbf{T}]/(P_j) \text{ if } i \leq j$$

If $i > j$, we have $P_i | P_j$ so the condition $(*)$ reads $P \equiv 0 \pmod{P_j/P_i}$ so that

$$\text{Hom}_{\mathbf{k}[\mathbf{T}]}(\mathbf{k}[\mathbf{T}]/(P_i), \mathbf{k}[\mathbf{T}]/(P_j)) \simeq P_j/P_i \mathbf{k}[\mathbf{T}]/(P_j) \simeq \mathbf{k}[\mathbf{T}]/(P_i) \text{ if } i > j$$

We therefore have

$$\begin{aligned} \dim_{\mathbf{k}}(\text{Com}(a)) &= \sum_{i \leq j} \deg(P_j) + \sum_{i > j} \deg(P_i) \\ &= \sum_j j \deg(P_j) + \sum_i (i - 1) \deg(P_i) \\ &= \sum (2i - 1) \deg(P_i) \end{aligned}$$

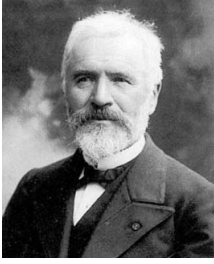
Using $n = \sum \deg(P_i)$, we get $\dim \text{Com}(a) - n = 2 \sum_{i=1}^n (i - 1) \deg(P_i) \geq 0$. Furthermore, equality implies $(i - 1) \deg(P_i) = 0$ for every i , thus $\deg(P_i) = 0$ if $i > 1$ so that equality is equivalent to the cyclicity of a . \square

Exercise(s) 5.11.0.2 (Bicommutant, difficult). *Show that the inclusion $\mathbf{k}[a] \subset \text{Com}(\text{Com}(a))$ is an equality where $\text{Com}(\text{Com}(a))$ is the set of endomorphisms that commute with all elements of $\text{Com}(a)$.*

5.12 Application: Jordan Reduction

Let $A \in M_n(\mathbf{k})$ and \underline{P} the similarity invariants of A . Assume χ_A splits over \mathbf{k} and denote by Λ the set of its distinct roots. One gets

$$\chi_A(\mathbf{T}) = \prod_{\lambda \in \Lambda} (\mathbf{T} - \lambda)^{d_\lambda}.$$



Camille Jordan

Let us explain why the Frobenius reduction immediately leads to the Jordan reduction of endomorphisms with a split characteristic polynomial. We retain the previous notations (and remind that a matrix of size ≤ 0 is an empty matrix).

If we specialize to the case $\chi_A = T^n$, we have $P_i = T^{d_i}$ with $d_i \geq 0$ decreasing and $\sum d_i = d$.

Definition 5.12.0.1. A partition of an integer $n \geq 0$ is a decreasing sequence $\underline{d} = (d_i)_{1 \leq i \leq n}$ of integers ≥ 0 such that $\sum d_i = n$.

Since each P_i divides χ_A , we have

$$(iv) \quad P_i = \prod_{\lambda} (T - \lambda)^{d_{\lambda,i}} \text{ where } \underline{d}_{\lambda} = (d_{\lambda,i})_i \text{ is a partition of } d_{\lambda}.$$

By applying the Chinese remainder theorem 3.7.0.6, we have

$$V_A = \bigoplus_{\lambda} \bigoplus_i \mathbf{k}[T]/((T - \lambda)^{d_{\lambda,i}}).$$

Let $\mathcal{B}_{\lambda,i} = ((T - \lambda_j) \bmod (T - \lambda)^{d_{\lambda,i}})_{j < d_{\lambda,i}}$. It is a \mathbf{k} -basis of $\mathbf{k}[T]/((T - \lambda)^{d_{\lambda,i}})$. The formula

$$T(T - \lambda)^j = (T - \lambda)^{j+1} + \lambda_j(T - \lambda)^j$$

ensures that the matrix $\text{Mat}_{\mathcal{B}_{\lambda,i}}(T)$ of multiplication by T on $\mathbf{k}[T]/((T - \lambda)^{d_{\lambda,i}})$ is $\lambda + J_{d_{\lambda,i}}$ where $J_m = C(T^m)$ is the standard Jordan block of size m (5.2.1). Thus, we have

Theorem 5.12.0.2 (Jordan Reduction). Under the assumptions and notations above, we have:

1. A is similar to a unique diagonal matrix $\text{diag}(\lambda + J_{d_{i,\lambda}})$ with for every λ the sequence $(d_{i,\lambda})_i$ being a partition of d_{λ} .
2. In particular, if $\chi_A = T^n$ (i.e., A is nilpotent), there exists a unique partition $\underline{d} = (d_i)$ of n verifying A is similar to the diagonal block matrix $J_{\underline{d}} = \text{diag}(J_{d_n}, \dots, J_{d_1})$. The similarity invariants of A are $T^{d_n}, T^{d_{n-1}}, \dots, T^{d_1}$.

Remark(s) 5.12.0.3. The uniqueness follows from the fact that once the Jordan reduction is given, the calculation of its similarity invariants follows. Indeed, if the Jordan form consists of blocks of the type $\lambda \text{Id}_r + J_r$, each such block is associated with the polynomial $(T - \lambda)^r$. For each eigenvalue λ , we order the blocks that appear in descending order, and write down the corresponding polynomials in columns

$$\begin{array}{cccc} (T - \lambda_1)^{d_{1,1}} & (T - \lambda_2)^{d_{1,2}} & \dots & \\ (T - \lambda_1)^{d_{2,1}} & (T - \lambda_2)^{d_{2,2}} & \dots & \\ \vdots & \vdots & & \end{array}$$

with $d_{i+1,j} \leq d_{i,j}$. We then read off the invariant factors P_1, P_2 , etc., from the rows (starting from the last one).

Exercise(s) 5.12.0.4. Let $M \in M_n(\mathbf{k})$ be a nilpotent matrix.

1. Show that $\text{rk}(M) = n - 1$ if and only if the Jordan reduction is J_n .
2. If $\mathbf{k} = \mathbf{R}$, show that the set of nilpotent matrices of rank $n - 1$ is the largest open set of the set of nilpotent matrices on which the Jordan reduction is continuous (with the topology defined by a norm on $M_n(\mathbf{R})$).
3. Show that $\text{rk}(M) = n - 2$ if and only if M has exactly two Jordan blocks J_p, J_{n-p} where p is the index of nilpotency of M . Show that $p \geq n/2$.
4. Let $p \geq n/2$, an integer $q = n - p$, and set for $t \in \mathbf{k}$, let $M_t = \text{diag}(J_p, J_q) + tE_{p+q,p}$ (adding t at the bottom of the p -th column). Calculate the index of nilpotency of M_t depending on t . Deduce that the Jordan reduction of M_t is $\text{diag}(J_{p+1}, J_{q-1})$ if $t \neq 0$ and $\text{diag}(J_p, J_q)$ otherwise.
5. Assume $\mathbf{k} = \mathbf{R}$. What is the set of continuity of the Jordan reduction application restricted to the subset of nilpotent matrices of rank $n - 2$ (with the topology defined by a norm on $M_n(\mathbf{R})$)?

5.12.1 Examples

(1) The elementary divisors of the Jordan reduction

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix}$$

(where $\lambda \neq \mu$), are

$$\begin{aligned} & (T - \lambda)^2 \quad (T - \mu) \\ & (T - \lambda)^2 \\ & (T - \lambda). \end{aligned}$$

The similarity invariants are thus

$$(T - \lambda), \quad (T - \lambda)^2, \quad (T - \lambda)^2(T - \mu).$$

(2) If $M = \begin{pmatrix} 0 & 4 & 2 \\ -1 & -4 & -1 \\ 0 & 0 & -2 \end{pmatrix}$, we have

$$TI - M = \begin{pmatrix} T & -4 & -2 \\ 1 & T+4 & 1 \\ 0 & 0 & T+2 \end{pmatrix}.$$

Let's perform elementary operations according to the algorithm - or rather its outline - described in the proof of the proposition 4.3.1.1 :

$$\begin{aligned} & \begin{pmatrix} T & -4 & -2 \\ 1 & T+4 & 1 \\ 0 & 0 & T+2 \end{pmatrix} \xrightarrow{L_1 \leftrightarrow L_2} \begin{pmatrix} 1 & T+4 & 1 \\ T & -4 & -2 \\ 0 & 0 & T+2 \end{pmatrix} \\ & \xrightarrow{L_2 \rightarrow L_2 - TL_1} \begin{pmatrix} 1 & T+4 & 1 \\ 0 & -4 - T(T+4) & -2 - T \\ 0 & 0 & T+2 \end{pmatrix} \xrightarrow{\begin{matrix} C_2 \rightarrow C_2 - (T+4)C_1 \\ C_3 \rightarrow C_3 - C_1 \end{matrix}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & (T+2)^2 & -2 - T \\ 0 & 0 & T+2 \end{pmatrix} \\ & \xrightarrow{L_2 \rightarrow L_2 + L_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & (T+2)^2 & 0 \\ 0 & 0 & T+2 \end{pmatrix} \xrightarrow{\begin{matrix} C_1 \leftrightarrow C_2 \\ L_1 \leftrightarrow L_2 \end{matrix}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & T+2 & 0 \\ 0 & 0 & (T+2)^2 \end{pmatrix}. \end{aligned}$$

The similarity invariants are thus $T + 2$ and $(T + 2)^2$ and the Jordan reduction is $\begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$. An endomorphism with matrix M is not cyclic.

(3) If $M = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix}$, we obtain as the reduction for $T I - M$ the matrix

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

The invariant factors are $(T-1)^2$ and $(T-1)^2$, and the Jordan reduction is $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. An endomorphism with matrix M is not cyclic.

(4) An endomorphism is cyclic if and only if, for each eigenvalue, there is only one Jordan block.

5.12.2 Supplement on nilpotent matrices

Let A be a nilpotent matrix and \underline{d} the associated partition (5.12.0.2). Since the Jordan block J_p is the matrix in the canonical basis of the multiplication by T of $\mathbf{k}[T]/(T^p)$, the image of J_p^i is identified with $T^i \mathbf{k}[T]/(T^p) \simeq \mathbf{k}[T]/(T^{p-i})$. We derive the equality $\text{rk}(J_p^i) = (p-i)_+$ and more generally

$$(v) \quad \text{rk}(A^i) = \sum_j (d_j - i)_+$$

We set

$$d_i^* = \dim(\text{Im}(A^{i-1})/\text{Im}(A^i)) = \text{rk}(A^{i-1}) - \text{rk}(A^i), \quad i = 1, \dots, n$$

(with $A^0 = \text{Id}$) so that we have $\sum d_i^* = n$ and $d_i^* \geq 0$. Moreover, the multiplication by A induces a surjection $\text{Im}(A^{i-1})/\text{Im}(A^i) \rightarrow \text{Im}(A^i)/\text{Im}(A^{i+1})$ so that d_i^* decreases. We have by construction $\text{rk}(A^i) = \sum_{j>i} d_j^*$.

$$(vi) \quad n - \text{rk}(A^i) = \sum_{j \leq i} d_j^*$$

Definition 5.12.2.1. The partition $\underline{d}^* = (d_i^*)$ is said to be the dual partition of \underline{d} .

We now rewrite of the dual partition and prove that partition duality is involutive as usual!

Lemma 5.12.2.2. *With the previous notations, we have $d_i^* = \text{Card}\{j|d_j \geq i\}$ and $\underline{d}^{**} = \underline{d}$.*

Proof. We first write

$$\begin{aligned} d_i^* &= \sum_j (d_j - i + 1)_+ - (d_j - i)_+ \\ &= \sum_{j|d_j \geq i} (d_j - i + 1)_+ - (d_j - i)_+ \\ &= \sum_{j|d_j \geq i} 1 \\ &= \text{Card}\{j|d_j \geq i\} \end{aligned}$$

giving the first equality. For the second, we write

$$\begin{aligned} d_i^{**} &= \text{Card}\{j|d_j^* \geq i\} \\ &= \text{Card}\{j| \text{Card}\{k|d_k \geq j\} \geq i\} \end{aligned}$$

But $\text{Card}\{k|d_k \geq j\} \geq i$ if and only if $d_i \geq j$. Indeed, if there is an ordered set of indices K of cardinality $\geq i$ such that $k \in K \Rightarrow d_k \geq j$, then its i -th element k is $\geq i$ and $d_i \geq d_k \geq j$ by the decreasing nature of \underline{d} . Conversely, if $d_i \geq j$, then $d_k \geq j$ for $k \leq i$ always by the decreasing nature and thus $\text{Card}\{i|d_i \geq j\} \geq i$. Thus $\text{Card}\{j| \text{Card}\{k|d_k \geq j\} \geq i\} = d_i$. \square

Remark(s) 5.12.2.3. *The usual argument uses Young's tableau giving proofs, more or less convincing, of a graphical nature. It is unnecessary for us to introduce these additional notations.*

5.13 Appendices

5.13.1 Algorithm for moving from equivalence to similarity

We know therefore that if $T\text{Id} - A$ and $T\text{Id} - B$ are equivalent, i.e., if there exist $P(T), Q(T)$ polynomial and invertible matrices such that

$$P(T)(T\text{Id} - A) = (T\text{Id} - B)Q(T)^{-1},$$

then there exists $P \in \text{GL}_n(\mathbf{k})$ such that $B = PAP^{-1}$.

Proposition 5.13.1.1 (Thanks to O. Debarre). *There exists an algorithm for computing such a P .*

Proof. We can perform the divisions by monic (here of degree one) in $\mathcal{R}[T]$ with $\mathcal{R} = M_n(\mathbf{k}[T])$ (5.2.0.1)

$$\begin{aligned} P(T) &= (T\text{Id} - B)P_1(T) + P_0, \\ Q(T)^{-1} &= \tilde{Q}_1(T)(T\text{Id} - A) + \tilde{Q}_0, \end{aligned}$$

with P_0 and \tilde{Q}_0 in $M_n(\mathbf{k})$ (let's stress that \mathcal{R} is not in a commutative ring). We obtain by substituting

$$((T\text{Id} - B)P_1(T) + P_0)(T\text{Id} - A) = (T\text{Id} - B)(\tilde{Q}_1(T)(T\text{Id} - A) + \tilde{Q}_0)$$

or also

$$(T\text{Id} - B)(P_1(T) - \tilde{Q}_1(T))(T\text{Id} - A) = (T\text{Id} - B)\tilde{Q}_0 - P_0(T\text{Id} - A).$$

The left-hand side is therefore of degree at most 1 in T , which is only possible if $P_1(T) = \tilde{Q}_1(T)$. Thus $(T\text{Id} - B)\tilde{Q}_0 = P_0(T\text{Id} - A)$ (argue by contradiction and look at the highest degree term). The equality of the coefficients of T gives $\tilde{Q}_0 = P_0$, that of the constant coefficients gives $B\tilde{Q}_0 = P_0A$. It remains to show that \tilde{Q}_0 is invertible. We perform another division in $\mathcal{R}[T]$

$$Q(T) = Q_1(T)(T\text{Id} - B) + Q_0$$

and we write

$$\begin{aligned} \text{Id} &= Q(T)^{-1}Q(T) \\ &= (\tilde{Q}_1(T)(T\text{Id} - A) + \tilde{Q}_0)Q(T) \\ &= \tilde{Q}_1(T)(T\text{Id} - A)Q(T) + \tilde{Q}_0Q(T) \\ &= \tilde{Q}_1(T)P(T)^{-1}(T\text{Id} - B) + \tilde{Q}_0(Q_1(T)(T\text{Id} - B) + Q_0) \\ &= (\tilde{Q}_1(T)P(T)^{-1} + \tilde{Q}_0Q_1(T))(T\text{Id} - B) + \tilde{Q}_0Q_0. \end{aligned}$$

Again, as \tilde{Q}_0Q_0 is constant, the factor of $T\text{Id} - B$ is zero and $\tilde{Q}_0Q_0 = \text{Id}$, hence the conclusion. \square

5.13.2 Jordan reduction by duality of nilpotents without modules

We are going to give a classical proof of the Jordan reduction theorem 5.12.0.2 by induction on dimension. We'll freely use using standard methods of duality (see chapter (7) below) to deal with the nilpotent matrices. The reader we'll prove the general case and of characteristic spaces (*i.e.* primary components of V_a) by reduction to characteristic spaces (see chapter 8 below) thanks to the kernel lemma 3.7.0.2 to prove the general case see chapter.

We start the induction in dimension 0. So let a be a nilpotent endomorphism on V of dimension $n \geq 1$ and $d = \deg \mu_a \geq 1$ its index of nilpotency. Since $a^{d-1} \neq 0$, we can choose v such that $a^{d-1}(v) \neq 0$. As the vector is non-zero, we further choose $\varphi \in V^*$ such that $\langle \varphi, a^{d-1}(v) \rangle \neq 0$.

By construction, the spaces $W = \mathbf{k}[a].v$ and $W_* = \mathbf{k}[{}^t a].\varphi$ are stable by a and ${}^t a$ respectively generated by $a^i(v), i \leq d-1$ and ${}^t a^i(\varphi), i \leq d-1$ respectively. Their dimension is therefore $\leq d$. We easily verify that these generating families are free, so they are of dimension d . In particular, W is cyclic and the matrix of (the restriction to W of) a in the previous base is the standard Jordan block J_d . As W^* is stable by ${}^t a$, its orthogonal $W_*^\perp \subset V^{**} = V$ is stable by ${}^{tt} a = a$, hence is of dimension $n-d < n$. We can thus apply the induction hypothesis to the restriction of a to W' .

It remains to verify that the sum $W + W_*^\perp$ is direct therefore that $W \cap W_*^\perp = \{0\}$ since the dimensions are complementary. So let $\sum_{i < d} \lambda_i a^i(v)$ be in the intersection. If any of the λ_i is non-zero, choose j the smallest index of non-zero coefficients and apply ${}^t a^{d-1-j} \varphi \in W_*$. We thus have

$$0 = \langle {}^t a^{d-1-j} \varphi, \sum_{i < d} \lambda_i a^i(v) \rangle = \langle \varphi, \sum_{i \geq j} \lambda_i a^{d-1-j+i}(v) \rangle = \lambda_j \langle \varphi, a^{d-1}(v) \rangle \neq 0,$$

a contradiction.

5.13.3 Frobenius decomposition without modules

We are now going to give a proof of the Frobenius theorem 5.9.0.2 by induction on dimension using standard methods of linear algebra (see also 5.9.1). This is in fact just an adaptation of the previous proof 5.13.2 of the Jordan reduction theorem. It will not be algorithmic² due to the use of the kernel lemma for the following key lemma.

Lemma 5.13.3.1. *Every endomorphism in finite dimension admits a cyclic vector.*

Proof. The ideal of $\mathbf{k}[T]$ of polynomial P such that $P(a)(v) = 0$ has a unique monic generator $\mu_{a,v}$ which divides $\mu = \mu_a$. Suppose that μ is the power P^d of an irreducible polynomial. Each minimal $\mu_{a,v}$ is therefore of the form P^{d_v} with $d_v \leq d$. If for every v , $d_v \leq d-1$, we would have $P^{d-1}(a)(v) = 0$, a contradiction with $\mu_a = P^d$.

In the general case, decompose $\mu_a = \prod P_i^{d_i}$ into powers of irreducible factors pairwise coprime. On each kernel $K_i = \text{Ker}(P_i^{d_i}(a))$, the minimal of $a|_{K_i}$ is $P_i^{d_i}$, so thanks to what precedes there exists a cyclic vector v_i for $a|_{K_i}$. It remains to apply the kernel lemma 3.7.0.2 to be convinced that the sum of the v_i is a cyclic vector for a . □

Exercise(s) 5.13.3.2. *Using theorem 5.5.0.2 and the corresponding algorithm, write an algorithm to find a cyclic vector. Implement it with SAGEMath.*

For the existence of the Frobenius decomposition, we will thus adapt the demonstration of the previous section. Suppose therefore that a is an arbitrary endomorphism of V of finite dimension of minimal

²For the interested reader, see http://www.lix.polytechnique.fr/~augot/CRAS_94.pdf for an algorithm. Compare with exercise 5.13.3.2 *infra*.

polynomial μ_a of degree d and choose v a cyclic vector for a . The subspace $W = \mathbf{k}[a].v$ is cyclic, stable of dimension d so that the minimal of $a|_W$ is μ_a . The same is true of its transpose ${}^t a|_W \in \text{End}_{\mathbf{k}}(W^*)$; let $\tilde{\varphi} \in W^*$ a cyclic vector for ${}^t a|_W \in \text{End}_{\mathbf{k}}(W^*)$. As the degree of its minimal is d which is also the dimension of W^* , we have $\mathbf{k}[{}^t a|_W].\tilde{\varphi} = W^*$.

Let $\varphi \in V^*$ be an arbitrary linear extension of $\tilde{\varphi} \in W^* = \text{Hom}_{\mathbf{k}}(W, \mathbf{k})$ to V and set $W_* = \mathbf{k}[{}^t a].\varphi \subset V^*$. As any monogenic subspace, W_* has dimension $\leq d$. But since the restriction of forms to W surjectively maps $W_* = \mathbf{k}[{}^t a].\varphi \subset V^*$ to $\mathbf{k}[{}^t a|_W].\tilde{\varphi} = W^*$ because it maps φ to $\varphi|_W = \tilde{\varphi}$, this restriction is an isomorphism $W_* \simeq W^*$. In particular, φ is cyclic for ${}^t a$.

As above, W and W_*^\perp are stable with complementary dimensions, with W being cyclic. Furthermore, the minimal polynomial of $a|_{W_*^\perp}$ divides μ_a . What remains is to prove that the sum of W and W_*^\perp is direct to conclude by induction. However, if $w \in W$ is orthogonal to W_* , then for all $\psi \in W_*$, the nullity of $\langle \psi, w \rangle$ is just $\langle \psi|_W, w \rangle$. Since the restriction $W_* \rightarrow W^*$ is an isomorphism, it follows that w is orthogonal to every form of its dual, thus is zero.

For uniqueness, suppose, with obvious notations, that we have two Frobenius decompositions

$$V = \bigoplus V_i(P_i) = \bigoplus W_i(Q_i).$$

Let us show by strong induction that $P_i = Q_i$ for every i . We already necessarily have $P_1 = \mu_a = Q_1$. Now assume $i > 1$ and $P_1 = Q_1, \dots, P_{i-1} = Q_{i-1}$.

On one hand, we have

$$P_i.V = \bigoplus_{j < i} P_i.V_j$$

because P_i divides $P_j = \mu_a|_{V_j}|P_i$ if $j \geq i$. On the other hand, we have

$$P_i.V = \bigoplus_j P_i.W_j$$

and

$$\dim P_i.V_j = \dim P_i.W_j \text{ if } j < i$$

because, in suitable bases, the matrices of the restrictions of a to V_j and W_j are the same companion matrices associated with $P_j = Q_j$ if $j < i$, and so it is the same for those of $P_i(a)$. By calculating the dimension of $P_i.V$ in two ways, it follows that $\dim P_i.W_i = 0$ and thus $Q_i|P_i$ since Q_i is the minimal polynomial of a on W_i . By symmetry of roles, we have $P_i = Q_i$.

5.14 Implementations in Sage

We have used SageMath version 9.7, Release Date: 2022-09-19, Using Python 3.10.5.

5.14.1 Elementary Divisor Calculations

```

2     # Function to calculate elementary divisors
3     def divelem(a):
4         m = a.nrows()
5         n = a.ncols()
6         d = []
7         # Matrices of size <= (1,1)
8         if m * n == 0 or a == zero(ZZ, m, n):
9             return d
10        a = zerotage(a)
11        d = d + [a[0, 0]]
12        # Row or column matrices
13        if (m - 1) * (n - 1) == 0:
14            return d
15        # Non-row or column matrices
16        d = d + divelem(a[1:m, 1:n])
17        # By recursion, d is defined and a is equivalent to diag(d)
18        return divdiag(d)
19
20    def divelem_norm(a):
21        return [p / p.leading_coefficient() for p in divelem(a) if p != 0]
22
23    d = divelem_norm(a)
24    tmps2 = time.time()
25    print((tmps2 - tmps1) * 1000, 'ms')
26

```

Remark(s) 5.14.1.1. *This Polynomial Gauss Pivot program quickly becomes deficient when, for example, \mathbf{Q} is replaced by \mathbf{R} . The reason is the structural numerical instability of the Gauss pivot. When pivoting on scalar matrices, we partially compensate for this by always choosing the largest pivot in absolute value to try to prevent the coefficients from exploding and exceeding the machine's capabilities, but this becomes impossible with polynomial pivoting. It's an interesting subject for reflection to see how we might overcome this difficulty. We chose to reprogram a number of native Sage functions, such as the elementary operations, to clearly illustrate the algorithm.*

5.14.2 Jordan-Chevalley Decomposition

```

1
2     # Thanks to Antoine Castellani for this SAGE code calculating Dunford's reduction
3     l = 5
4     d = 4
5     k.<u> = GF(l^d)

```

```

6 # To switch fields, uncomment and comment the above line for field k = QQ
7 # Adaptation is necessary if the field is not perfect: a good exercise
8
9 R.<x> = PolynomialRing(k) # Replace 'k' with 'QQ' or 'k.<u> = GF(p^n)' as needed
10
11 # Version to avoid using the factor() command in SAGE
12 def remove_square_factors(P):
13     p = k.characteristic()
14     if p == 0:
15         return P / gcd(P, diff(P))
16     else:
17         u = gcd(P, diff(P))
18         v = P / u # Returns terms without square when power is not divided by p
19         if v == P:
20             return v
21         else:
22             w = u / gcd(u, v^(P.degree())) # Returns the other terms, i.e., those whose power is
23             divided by p. We can therefore take the p-th root
24             root_w = (w.numerator()).nth_root(p) # Returns the p-th root of w and iterate
25             return v * remove_square_factors(root_w)
26
27 def Hensel(P, x_0, n):
28     # P polynomial
29     # x_0 from lemma
30     # n integer
31     solutions = [x_0]
32     N = valuation(n, 2) + 2
33     for j in range(N):
34         i = P(solutions[j])
35         r = (diff(P)(solutions[j])).inverse()
36         solutions.append(solutions[j] - r * i)
37     return solutions[N]
38
39 def Jordan(a):
40     # a square matrix
41     n = a.nrows()
42     pi = remove_square_factors(minpoly(a))
43     Delta = Hensel(pi, a, n)
44     return "D =", Delta, "N =", a - Delta, "test =", a * Delta - Delta * a
45
46 # Testing
47 # a = matrix(k, [[1, k.random_element(), 1], [0, 1, 1], [0, 0, 2]])
48 # Jordan(a)

```

5.15 Additional Exercises

Exercise(s) 5.15.0.1. Prove that $a \in \text{End}_k(V)$ is triangulable if and only if χ_a is split. Demonstrate that a family of commuting and triangulable endomorphisms admits a common triangulation base (inspired by the proof of 5.7.0.4).

Exercise(s) 5.15.0.2. Consider G , a commutative subgroup of $\text{GL}_n(\mathbf{R})$ where every element is squared to Id .

1. Prove that G is finite with cardinality $\leq n$ (refer to 5.7.0.2 and 5.7.0.4).
2. Prove that if $\text{GL}_n(\mathbf{R})$ and $\text{GL}_m(\mathbf{R})$ are isomorphic, then $n = m$.
3. Can you generalize this to other fields?
4. What happens if G is no longer assumed to be commutative?

Exercise(s) 5.15.0.3. Using the notations and results from 3.8.0.13.

1. Let $P \in \mathbf{C}[T]$. Show that the roots of P are simple if and only if the discriminant of P , defined by $\text{Res}(P, P')$, is non-zero.
2. Considering the discriminant of the characteristic polynomial of a complex square matrix, prove that the set of $M_n(\mathbf{C})$ matrices with distinct eigenvalues is dense and that its complement is of zero Lebesgue measure.

Exercise(s) 5.15.0.4. In this exercise, we assume that $\mathbf{k} = \mathbf{R}$ and accept that every irreducible polynomial in $\mathbf{R}[T]$ has a degree of at most 2. Let a be an endomorphism of V .

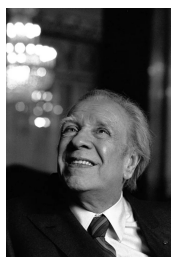
1. Using the Chinese remainder theorem, demonstrate that V_a can be written as a direct sum of $\mathbf{R}[T]$ -modules of the form $\mathbf{R}[T]/(T - \alpha)^m$ or $\mathbf{R}[T]/(T^2 - bT - c)^n$ with α, b, c being real numbers such that $b^2 + 4c < 0$ and m, n being strictly positive integers.
2. Show that $(T - \alpha)^{m-1}, \dots, (T - \alpha), 1$ is a basis of the \mathbf{R} -vector space $\mathbf{R}[T]/(T - \alpha)^m$ and that $T(T^2 - bT - c)^{n-1}, (T^2 - bT - c)^{n-1}, \dots, T(T^2 - bT - c), T^2 - bT - c, T, 1$ is a basis of the \mathbf{R} -vector space $\mathbf{R}[T]/(T^2 - bT - c)^n$.
3. In both cases, write the matrix of multiplication by T in the basis given in the previous question. Such matrices are called real Jordan blocks.
4. We assume that the matrix of a is

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

Give the decomposition of this matrix into real Jordan blocks.

Chapter 6

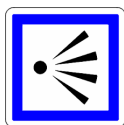
Semisimplicity in $M_n(\mathbf{k})$



Jorge Luis Borges

Simplicity// It opens, the gate to the garden/ with the docility of a page/ that frequent devotion questions and inside, my gaze/ has no need to fix on objects/ that already exist, exact, in memory.// I know the customs and souls/ and that dialect of allusions/ that every human gathering goes weaving./ I've no need to speak/ nor claim false privilege;/ they know me well who surround me here,/ know well my afflictions and weakness.// This is to reach the highest thing,/ that Heaven perhaps will grant us:/ not admiration or victory/ but simply to be accepted/ as part of an undeniable Reality,/ like stones and trees.

6.1 Perspective



We show algorithmically how the study of similarity classes reduces to the classes of nilpotent matrices in $M_n(\mathbf{k})$ (5.12.0.2) and to those of diagonalizable matrices over an algebraically closed field Ω containing \mathbf{k} , the (absolutely) semisimple matrices.

The Jordan-Chevalley decomposition¹ (over fields of characteristic zero and more generally perfect fields) is the cornerstone of the theory of algebraic groups and Lie algebras. In the complex case, it will also be specified how this decomposition is generally discontinuous, continuous if the characteristic polynomial is fixed.

In this chapter, unless expressly stated otherwise,
 \mathbf{k} denotes a perfect field (6.2.3.1) and Ω an algebraically closed field that contains it.

¹Interestingly, the term Duford decomposition is found for Jordan-Chevalley decomposition in French literature.

6.2 Semisimplicity

Semisimplicity is the right generalization of diagonalizability in the case of perfect fields as we will see. Let's start with some formal generalities.



Ryoan-ji, Kyoto

6.2.1 General Semisimple Modules

In this paragraph, R denotes an arbitrary commutative unit ring.

Definition 6.2.1.1. *A R -module is said to be*

- *semisimple if every submodule has a complement;*
- *simple if it is non-zero and has no non-trivial submodules.*

An endomorphism $a \in \text{End}_{\mathbf{k}}(V)$ is said to be semisimple if the $\mathbf{k}[T]$ -module V_a is.

Exercise(s) 6.2.1.2. *Demonstrate the following points.*

1. *A vector space is semisimple, and is simple if and only if it is of dimension 1.*
2. *There exists $a \in \text{End}_{\mathbf{k}}(V)$ such that V_a is not semisimple².*
3. *The \mathbf{Z} -module (i.e., abelian group) $\mathbf{Z}/4^2\mathbf{Z}$ is not semisimple (cf. 6.2.1.3 infra).*
4. *A principal ring that is not a field is never semisimple as a module over itself.*

The following observations are elementary but very useful.

²Consider a nilpotent in dimension 2 for example (6.2.4.1).

Proposition 6.2.1.3. *Let N be a submodule of a semisimple module M .*

- M is isomorphic to $N \oplus M/N$.
- Every submodule is isomorphic to a quotient module and every quotient module is isomorphic to a submodule.
- Both N and M/N are semisimple.

Proof. Let S be a complement of N in M . The canonical surjection $f : M \rightarrow M/N$ defines by restriction an isomorphism $S \xrightarrow{\sim} M/N$ hence the first point. But then, the first projection $M = N \oplus M/N$ is surjective and identifies N to the quotient $\text{Coker}(p)$. Similarly for the quotients with the inclusion $M/N \hookrightarrow M = N \oplus M/N$, hence the second point. For the last point, if M' is a submodule of M/N , a complement S is chosen for the submodule $f^{-1}(M')$ and it is verified that $f(S)$ is a complement of M' in M/N so that M/N is semisimple. But as N is identified to a quotient of M , the same is true for N . ■ □

Exercise(s) 6.2.1.4. *Show that M is semisimple if and only if every short exact sequence is split (cf. 3.8.0.3).*

6.2.2 Semisimple Modules over Principal R

If m is a torsion element of some R -module, its annihilator $\text{Ann}_R(m)$ has a non-zero generator, well defined up to invertible elements: its minimal μ_m (see the proof 5.13.3.1 for a special case of this notion).

Proposition 6.2.2.1. *Suppose M is a module over principal ring R which is not a field.*

1. *There exists an irreducible element p of R .*
2. *$R/(p^2)$ and therefore R is not semisimple.*
3. *M is semisimple if and only if M is torsion and if the minimal of every element is square-free.*

Proof. As R is not a field, R has a non-null non-invertible element whose any of its irreducible factors meets (1).

If $R/(p^2)$ were semisimple, the exact sequence $0 \rightarrow R/(p) \xrightarrow{p} R/(p^2) \rightarrow R/(p) \rightarrow 0$ would be split since $pR/(p)$ would have a complement in $R/(p^2)$ and thus $R/(p^2) \simeq R/(p) \oplus R/(p)$ (3.8.0.3). But this would imply that p kills $R/(p^2)$, which it does not hence (2). Moving on to (3).

\Rightarrow If $m \in M$ has a trivial annihilator, $R \xrightarrow{m} R$ is injective so that R is a submodule of M and therefore should be semisimple, which it is not (6.2.1.2)). Hence, every element is torsion. Let then m whose

minimal μ_m is divisible by p^2 with p an irreducible so that $R/(\mu_m)$ is a submodule of M . Assume absurdly M semisimple. Then, $R/(\mu_m)$ is also a quotient of M (6.2.1.3) and therefore the same for $R/(p^2)$ (as a quotient of $R/(\mu_m)$ therefore of M) which then would be semisimple, which it is not according to (2).

\Leftarrow Assume the minimal of every element is square-free and let N submodule of M . Then, for every p irreducible, we have $M[p] = \bigcup_{n \geq 1} \text{Ann}_M(p^n) = \text{Ann}_M(p)$ (3.7). As M is torsion, the Chinese lemma (3.7.0.1) ensures $M = \bigoplus_p \text{Ann}_M(p)$ for p describing the irreducibles up to a unit (**exercise**). But the structure of R -module of $\text{Ann}_M(p)$ is factored through $R \rightarrow R/pR = \mathbf{k}(p)$ which is a field because R is principal : $\text{Ann}_M(p)$ is a $\mathbf{k}(p)$ -vector space. Likewise, we have $N = \bigoplus_p \text{Ann}_N(p)$. Let then for every p a complement S_p of the $\mathbf{k}(p)$ -vector subspace $\text{Ann}_N(p)$ of $\text{Ann}_M(p)$. The R -module $\bigoplus S_p$ is a complement of N . \square

6.2.3 «Reminder» on Perfect Fields

On a general field K , it may happen that a polynomial without squared factors has multiple roots in a larger field. For example, this is the case with $T^2 + t$ in $K = \mathbf{F}_2(t)$, the field of fractions of the polynomial ring $\mathbf{F}_2[t]$ [t is assumed to be transcendental over \mathbf{F}_2]. This does not occur in perfect fields. Let p be a prime number and R a ring such that $pR = \{0\}$. The well-known divisibility $p \mid \binom{p}{n}$ for $1 \leq n \leq p-1$ and the binomial formula ensure that the application $F : r \mapsto r^p$ is a ring morphism called the Frobenius morphism. If R is a field, it is additionally injective as any morphism of fields.

Definition 6.2.3.1. *A field of characteristic p is said to be perfect if $p = 0$ or if every element admits a p -th root, i.e. if its Frobenius morphism is an isomorphism.*

Thus, every finite field is perfect since an injection between finite sets is bijective. Therefore, we must prove the following statement.

Lemma 6.2.3.2. *Let \mathbf{k} be a perfect field and $P \in \mathbf{k}[T]$. Then, P is square-free if and only if $\text{GCD}(P, P') = 1$. In particular, if \mathbf{k} is perfect and P irreducible, then $\text{GCD}(P, P') = 1$.*

Proof. The direction \Leftarrow immediately follows from Bézout's identity. Let's consider the direct direction. Suppose P is without squared factors and write $P = \prod P_i$ with P_i irreducible. If $\text{GCD}(P, P') \neq 1$, one of the P_i divides $P' = \sum_i P'_i \prod_{j \neq i} P_j$ and thus $P_i | P'_i$. By comparing degrees, we have $P'_i = 0$. This implies that the characteristic of \mathbf{k} is a prime number p and that all coefficients of P_i of indices not multiples of p are zero: $P_i = \sum_n a_{np} T^{np}$. But in this case, we have $P_i = (\sum_n a_{np}^{1/p} T^n)^p$ because the Frobenius of $\mathbf{k}[T]$ is a ring morphism. A contradiction with the irreducibility of P_i \square

Exercise(s) 6.2.3.3. Let V be a \mathbf{k} -vector space of finite dimension and φ an automorphism of \mathbf{k} . Denote $[\varphi] \otimes V$ as the vector space with underlying group V and external law $\lambda \cdot_{[\varphi]} v = \varphi(\lambda)v$. Show $\dim(V) = \dim([\varphi] \otimes V)$. Deduce that any field of finite dimension over a perfect field is still perfect.

6.2.4 Criterion for Semisimplicity of V_a

The calculation of GCD of polynomials does not depend on the base field (for example because Euclid's algorithm does not depend on it) nor does that of the minimal of the matrix. According to 5.7.0.1, the condition $\text{GCD}(\mu_a, \mu'_a) = 1$ therefore means that the matrix of a is diagonalizable in $M_n(\Omega)$. In the case of V_a , this can be summarized as follows.

Proposition 6.2.4.1. Let $a \in \text{End}_{\mathbf{k}}(V)$ (with \mathbf{k} perfect) whose matrix $A \in M_n(\mathbf{k})$ is in a given base. The following propositions are equivalent:

1. The minimal μ_a of a is without squared factors.
2. $\text{GCD}(\mu_a, \mu'_a) = 1$.
3. A is diagonalizable in $M_n(\Omega)$.
4. V_a is semisimple.
5. Every submodule of V_a is semisimple.

If these equivalent conditions are met, a is said to be semisimple (ditto for a matrix of a).

Families of commuting diagonalizable endomorphisms being simultaneously diagonalizable (5.7.0.4), we deduce

Corollary 6.2.4.2. Let $a, b \in \text{End}_{\mathbf{k}}(V)$ with a, b semisimple that commute (\mathbf{k} perfect) and $P \in \mathbf{k}[X, Y]$. Then, $P(a, b)$ is semisimple.

This corollary is false in the imperfect case.

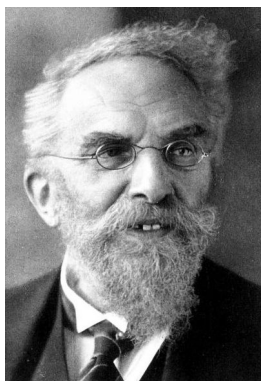
Remark(s) 6.2.4.3. When the base field \mathbf{K} is not perfect, there are semisimple matrices over \mathbf{K} which, considered in a superfield, are no longer so. With the notations of 6.2.3, this is the case with $A = \begin{pmatrix} 0 & t \\ 1 & 0 \end{pmatrix}$ over $\mathbf{K} = \mathbf{F}_2(t)$ because $\chi_A(T) = T^2 + t$ is irreducible over \mathbf{K} but not over $\mathbf{K}(t^{1/2}) = \mathbf{K}[\tau]/(\tau^2 - t)$ and a fortiori over $\Omega \supset \mathbf{K}$. Moreover, $A + t^{1/2} \text{Id}$ is even nilpotent! The correct notion in the non-perfect case is that of absolute simplicity defined by the condition $\text{GCD}(\mu_a, \mu'_a) = 1$, stronger than semisimplicity.

Exercise(s) 6.2.4.4. Let p be prime, K the field of fractions of $\mathbf{F}_p[T]$ and $V = K[X, Y]/(X^p - T, Y^p - T)$. Show that V is of finite dimension over K and that the K -endomorphisms of V multiplying by X and Y respectively are semisimple, commute but their difference is nilpotent (this is exercise 14 chapter VII.5 [8] rewritten without tensor product). Prove without recourse to simultaneous diagonalization that the sum of two absolutely semisimple matrices is absolutely semisimple using the exercise 5.15.0.3.

6.3 Jordan-Chevalley Decomposition

Let's begin with a very important result, although easily demonstrated, which allows the construction of polynomial roots step-by-step (adaptation of Newton's method).

6.3.1 Hensel's Lemma and Existence



Kurt Hensel

Kurt Hensel



© Gotlib

Isaac Newton

Lemma 6.3.1.1 (Hensel-Newton). Let I be a nilpotent ideal ($I^N = 0$) of an arbitrary ring R and $P \in R[T]$. Assume there exists $x_0 \in R$ such that $P(x_0) \equiv 0 \pmod I$ and $P'(x_0) \pmod I$ is invertible. Then, there exists $x \in R$ such that $x \equiv x_0 \pmod I$ and $P(x) = 0$.

Proof. First, observe that if $a \pmod I$ is invertible, then a is invertible in a . Indeed, if $b \pmod I$ is its inverse, $ab = 1 - i$ with $i \in I$. Formally expanding $1/(1 - i)$ into a series, we deduce that $1 - i$ is invertible with the inverse $\sum_{k < N} i^k$ since $i^k = 0$ for $k \geq N$ and thus $b/(1 - i)$ is the inverse of a .

We will compute (algorithmically) an approximate root

$$x_k \pmod{I^{2^k}} | P(x_k) \equiv 0 \pmod{I^{2^k}} \text{ and } x_k \equiv x_0 \pmod{I}$$

by successive approximations. Proceed by induction on $k \geq 0$ (with tautological initialization). Assuming the property holds at rank k , we then seek x_{k+1} in the form $x_{k+1} + \varepsilon$, $\varepsilon \in I^{2^k}$ so that x_{k+1} is indeed an approximation of $x_k \pmod{I^{2^k}}$.

The entire Taylor formula gives

$$P(x_{k+1}) = P(x_k) + \varepsilon P'(x_k) + \varepsilon^2 Q(x_k, \varepsilon)$$

with $Q[T, Y] \in R[T, Y]$ (check this!). Since $x_k \equiv x_0 \pmod{I}$, we have $P'(x_k) \equiv P'(x_0) \pmod{I}$ and therefore $P'(x_k) \pmod{I^{2^k}}$ is invertible. We then set $\varepsilon = -P(x_k)/P'(x_k)$. $\varepsilon \in I^{2^k}$ is guaranteed by the construction of x_k . As $\varepsilon^2 \in I^{2^{k+1}}$, this choice is suitable. To conclude, we choose k such that $2^k \geq N + 1$ and set $x = x_k$: the algorithm converges exponentially! \square

Corollary 6.3.1.2 (Existence). *Let $a \in \text{End}_{\mathbf{k}}(V)$ (with \mathbf{k} a perfect field). There exist $d, \nu \in \mathbf{k}[a] \subset \text{End}_{\mathbf{k}}[a]$ such that $a = d + \nu$ and d semisimple, ν nilpotent. In particular, d and ν commute.*

Proof. Let $\pi \in \mathbf{k}[T]$ be the product of the irreducible factors of the minimal μ_a of a . As it is without squared factors, it is coprime with its derivative. Choose $\alpha, \beta \in \mathbf{k}[T]$ such that $\alpha\pi + \beta\pi' = 1$.

Let I be the ideal $\pi(a)\mathbf{k}[a]$ of $\mathbf{k}[a]$. We have $\mu_a | \pi^n$ and therefore $\pi^n(a) = 0$ so that $I^n = 0$. Furthermore, we have $\beta(a)\pi'(a) = 1 \pmod{I}$ and thus $\pi'(a) \pmod{I}$ is invertible. By setting $x_0 = a \in \mathbf{k}[a]$, we deduce the existence of $x \in \mathbf{k}[a]$ such that $x = a \pmod{I}$ and $\pi(x) = 0 \pmod{I^n} = (0)$. We then set $d = x$ and $\nu = a - P(a)$. As $\pi(d) = 0$, d is absolutely semisimple. Since $\nu = a - P(a) \in I$ and $I^n = 0$, ν is nilpotent. \blacksquare

Remark(s) 6.3.1.3. *This is essentially Chevalley's proof. Beyond its algorithmic character (very fast), it is important because it allows the definition of semisimple and nilpotent parts within the context of Lie algebras and algebraic groups (on a perfect field), see for example the excellent [7].*

6.3.2 Uniqueness

Theorem 6.3.2.1 (Jordan-Chevalley). *We still assume \mathbf{k} is a perfect field.*

1. *Let $a \in \text{End}_{\mathbf{k}}(V)$. There exists a unique pair (d, ν) with d semisimple, ν nilpotent, d and ν commuting with $a = d + \nu$.*

2. Let $\chi \in \mathbf{k}[T]$ be a monic polynomial of degree n . There exists $P \in \mathbf{k}[X]$ (depending only on χ) such that if $\chi_a = \chi$, then $d = P(a)$ and in particular $d, \nu \in R = \mathbf{k}[a] \subset \text{End}_{\mathbf{k}}[a]$.

Proof. Only uniqueness requires an argument given the above. Suppose d, ν as in the theorem and a pair $d', \nu' \in \mathbf{k}[a]$ as in Corollary 6.3.1.2. Since d, ν commute with each other, they commute with $d + \nu = a$. They therefore also commute with d', ν' because these are polynomials in a . But $d + \nu = d' + \nu'$ i.e., $d - d' = \nu' - \nu$. However, $\nu' - \nu$ is nilpotent (as a sum of commuting nilpotents) and $d - d'$ semi-simple (as a sum of commuting semi-simples, 6.2.4.2); an endomorphism that is both semi-simple and nilpotent being zero since its minimal polynomial has no squared factors and divides T^n , we indeed have $d = d'$ and $\nu = \nu'$. □

A diagonalizable endomorphism a thus decomposes into $d = a$ and $\nu = 0$. Thus $a = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$



decomposes into $a + 0$ and not into $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ as one might be tempted to write.

Furthermore, the assumption of \mathbf{k} being a perfect field cannot be relaxed: the matrix $\begin{pmatrix} 0 & t \\ 1 & 0 \end{pmatrix}$ from 6.2.4.3 does not have a Jordan-Chevalley decomposition. If one wants such a decomposition in the imperfect case, one must restrict to endomorphisms with *separable* characteristic polynomials and replace semi-simple with absolutely semi-simple. The proof is then identical.

6.3.3 Similarity class of the components

We retain the previous notation. $a = d + \nu$. The invariant factors of the semi-simple part d are entirely determined by χ_a since two diagonalizable endomorphisms with the same characteristic polynomials are similar over Ω and the invariants do not depend on the base field (cf. 6.3.4.1). Similarly, the similarity invariants of a determine the nilpotent type \underline{d}_a of ν . One way to see this is to observe that the nilpotent parts of two similar matrices have similar nilpotent parts by uniqueness of the Jordan-Chevalley decomposition.

6.3.4 Appendix: What about the algorithmic nature of the decomposition?

On re-examining the proofs *supra*, one easily convinces oneself that finding d and ν is algorithmic once one knows the product π of the distinct irreducible factors of P_n . SageMath does this very well thanks to the *factor* command. But what if this command did not exist? In characteristic zero, one is easily convinced of the formula

$$\pi = P_n / \text{GCD}(P_n, P'_n)$$

so that the process is algorithmic thanks to Euclid's GCD algorithm in $\mathbf{k}[T]$. In characteristic $p > 0$, it is more complicated because there are polynomials with a null derivative: the polynomials in T^p . The

following exercise provides an «algorithm» to find π for a perfect field of characteristic $p > 0$. The quotes are justified by the assumption that the inverse of the Frobenius³ $F : x \mapsto x^p$ of \mathbf{k} is known algorithmically.

Exercise(s) 6.3.4.1. Let \mathbf{k} be a field and $\chi = \prod \pi_i^{n_i}$ the decomposition into unitary irreducible factors of P a unitary polynomial of degree n . We denote $\chi_{\text{red}} = \prod \pi_i$. In the first four questions, \mathbf{k} is assumed to be a perfect field of characteristic $p > 0$ and I the set of indices i such that n_i is coprime with p .

1. Show that $\chi / \text{GCD}(\chi, \chi') = \prod_{i \in I} \pi_i$.
2. Show that $\prod_{i \notin I} \pi_i$ is a p -th power in $\mathbf{k}[T]$.
3. Write an algorithm computing $\prod_{i \in I} \pi_i$ and $\prod_{j \notin I} \pi_j^{n_j/p}$.
4. Deduce an algorithm computing χ_{red} .
5. What is χ_{red} in characteristic zero?
6. Program the algorithm on \mathbf{F}_p ? On \mathbf{F}_{p^n} ? On a general perfect field?
7. How to generalize on a non-perfect field?
8. Always for \mathbf{k} a general field, consider the sequence of polynomials $\underline{\chi}_{\text{red}} = (\chi_i)_{1 \leq i \leq n}$ defined by $\chi_1 = \chi_{\text{red}}$, $\chi_{i+1} = (\chi / (\prod_{j \leq i} \chi_j))_{\text{red}}$. Show that $\underline{\chi}_{\text{red}}$ is the sequence of invariant factors of the semisimple endomorphisms with characteristic polynomial χ .
9. Assuming again \mathbf{k} perfect and let D, N be the Jordan-Chevalley decomposition of $M \in M_n(\mathbf{k})$. What are the similarity invariants of D based on the invariants \underline{P} of M [Use the previous question]? Can you similarly describe the invariants of N based on P_i [Place yourself in \bar{k} and study the application $P_i \mapsto P_i / P_{i, \text{red}}$ and its iterates]? Program the obtained algorithm for example on \mathbf{F}_p .

Regarding Hensel's lemma, the very writing of the proof is an algorithm that lives in $\mathbf{k}[a] \subset M_d(\mathbf{k})$ where $d = \dim(V)$. It involves calculating the inverse of $P'(x_n)$ as long as $2^n < d$. This is a small number of times, but if the matrices are large, the calculation is heavy. One way to lighten it is to consider the algebra isomorphism $k[T]/\mu_a \xrightarrow{\sim} k[a]$ that sends T to a (exercise) and to work within this quotient, which is less computationally demanding.

Despite this, these algorithms are very unstable. For two reasons. The first is that the Gaussian pivot is a numerically unstable algorithm. And working with polynomial coefficients does not help. The second is more serious. As will be seen below, the similarity invariants do not vary continuously with the coefficients of the matrix (see, for example, the theorem 9.2.0.2). Therefore, approximating the values of the coefficients becomes perilous. When the matrices have rational coefficients, or are in finite fields, one can, with great care, control the height of the coefficients and thus work with true equalities. Even though these algorithms tend to explode the sizes of the integers involved... In short, a real subject for reflection, one of the motivations that led us to include the topological study of similarity classes in chapter 9.

³Which is the case, for example, for finite fields.

6.4 Additional Exercises

Exercise(s) 6.4.0.1. Let M be a complex square matrix of size $n > 1$. We denote by M_{nil} the nilpotent component of its Jordan-Chevalley decomposition. The goal is to give some properties of M_{nil} . Recall that the exponential of M is defined by the absolutely convergent series (for any norm on $M_n(\mathbf{C})$):

$$\exp(M) = \sum_{k=0}^{\infty} \frac{M^k}{k!}$$

and that the exponential of the sum of two commuting matrices is the product of their exponentials.

1. Compute $\exp(M)_{nil}$ in terms of M_{nil} and M .
2. Show that $\exp(M)_{nil} = 0$ if and only if $M_{nil} = 0$. What can be deduced from this?
3. Show that the set of diagonalizable complex matrices is dense in $M_n(\mathbf{C})$.
4. Show that the map $M \mapsto M_{nil}$ is not continuous on $M_n(\mathbf{C})$.
5. What is the set of points of continuity of the map $M \mapsto M_{nil}$ (Difficult)?

Exercise(s) 6.4.0.2. Recall that the exponential of a complex square matrix of M is defined by the absolutely convergent series (for any norm on $M_n(\mathbf{C})$):

$$\exp(M) = \sum_{k=0}^{\infty} \frac{M^k}{k!}$$

and that the exponential of the sum of two commuting matrices is the product of their exponentials.

1. If $M \in M_n(\mathbf{R})$, prove that $\det(M) \geq 0$.
2. Show that $\exp(M_n(\mathbf{R}))$ is the set of square of real matrices.
3. If $n > 1$, show that there exists real matrices of size n with positive determinant but who are not square of any real matrix.

Chapter 7

Reminder on Duality in Finite Dimension



René Magritte

7.1 Basic notions

As always, V denotes in this chapter a finite dimensional¹ \mathbf{k} -vector space and $V^* = \text{Hom}(V, \mathbf{k})$ denotes its dual; the vector space of linear applications from V to \mathbf{k} , *i.e.* linear forms of V .

If $\varphi \in V^*$, $v \in V$, we note $\langle \varphi, v \rangle = \varphi(v)$ the duality bracket² $V^* \times V \rightarrow \mathbf{k}$.

A hyperplane is the kernel of a non-zero linear form φ . Conversely, any hyperplane H determines φ up to multiplication by a non-zero scalar: choosing any $v \notin H$ defines a direct sum decomposition $H \oplus \mathbf{k}v = V$ and φ is unambiguously defined by any (nonzero) value of v .

We recall that any any free family of V can be completed in a basis of V . In particular, any proper subspace of V is contained in some hyperplane and in fact is precisely the intersection of hyperplanes that contain it (i).

¹Unless otherwise stated.

²Be careful, the dual acts to the right on vectors, cf. [9].

Proposition 7.1.0.1. *Let V be a n -dimensional vector space and let V_i finitely many proper sub-vector spaces. If \mathbf{k} is infinite or if the number of subspaces is ≤ 2 , then $\cup V_i \neq V$.*

Proof. By the above remark, we can assume that all the V_i 's are hyperplanes $\text{Ker}(\varphi_i)$. Choosing a (finite) basis of V , these linear forms φ_i are nothing but (homogeneous) degree one polynomial in the coordinates. By assumption $\prod \varphi_i$ is zero on \mathbf{k}^n and therefore the polynomial $\prod \varphi_i(X_1, \dots, X_n)$ is zero in $\mathbf{k}[X_1, \dots, X_n]$ because \mathbf{k} is infinite. But a polynomial ring is an integral domain, showing that one the φ_i is zero, a contradiction. If \mathbf{k} is a finite field (of characteristic $p \geq 2$), the cardinality of V is p^n . The union of two hyperplanes has cardinality at worst $2p^{n-1} - 1 \leq p^n - 1$ (because 0 belongs to both hyperplanes) and the proposition follows. \square

We recall that if $\mathcal{B} = (e_i)$ is a (finite) basis of V , we define the dual basis $\mathcal{B}^* = (e_i^*)$ of V^* by the formula $\langle e_i^*, e_j \rangle = \delta_{i,j}$. In other words, e_i^* is the i -th coordinate function and we have $v = \sum_j \langle v, e_j^* \rangle e_j$. In particular, $\dim(V^*) = \dim(V)$.

If $V = \mathbf{k}^n = M_{n,1}(\mathbf{k})$ (column vectors), we have $M_{1,n}(\mathbf{k}) = \mathbf{k}^n = V^*$ (row vectors) and the duality bracket is $\langle L, C \rangle = L^t C$ where $L \in V^*$ is a row and $C \in V$ a column. If $\mathcal{B} = (e_i = [\delta_{i,j}]_{1 \leq j \leq n})$ is the canonical basis ($E_{i,1} = e_i$) of $\mathbf{k}^n = M_{n,1}(\mathbf{k}) = V$, its dual basis \mathcal{B}^* is formed from the rows $e_i^* = {}^t e_i$, which is the canonical basis ($E_{1,i} = e_i^*$) of $M_{1,n}(\mathbf{k}) = \mathbf{k}^n = V^*$.

If \mathcal{B} is a basis of an infinite dimensional vector space, the family \mathcal{B}^* is still free but is never a basis. For instance, the linear form φ defined by $\langle \varphi, e_i \rangle$ for all i is certainly not in the span of \mathcal{B}^* . Even as a set, $\text{Card}(V^*) > \text{Card}(V)$ (**exercice**). In fact, in the infinite dimensional case, the algebraic dual is not the good notion. As the reader who has notion in functional analysis knows, the good notion is the appropriate topological dual of topological vector spaces.



If W is a subspace of V (or even a subset), we recall that its orthogonal is defined by

$$W^\perp = \{\varphi \in V^* \mid \langle \varphi, w \rangle = 0 \text{ for all } w \in W\} \subset V^*.$$

If now W_* is a subspace of V^* (or even a subset) its polar in V is defined by

$$W_*^\circ = \{v \in V \mid \langle \varphi, v \rangle = 0 \text{ for all } \varphi \in W_*\} \subset V.$$

Example(s) 7.1.0.2. *An important example comes from differential geometry. If f is a regular function on an open Ω of \mathbf{R}^n , its differential at $\omega \in \Omega$ is a linear form on $T_\omega \Omega = \mathbf{R}^n$: the differential $df(\omega)$. In the canonical basis $(\frac{d}{dx_i}(\omega))_i$ of $T_x \Omega$, this form is the Jacobian $J(\omega) = (\frac{df}{dx_j}(\omega))_j$ thus seen as a row matrix. The kernel of $df(\omega)$ is none other than the tangent hyperplane at ω to the hypersurface defined*

by the equation $f = 0$ as long as the differential is non-null at that point. The generalization to several functions is contained in the notion of higher-dimensional submanifolds.

7.2 Motivation

Two useful ways compete to define a vector subspace W of $V = k^n$.

1. Via generators $v_i \in V$: $W = \text{Vect}\{v_i\}$.
2. Via equations $eq_i \in V^*$: $W = \{v \mid \langle eq_i, v \rangle = 0\}$ with

$$\left\langle eq_i, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle = \sum_j a_{i,j} x_j = (a_{i,1}, \dots, a_{i,n}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

The duality first focus on the second point of view, thus on the dual V^* and the set of all possible equations of W : the orthogonal $W^\perp = \{\varphi \in V^* \mid \varphi(W) \equiv 0\}$ and then to the link with the first point of view.

7.3 Formal Biorthogonality

Whether V is of finite dimension or not, any subspace W is tautologically contained in the space defined by the set of its equations

$$W \subset (W^\perp)^\circ \subset \{v \mid \langle \varphi, v \rangle = 0 \text{ for all } \varphi \in W^\perp\}.$$

In general, this inclusion is formal in the sense that it is always an equality, without any further assumption about the dimensionality of V .

$$(i) \quad W = (W^\perp)^\circ = \{v \mid \langle \varphi, v \rangle = 0 \text{ for all } \varphi \in W^\perp\}.$$

Indeed, if $v \notin W$, one can choose a complement S of $W \oplus kv$ in W and define for example $\varphi \in W^\perp$ by the conditions $\langle \varphi, W \rangle = \langle \varphi, S \rangle = \{0\}$ and $\langle \varphi, v \rangle = 1$ which implies $v \notin (W^\perp)^\circ$ proving the reverse inclusion.

7.4 Ante-dual Basis: Biduality

Henceforth, in this chapter, V is finite-dimensional.

Proposition 7.4.0.1. *Let V be of dimension $n < \infty$. Then*

1. *The evaluation linear application*

$$ev : \begin{cases} V & \rightarrow & V^{**} \\ v & \mapsto & (\varphi \mapsto (\langle \varphi, v \rangle)) \end{cases}$$

is an isomorphism.

2. *For any basis \mathcal{B}_* of V^* , there exists a unique basis \mathcal{B} of V called ante-dual whose dual is \mathcal{B}_* , i.e. such that $\mathcal{B}^* = \mathcal{B}_*$.*

Proof. For (1), note that ev is injective between spaces of the same finite dimension.

For (2), note that $\mathcal{B} = ev^{-1}((\mathcal{B}_*)^*)$ is the unique solution to the problem posed. □

7.5 Orthogonal and Polar in Finite Dimension

Proposition 7.5.0.1. *Let W, W_* be two subspaces of V, V^* respectively. We have*

1. $\dim(W) + \dim(W^\perp) = n$.

2. $\dim(W_*) + \dim(W_*^\circ) = n$.

3. $W_* = (W_*^\circ)^\perp$.

4. $W = (W^\perp)^\circ$.

5. $ev(W_*^\circ) = W_*^\perp$.

6. $ev(W) = W^{\perp\perp}$.

Proof. For (1), choose a basis $(e_i, 1 \leq i \leq d)$ of W and complete it to a basis $\mathcal{B} = (e_i, 1 \leq i \leq n)$ of V . If $\mathcal{B}^* = (e_i^*)$ is the dual basis, then by construction $W^\perp = \text{Vect}(e_i, i > d)$.

For (2), choose a basis $(\varphi_i, 1 \leq i \leq d)$ of W_* and complete it to a basis $\mathcal{B}_* = (\varphi_i, 1 \leq i \leq n)$ of V^* . If $\mathcal{B} = (e_i)$ is the ante-dual basis, then by construction $W_*^\circ = \text{Vect}(\varphi_i, i > d)$.

Applying the argument from (1) to $W = W_*^\circ$ and using the basis $\varepsilon_i = e_{n-i}$, we get $W^\perp = (W_*^\circ)^\perp = \text{Vect}(\varphi_i, i \leq d) = W_*$ which gives (3).

(4) is added for reference and does not use finite dimension (i).

For (5), if $\varphi \in W_*^\circ$ and $w \in W$, then $ev(w)(\varphi) = \varphi(w)$ which is null because $\varphi \in W_*^\circ$ and therefore $ev(W_*^\circ) \subset W^\perp$. Since these two spaces have the same dimension as established previously, this inclusion is an equality.

For (6), if $w \in W$, and $\varphi \in W^\perp$, then $ev(v)(\varphi) = \langle \varphi, v \rangle = 0$ so that $W \subset W^{\perp\perp}$. As these two spaces have the same dimension as established previously, this inclusion is an equality. \square

Example(s) 7.5.0.2. If V is an euclidean space with scalar product $(v, w) \mapsto v.w$, the partial linear map $w \mapsto (v \mapsto v.w)$ has zero kernel and is therefore an isomorphism $W \mapsto W^*$. One checks that this isomorphism identifies W^\perp with the usual Euclidean orthogonal $\{v \in V \mid v.W = \{0\}\}$ recovering the classical dimension formula in Euclidean geometry $\dim(W^\perp) = n - \dim(W)$. Moreover, with this identification, $w \in W \cap W^\perp$ satisfies $w.w = 0$ and therefore is zero ensuring in the Euclidean space the so called usual orthogonal decomposition $W \oplus W^\perp = V$.

Remark(s) 7.5.0.3. Note that orthogonality and polarity are strictly decreasing applications for inclusion.

Corollary 7.5.0.4. Let $\varphi_i \in V^*$, $i = 1, \dots, m$. Then, the rank of $\text{Vect}\{\varphi_i\}$ is that of the evaluation application $\left\{ \begin{array}{l} V \rightarrow k^m \\ v \mapsto (\varphi_i(v))_i \end{array} \right.$

Proof. It suffices to observe that the kernel of the evaluation is the polar of $\text{Vect}\{\varphi_i\}$ and then to invoke the previous proposition and the rank theorem. \square

Exercise(s) 7.5.0.5. Let V be the real vector space of polynomial of degree ≤ 3 . Let $a < c < b$ be reals and define $I \in V^*$ by

$$\langle I, P \rangle = \int_a^b P(t) dt.$$

Compute $\dim \text{Span}(ev_a, ev_c, ev_b, I)$ depending on the value of c . Deduce a formula for I depending only on evaluation forms.

7.6 Biduality Conventions (Finite Dimension)

The previous paragraph allows, in finite dimension therefore, thanks to ev to identify V and its bidual, polar W_*° of W_* and orthogonal W_*^\perp , W and biorthogonal $W^{\perp\perp}$. We generally simply note W_*^\perp for W_*° . Generally, in finite dimension, we consider spaces and dual, but we do not dualize the dual thanks to ev and we simply write $W = W^{\perp\perp}$ whether W is a subspace of V or of V^* .

As an illustration, let's give the algebraic lemma, easy but important, which in real cases is the algebraic content of the theorem of linked extrema in differential geometry (interpret the result in terms of tangent spaces of submanifolds of \mathbf{R}^n in the spirit of the example 7.1.0.2).

Exercise(s) 7.6.0.1. Compare the orthogonal of a sum or intersection of sub vector spaces with the sum or intersection of their orthogonals.

The following lemma is the algebraic part of the search of extrema through constraints equalities (see 16.3 for constraint inequalities).

Lemma 7.6.0.2. Let φ and φ_i , $i \in I$ be linear forms of V . Then, φ is a linear combination of the φ_i if and only if $\cap_i \text{Ker}(\varphi_i) \subset \text{Ker}(\varphi)$.

Proof. By strict decrease of the orthogonal, the condition

$$\cap_i \text{Ker}(\varphi_i) = \text{Vect}(\varphi_i)^\perp \subset \text{Ker}(\varphi) = \text{Vect}(\varphi)^\perp$$

is equivalent to the inclusion

$$\text{Vect}(\varphi) = \text{Vect}(\varphi)^\perp{}^\perp \subset \text{Vect}(\varphi_i)^\perp{}^\perp = \text{Vect}(\varphi_i).$$

□

Exercise(s) 7.6.0.3. Les $\varphi_i, i = 1, \dots, N$ linear forms on V and $\Psi \in \text{Hom}(V, \mathbf{k}^N) = (\varphi_i)$. Prove that the rank of Ψ is the dimension of the span of the φ_i 's.

Remark(s) 7.6.0.4 (Farkas' Lemma). If $\mathbf{k} = \mathbf{R}$, we have an analogous result for finite families of half-spaces H^+, H_i^+ defined by the inequalities $f \geq 0, f_i \geq 0$. Indeed, $\cap_i H_i^+ \subset H^+$ if and only if φ is a linear combination with positive coefficients of the φ_i . See, for example, David Bart, "A short algebraic proof of the Farkas lemma", *Siam Publications SIAM journal on optimization*, 2008, Vol.19 (1), p.234-239.

7.7 Contravariance

Let $V_i, i = 1, 2, 3$, be arbitrary vector spaces,

Definition 7.7.0.1. If $f \in \text{Hom}_{\mathbf{k}}(V_1, V_2)$, we note ${}^t f \in \text{Hom}_{\mathbf{k}}(V_2^*, V_1^*)$ the transpose of f defined by ${}^t f(\varphi_2) = \varphi_2 \circ f$, in other words, $\langle {}^t f(\varphi_2), v_1 \rangle = \langle \varphi_2, f(v_1) \rangle$ for every $\varphi_2 \in V_2^*, v_1 \in V_1$.

Let's recall that a matrix and its transpose have the same rank: this is for instance an immediate consequence of the fact that equivalent matrices have equivalent transpose and that equivalence classes of matrices (with coefficients in a field) are classified by the rank).

We have the following (formal) proposition

Proposition 7.7.0.2. *If $f \in \text{Hom}_{\mathbf{k}}(V_1, V_2)$ and \mathcal{B}_i are bases of V_i .*

1. *The application $f \mapsto {}^t f$ is linear injective.*
2. *If $f_i \in \text{Hom}_{\mathbf{k}}(V_i, V_{i+1})$, we have (contravariance of the transpose) ${}^t(f_2 \circ f_1) = {}^t f_1 \circ {}^t f_2$.*

Assuming further that the V_i 's are finite dimensional, we have

3. *We have $\text{Mat}_{\mathcal{B}_2, \mathcal{B}_1}({}^t f) = {}^t \text{Mat}_{\mathcal{B}_1, \mathcal{B}_2}(f)$.*
4. *$\text{rk}(f) = \text{rk}({}^t f)$.*
5. *With the identifications (7.6), the transposition is involutive.*
6. *$\text{Im}({}^t f) = \text{Ker}(f)^\perp$ and $\text{Ker}({}^t f) = \text{Im}(f)^\perp$.*
7. *If $V_1 = V_2 = V$, a subspace W of V is stable by f if and only if W^\perp is stable by ${}^t f$.*

Proof. Let's just give an argument for 5)(the verification of the rest is left as an **exercise**). First, it suffices to show one of the two formulas (change f to ${}^t f$ and use the involution of the transposition and of the orthogonal). Then, $\text{Im}({}^t f)$ and $\text{Ker}(f)^\perp$ having the same dimension according to 1) and 7.5.0.1, it suffices to prove $\text{Im}({}^t f) \subset \text{Ker}(f)^\perp$. Now, if $f(v_1) = 0$, then $\langle {}^t f(\varphi_2), v_1 \rangle = \langle \varphi_2, f(v_1) \rangle = 0$.

□

7.7.1 Review of Transvections

Proposition 7.7.1.1. *Let $\tau \in \text{End}_{\mathbf{k}}(V)$.*

1. *$H(\tau) = \text{Ker}(\tau - \text{Id})$ is a hyperplane of V containing $D(\tau) = \text{Im}(\tau - \text{Id})$, which is a line in V .*
2. *There exist $\varphi \in V^*$ and $v \in V$, both nonzero, such that $\tau(x) = x + \varphi(x)v$ with $\varphi(v) = 0$.*
3. *The restriction to the affine hyperplane defined by the equation $\varphi(x) = 1$ is a translation by the vector v .*

4. *The matrices of τ are similar to $\text{Id}_n + E_{1,2} = \begin{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & \text{Id}_{n-2} \end{pmatrix}$.*

We say that τ is a transvection of V of type $(D(\tau), H(\tau)) \in \mathbf{P}V \times \mathbf{P}V^*$. If φ, v are as above, let us define $\tau_\lambda(x) = x + \lambda\varphi(x)v$, $\lambda \in \mathbf{k}$. Under these conditions, we have:

- $H(\tau) = \text{Ker}(\varphi), D(\tau) = \langle v \rangle$,
- Transvections of type $(\langle v \rangle, \langle \varphi \rangle)$ are given by τ_λ , $\lambda \in \mathbf{k}^*$, and $\lambda \mapsto \tau_\lambda$ is an injective group morphism $(\mathbf{k}, +) \rightarrow (\text{SL}(V), \times)$,
- ${}^t\tau$ is a transvection of V^* of type $(H(\tau), D(\tau)) \in \mathbf{P}V^* \times \mathbf{P}V$.
- $D(\text{GL}(V)) = \text{SL}(V)$.

Proof. TBD

□

7.8 Additional Exercises

Exercise(s) 7.8.0.1. Let X be any set and V a finite dimensional vector subspace of the \mathbf{R} -vector space of functions from X to \mathbf{R} . Let $n = \dim(V)$.

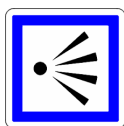
1. Show that the family $(e_{v_x}), x \in X$ generates V^*
2. Show that there exists $f_i \in V, x_i \in X, i = 1, \dots, n$ such that $\det(f_i(x_j)) \neq 0$.
3. Assume that all the functions of V are bounded on X . Show that any pointwise convergent sequence of elements of V is uniformly convergent on X .
4. Does the result previous remain true if one no longer with no boundeness assumption?

Chapter 8

Stable Subspaces



8.1 Perspective



As we will see, the existence of stable spaces strongly depend on the base field, unlike everything that precedes essentially. This explains why this chapter will partly deviate from our desire to provide concrete algorithmic proof through the consideration of particular stable subspaces: the characteristic subspaces. They are indeed defined by the irreducible factor decomposition of the characteristic polynomial of an endomorphism, which depends on the base field and generally cannot be obtained concretely.

Even in the complex case, it is well known that it is not possible to explicitly compute the roots of a polynomial. Nevertheless, even practically, this chapter remains important as there are significant cases where we have access to eigenvalues. It allows, in particular, to understand the topology of matrix similarity classes in the complex case. More generally, we will discuss the aspects of continuity of the constructions in play to the extent that we can only approximate the roots of a polynomial in general.

8.2 Generalities

We know that the stable subspaces by $a \in \text{End}_{\mathbf{k}}(V)$ are its submodules (3.2.4). According to 3.2.3.1, if a is cyclic, these are exactly the $P(a)(V)$ with P being monic divisors of χ . In particular, they are finite in number. Remarkably, the converse is essentially true.

Proposition 8.2.0.1. *If \mathbf{k} is infinite, an endomorphism that has only a finite number of stable subspaces is cyclic.*

Proof. Let a be such an endomorphism. We have to find some cyclic vector for a . The family of stable strict subspaces) of V is a finite family of strict subspaces. Since \mathbf{k} is infinite, their union is not the entire V . Indeed, in the opposite case, their union would be the entire V . Let us then choose for each of these strict subspaces W a non-zero linear form that vanishes on W . The product of these forms is a polynomial function which is identically zero. Since \mathbf{k} is infinite, the ring of polynomial functions on $V = \mathbf{k}^n$ is isomorphic to the ring of polynomials in n variables, a ring that is integral. Thus, one of the forms that is a factor of the product would be identically zero, a contradiction. \square

Obviously, if \mathbf{k} is finite the proposition is false since there is only a finite number of subspaces of V in this case, stable or not.

Remark(s) 8.2.0.2. *When $\mathbf{k} = \mathbf{C}$, any endomorphism a in dimension > 1 admits non-trivial stable spaces (take proper lines). When $\mathbf{k} = \mathbf{R}$, either it admits stable lines (real eigenvalues) or stable planes (take for example the plane defined by the real and imaginary parts of the coordinates of a non-zero eigenvalue vector of the matrix of a in a base or, what comes to the same, consider an irreducible degree 2 polynomial characteristic factor). If $\mathbf{k} = \mathbf{Q}$ and if $P \in \mathbf{Q}[X]$ is irreducible of degree n (take for example $P(X) = X^n - 2$), then the multiplication endomorphism by X on $\mathbf{Q}[X]/(P)$ has no non-trivial stable subspaces since it is cyclic and its minimal does not have a strict divisor: the stable subspaces of an endomorphism depend strongly on the arithmetic of the base field.*

We might hope to treat the general case by reducing to the cyclic case thanks to the Frobenius decomposition 5.13.3. The problem is that the cyclic subspaces that appear are not canonical and thus behave poorly when intersected with a stable subspace. This is not the case for characteristic subspaces.

8.3 Characteristic Subspaces

Let $a \in \text{End}(V)$. We recall (5.6.0.1) that μ_a and χ_a have the same irreducible factors and that the module V_a is annihilated by its characteristic polynomial (Cayley-Hamilton).

Definition 8.3.0.1. Let $\pi \in \mathbf{k}[\mathbf{T}]$ be an irreducible unitary factor of its characteristic polynomial $\chi_a(\mathbf{T})$. The characteristic subspace of a associated with π is the π -primary submodule $V_a[\pi]$ of V_a

$$V_a[\pi] = \cup_{i \geq 1} \text{Ann}_{V_a}[\pi^i] = \cup_{i \geq 1} \text{Ker}(\pi^i(a)).$$

This is a subspace stable by a .

We then have

According to the Chinese remainder theorem 3.7.0.6 and , we then have

Proposition 8.3.0.2. Let $\chi = \prod \pi^{v_\pi(\chi)}$ be the decomposition of $\chi = \chi_a$ into irreducible factors and $\mu_a = \prod \pi^{v_\pi(\mu_a)}$ that of μ_a .

1. We have $v_\pi(\mu_a) \leq v_\pi(\chi)$.
2. There exist $u_{\pi, \chi} \in \mathbf{k}[\mathbf{T}]$ depending only on χ such that $\sum u_{\pi, \chi}(\chi/\pi^{v_\pi(\chi)}) = 1$.
3. We have $V_a = \bigoplus V_a[\pi]$ and the projection p_π on $V_a[\pi]$ parallel to $\bigoplus \pi' \neq \pi V_a[\pi']$ is the homothety of ratio $e_\pi(a) =$ with $e_\pi = u_{\pi, \chi}(\chi/\pi^{v_\pi(\chi)}) \in \mathbf{k}[\mathbf{T}]$.
4. The p_π form an orthogonal family of spectral projectors of V_a i.e. $\sum p_\pi = \text{Id}$ and $p_\pi p_{\pi'} = \delta_{\pi, \pi'} p_\pi$.
5. Each $V_a[\pi]$ is stable by a and $V_a[\pi] = \text{Ker}(\pi^{v_\pi(\chi)}(a)) = \text{Ker}(\pi^{v_\pi(\mu_a)})$.
6. If W is stable by a , we have $W_a[p_i] = V_a[\pi] \cap W$: : any subspace stable by a is the direct sum of its intersections with the characteristic subspaces.
7. We have $\dim_{\mathbf{k}} V_a[\pi] = \deg(\pi^{v_\pi(\chi)}) = v_\pi(\chi) \deg(\pi)$.

Proof. The first 6 points are a rewrite of the Chinese lemma 3.7.0.6 and the functoriality of primary components. For (7), let's recall that each characteristic subspace is stable by a . Since a power of π annihilates $V_a[\pi]$, the characteristic polynomial $\chi_{a|V_a[\pi]}$ of the restriction of a to $V_a[\pi]$ is a power π^{w_π} . But since V_a is the direct sum of the $V_a[\pi]$, we have

$$\prod_{\pi|\chi_a} \pi^{v_\pi(\chi)} = \chi_a = \prod_{\pi|\chi_a} \chi_{a|V_a[\pi]} = \prod_{\pi|\chi_a} \pi^{w_\pi}$$

so that $w_\pi = v_\pi(\chi)$. But

$$\dim_{\mathbf{k}} V_a[\pi] = \deg \chi_{a|V_a[\pi]} = w_\pi \deg(\pi) = v_\pi(\chi) \deg(\pi).$$

□

Note that the spectral projectors are defined by $e_{\pi,\chi} \in \mathbf{k}[T]$ which depends only on χ_a . This illustrates the fact that $p_{\pi,\chi} = e_{\pi,\chi}(a)$ «varies continuously» when a varies continuously, *i.e.* when a is defined by a continuous matrix function $A : \Omega \rightarrow M_n(\mathbf{k})$, whenever the characteristic polynomial $\chi_{A(\omega)}(T)$ is independent of $\omega \in \Omega$. This is also true for characteristic spaces. In particular, their dimension is (locally) constant over Ω under this condition (very strong obviously). The reader will specify the meaning of this statement when $\mathbf{k} = \mathbf{R}$ or $\mathbf{k} = \mathbf{C}$ or, for the learned reader, in the general case for the Zariski topology. This point is crucial, even if it will be somewhat hidden, in the topological study of similarity classes (9 and 8.3.1).

The following lemma is important and follows immediately from the fact that a stable subspace is the sum of its intersections with the characteristic subspaces. This type of result will allow us to reduce the study of the topology of similarity classes to the case of the topology of similarity classes of nilpotent matrices.

Lemma 8.3.0.3 (Invariance by field extension). *Let $A \in M_n(\mathbf{k})$ and $\chi_A = \prod_{\pi} \pi^{v_{\pi}}$ its decomposition into irreducible (unitary) factors. We denote by A, A_K the corresponding endomorphisms of \mathbf{k}^n, K^n . Then $\text{Ker}(\pi^{v_{\pi}(\chi)}(A_K)) = \bigoplus_{\tilde{\pi}|\pi} V_{A_K}[\tilde{\pi}]$ where $\tilde{\pi}$ describes the irreducible unitary divisors of π in $K[T]$.*

8.3.1 Topological properties in the complex case

Consider in this section a sequence of matrices $A_n \in M_d(\mathbf{C})$ (identified as endomorphisms of \mathbf{C}^d) whose eventual convergence to a matrix denoted $A_{\infty} \in M_d(\mathbf{C})$ for the topology defined by a norm¹ based on their projections onto their spectral spaces. As we will focus (see chapter 9) on the case where the A_n are all in the same similarity class (whose closure we seek to study), we further assume that the characteristic polynomial of A_n is a constant polynomial.

By the continuity of the characteristic polynomial in the matrix coefficients (which are polynomials in the coefficients), the convergence of A_n to A_{∞} imposes $\chi_{A_{\infty}}(T) = \det(T \text{Id} - A_{\infty}) = \chi_{A_n}(T)$, a condition that is therefore assumed to be fulfilled.

Let Λ be the spectrum of A_{∞} , the set of complex roots of $\chi_{A_{\infty}}$ and v_{λ} their corresponding multiplicities. As in 8.3.0.2, choose $u_{\lambda}(T) \in \mathbf{C}[T]$ such that

$$\sum_{\lambda \in \Lambda} u_{\lambda}(T) \frac{\chi_{\lambda}(T)}{(X - \lambda)^{v_{\lambda}}} = 1$$

so that the polynomials

$$e_{\lambda}(T) = u_{\lambda}(T) \frac{\chi_{\lambda}(T)}{(X - \lambda)^{v_{\lambda}}}$$

¹As mentioned above, the knowledgeable reader may usefully discuss the case of any infinite field with the matrix space $M_n(\mathbf{k})$ on $V = \mathbf{C}^d$ equipped for example with the Zariski topology, all essential closed sets being defined by polynomial equations as will naturally appear.

define the spectral projectors

$$(i) \quad p_{\lambda,n} = e_{\lambda}(A_n), \quad n \in \bar{\mathbf{N}} = \mathbf{N} \cup \{\infty\}$$

associated with A_n . Since we have

$$\sum_{\Lambda} p_{\lambda,n} = \text{Id}_V$$

it follows that

$$(ii) \quad A_n = \sum_{\Lambda} A_{n,\lambda} \quad n \in \bar{\mathbf{N}}$$

where $A_{n,\lambda} = A_n p_{\lambda,n}$, $n \in \bar{\mathbf{N}}$ is the restriction of A_n on the characteristic space associated with λ and 0 on the others so that $A_{n,\lambda} - \lambda \text{Id}$ is nilpotent. Another way to say is that the semi-simple part of A_n is $\sum \lambda e_{\lambda}(A_n)$.

Proposition 8.3.1.1. *With the previous notations and the assumption $\chi_{A_n}(\mathbb{T})$ independent of $n \in \bar{\mathbf{N}}$, we have $\lim A_n = A_{\infty}$ if and only if for every $\lambda \in \Lambda$, $\lim A_{n,\lambda} = A_{\infty,\lambda}$.*

Proof. This is an immediate consequence of the formulas (i) and (i). □

Lemma 8.3.1.2. *Let $P_{n,d} | \cdots | P_{n,1}$ be the similarity invariants of A_n . Then, the similarity invariants of $A_{n,\lambda}$ are $1, \dots, 1, (X - \lambda)^{v_{\lambda}(P_{n,i})}$, $i = d, \dots, 1$ where the 1s are repeated $d - v_{\lambda}(\chi)$ times.*

Proof. This is another way of writing remark 5.12.0.3. Recall the argument without explicitly invoking the Jordan decomposition. For $A = A_n$ of invariants $P_i = P_{n,i}$, the module V_A is isomorphic to $\oplus k[\mathbb{T}]/P_i(\mathbb{T})$, write the decomposition $P_i = \prod_{\Lambda} (X - \lambda)^{v_{\lambda}(P_i)}$ (since P_i divides χ_A) then invoke the Chinese lemma to write

$$V_A \simeq \oplus_i \oplus_{\Lambda} k[\mathbb{T}]/(\mathbb{T} - \lambda)^{v_{\lambda}(P_i)}.$$

But A_{λ} acts by A on the $(X - \lambda)$ -primary component $V_{A,\lambda} = e_{\lambda}(\mathbb{T})V_A$ and by 0 on the $(X - \tilde{\lambda})$ -primary components $V_{A,\tilde{\lambda}} = e_{\tilde{\lambda}}(\mathbb{T})V_A$ if $\tilde{\lambda} \neq \lambda$. The $(\mathbb{T} - \lambda)$ -primary component (of dimension the multiplicity $v_{\lambda}(\chi)$ of the root λ of χ according to (8.3.0.2) is then written as

$$V_{A,\lambda} = V_A[\mathbb{T} - \lambda] \simeq \oplus_i k[\mathbb{T}]/(\mathbb{T} - \lambda)^{v_{\lambda}(P_i)}$$

and we conclude thanks to the uniqueness of similarity invariants. □

A slightly different form of the previous statement can also be given.

Proposition 8.3.1.3. *With the previous notations and the assumption $\chi_{A_n}(\mathbb{T})$ independent of $n \in \overline{\mathbb{N}}$, we have $\lim A_n = A_\infty$ if and only if the semi-simple (resp. nilpotent) parts of the Jordan-Chevalley decomposition converge to the semi-simple (resp. nilpotent) part of $A_{\infty, \lambda}$.*

Proof. This is an immediate consequence of the fact that there exists $P \in \mathbf{k}[\mathbb{T}]$ depending only on χ such that the semi-simple and nilpotent parts of $A_n, n \in \overline{\mathbb{N}}$ are $P(A_n)$ (resp. $A - P(A_n)$) according to (6.3.2.1). \square

8.3.2 d -th roots in GL_n

If $A \in M_n(\mathbf{k})$ with $\chi_A(\mathbb{T}) = \prod (X - \lambda)^{v_\lambda}$ split, we thus find the usual definition encountered in linear algebra. If $\text{pr}_\lambda = e_\lambda(A)$ is as above, the spectral projector on $V[\mathbb{T} - \lambda] = \text{Ker}(A - \lambda)^{v_\lambda(x)}$, the Jordan-Chevalley decomposition $A = D + N$ is simply calculated by

$$d = \sum \lambda e_\lambda(A) \text{ and } N = A - D$$

as we have just seen. An immediate and useful application is the existence of polynomial d -th roots in the algebraically closed case.

Proposition 8.3.2.1. *Let d be an integer > 0 and assume \mathbf{k} is algebraically closed with characteristic prime to d . Let χ be unitary of degree n . There exists $P_{d, \chi} \in \mathbf{k}[\mathbb{T}]$ such that for any matrix $A \in GL_n(\mathbf{k})$ with $\chi_A = \chi$ we have $P_{d, \chi}(A)^d = A$.*

Proof. Since $\chi(0) \neq 0$, the polynomials χ and \mathbb{T} are coprime and we can write a Bézout identity $U\mathbb{T} + V\chi = 1$ in $\mathbf{k}[\mathbb{T}]$. With the previous notations, since $\chi_D = \chi_A = \chi$, the matrix D is invertible with inverse $U(D)$. Since D and N commute,

$$A = D(\text{Id} + D^{-1}N) = D(\text{Id} + U(D)N)$$

with $D^{-1}N$ being nilpotent. We can then write a d -th root of D as

$$D^{1/d} = \sum \lambda^{1/d} e_\lambda(A)$$

which is therefore a polynomial depending only on χ and d evaluated in A . Furthermore, the coefficients of the power series $(1 + z)^{1/d}$ are the generalized binomial coefficients $\binom{1/d}{i}$, $i \geq 0$ and thus are in $\mathbb{Z}[1/d]$. Since d is invertible in \mathbf{k} and $(D^{-1}N)^n = 0$, we have a d -th root

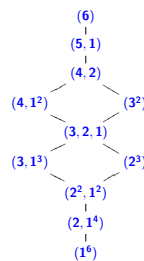
$$(D^{-1}N)^{1/d} = \sum_{i < d} \binom{1/d}{i} (D^{-1}N)^i$$

which is indeed a polynomial depending only on χ and d evaluated in A as are D^{-1} and N , which is what we wanted. \square

We cannot hope for better. On one hand, the statement is clearly false in the general case of non-algebraically closed fields, already in the case $n = 1$. On the other hand, a non-zero nilpotent matrix N does not admit a d -th root. Indeed, it would be nilpotent so that its n -th power would be zero but also equal to $n!$

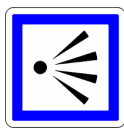
Chapter 9

Topology of Similarity Classes*



Hasse Diagram of GL_6

9.1 Perspective



Here we provide a perspective on the geometry of similarity classes through their topology. To avoid formalism, we restrict ourselves to the usual topology on complex matrices even though the so called Zariski topology whose closed sets are defined by families of polynomial equations would have been more natural¹.

9.2 Introduction

¹As mentioned above, in the case of a general infinite field, the Zariski topology should be considered, which adds no real difficulty once its definition is known. In fact, the topology must be finer than that of Zariski, the usual operations on matrices must be continuous, and the points of \mathbf{k} must not be open, ensuring that the closure of \mathbf{k}^* is \mathbf{k} . This is where the infinitude of the field comes into play in the case of the Zariski topology.

Definition 9.2.0.1. An n -type is a sequence $\underline{P} = (P_n | P_{n-1} | \cdots | P_1)$ of monic polynomials of $\mathbf{k}[T]$ such that $\sum \deg(P_i) = n$. We denote $O(\underline{P})$ the set of matrices in $M_n(\mathbf{k})$ similar to the companion matrix $C(\underline{P})$.

Thus, $O(\underline{P})$ is the orbit of $C(\underline{P})$ under the action of $GL_n(\mathbf{k})$ by conjugation. The theory of similarity invariants tells us that $O(\underline{P})$ consists of matrices with similarity invariants \underline{P} and that $M_d(\mathbf{k})$ is the disjoint union of $O(\underline{P})$ as \underline{P} covers all the n -types (5.9.0.2).

Our goal is to study the closure $\overline{O(\underline{P})}$ of the orbits $O(\underline{P})$. We will therefore assume in the remainder of this chapter that \mathbf{k} is the field of complex numbers \mathbf{C} , with matrix spaces equipped with some norm (let's recall that all matrix norms are equivalent).

We then define a (topological) relation on complex n -types by

$$\underline{P} \preceq \underline{Q} \text{ if and only if } O(\underline{P}) \text{ is contained in the closure } \overline{O(\underline{Q})}.$$

It is clearly a order. Since $\overline{O(\underline{Q})}$ is invariant by conjugation, it is a union of orbits and we have $\overline{O(\underline{Q})} = \cup_{\underline{P} \preceq \underline{Q}} O(\underline{P})$. We will characterize this order in a combinatorial manner as follows.

We define a (combinatorial²) relation on complex n -types by

$$\underline{P} \leq \underline{Q} \text{ if and only if we have the divisibility } \prod_{j \leq i} P_j | \prod_{j \leq i} Q_j \text{ for every } i = 1, \dots, n.$$

It is also a (partial) order. Note that necessarily then $\prod_{i=1}^n P_i = \prod_{i=1}^n Q_i$ for degree reasons.

Theorem 9.2.0.2. Let $\underline{P}, \underline{Q}$ be two complex n -types. Then, $\underline{P} \preceq \underline{Q}$ if and only $\underline{P} \leq \underline{Q}$. In other words, the topological and combinatorial orders on n -types coincide.

Remark(s) 9.2.0.3. This theorem is a reformulation, more transparent in my opinion, of Theorem 4 from [16]. Indeed, to my knowledge, it was Gerstenhaber who fully elaborated the structure of orbit closures, although I have not been able to find this statement *stricto sensu*.

We will proceed by reduction to the nilpotent case using topological results from 8.3.1. Let's start with the crucial case.

9.3 Closure of a Nilpotent Orbit

Thus, we have again a topological order on the partitions of n defined by

²Compare with cf. 9.3.



Nilpotent orbits are classified by partitions \underline{d} of n (5.12.0.2), the dictionary between type and partition being given by $\underline{d} \mapsto T^{\underline{d}} = (T_{d_n}, \dots, T_{d_1})$. We then denote $O(\underline{d})$ the orbit $O(T^{\underline{d}})$ accordingly.

$$\underline{d} \preceq \underline{\delta} \text{ if and only if } O(\underline{d}) \text{ is contained within the closure } \overline{O(\underline{\delta})}$$

and a combinatorial order

$$\underline{d} \leq \underline{\delta} \text{ if and only if for every } i = 1, \dots, n \text{ we have the inequality } \sum_{j \leq i} d_j \leq \sum_{j \leq i} \delta_j.$$

In the nilpotent case, the theorem 9.2.0.2 then becomes

Theorem 9.3.0.1 (Nilpotent Case). *Let $\underline{d}, \underline{\delta}$ be two partitions of n . Then, $\underline{d} \preceq \underline{\delta}$ if and only if $\underline{d} \leq \underline{\delta}$.*

Thus, we aim to show that the topological order \preceq and the combinatorial order \leq on the partitions coincide.

Remark(s) 9.3.0.2. *A partition is always defined by indicating the number of times an integer is repeated, often in ascending order. For $n = 6$, for example, the partition $(3, 1, 1, 1, 0, 0)$ is then denoted $(1^3, 3)$ while the partition $(6, 0, 0, 0, 0, 0)$ is noted as (6) . The diagram describing the order is then called a Hasse diagram. We will not use these notations except in the picture at the beginning of this chapter.*

9.3.1 Order and Duality on Partitions



We use notations and results on nilpotent matrices from 5.12.2. We will demonstrate that the duality of partitions is decreasing for the combinatorial order \leq . For this, and what follows, the key is the classic lemma of disassembling whose proof I reproduce from [27].

We say that $\underline{d} \leq_e \underline{\delta}$ (\underline{d} elementarily inferior to $\underline{\delta}$) if there are indices $i < j$ such that

$$(\delta_1, \dots, \delta_n) = (d_1, \dots, d_{i-1}, d_i + 1, \dots, d_j - 1, \dots, d_n).$$

Obviously

$$\underline{d} \leq_e \underline{\delta} \Rightarrow \underline{d} \leq \underline{\delta}$$

Lemma 9.3.1.1. *Let $\underline{d}, \underline{\delta}$ be two partitions of n . Then, $\underline{d} \leq \underline{\delta}$ if and only if there exists a series of elementary inequalities $\underline{d} = \nu_0 \leq_e \nu_1 \leq_e \dots \leq_e \nu_{N-1} \leq_e \nu_N = \underline{\delta}$.*

Proof. It suffices to prove the existence of a partition $\underline{\nu}$ such that $\underline{d} \leq_e \underline{\nu} \leq_e \underline{\delta}$ when $\underline{d} \neq \underline{\delta}$ and to iterate the process (which stops when $\underline{\nu}_N = \underline{\delta}$.) We thus seek $i < j$ such that $\underline{\nu} \leq_e \underline{\delta}$ with

$$\underline{\nu} = (d_1, \dots, d_{i-1}, d_i + 1, \dots, d_j - 1, \dots, d_n).$$

If $\underline{\nu} = \underline{\delta}$, we are done. Otherwise, $\underline{\nu} < \underline{\delta}$.

There exists therefore k such that

$$(1) \quad d_1 + \dots + d_k < \delta_1 + \dots + \delta_k$$

Let i be the smallest integer k satisfying (1)

Furthermore, as $\sum d_k = \sum \delta_k$, there must exist $k > i$ such that

$$(2). \quad d_1 + \dots + d_k \geq \delta_1 + \dots + \delta_k$$

Let j be the smallest integer $k > i$ satisfying (2).

We have

$$(3) \quad d_1 + \dots + d_k + 1 \leq \delta_1 + \dots + \delta_k \text{ for all } k \in [i, j-1]$$

and

$$(4) \quad d_1 + \dots + d_j = \delta_1 + \dots + \delta_j$$

With these values of i and j , we demonstrate that $\underline{\nu}$ is a partition, *i.e.* $d_{i-1} > d_i$ (or $i = 1$) on one hand and $d_j > d_{j+1}$ on the other.

By construction, i is the smallest integer such that $d_i < \delta_i$ and thus $d_i < \delta_i \leq \delta_{i-1} = d_{i-1}$ (or $i = 1$).

Furthermore, since $d_1 + \dots + d_{j-1} < \delta_1 + \dots + \delta_{j-1}$ and $d_1 + \dots + d_j = \delta_1 + \dots + \delta_j$ $d_j > \delta_j$; since furthermore and $d_1 + \dots + d_{j+1} \leq \delta_1 + \dots + \delta_{j+1}$ we also have $d_{j+1} \leq \delta_{j+1}$. Combining both, we get $d_{j+1} \leq \delta_{j+1} \leq \delta_j < d_j$, which is what we wanted.

We then observe that the inequality $\underline{\nu} \leq_e \underline{\delta}$ is equivalent to (3). □

Corollary 9.3.1.2. *The duality of partitions is strictly decreasing.*

Proof. It suffices to show the decrease in the elementary case $\underline{d} \leq_e \underline{\delta}$. For this, we observe that $\underline{\delta}^*$ satisfies

$$\delta_k^* = \begin{cases} d_k & \text{if } k \neq d_i, d_j \\ d_k - 1 & \text{if } k = d_i \\ d_k + 1 & \text{if } k = d_j \end{cases}$$

so that $\underline{\delta}^* \leq_e \underline{d}^*$. To see this, we note that $d_i > d_j$ and consider the following table

k	\underline{d}^*	$\underline{\delta}^*$	comparison	$\text{Card}(\underline{\delta}^*) - \text{Card}(\underline{d}^*)$
[1,i-1]	$d_k \geq \alpha$	$d_k \geq \alpha$	same	0
i	$d_k \geq \alpha$	$d_k \geq \alpha + 1$	same except if $\alpha = d_i$	-1
[i-1,j-1]	$d_k \geq \alpha$	$d_k \geq \alpha$	same	0
j	$d_k \geq \alpha$	$d_k \geq \alpha - 1$	same except if $\alpha = d_j$	+1
[j+1,n]	$d_k \geq \alpha$	$d_k \geq \alpha$	same	0

using the formula for calculating the dual partition $d_\alpha^* = \text{Card}\{k | d_k \geq \alpha\}$ (5.12.2.2). The proof also provides strict decrease (even though the strict character follows from the fact that duality is involutive)

□

9.3.2 Rank and Nilpotent Orbits



Let M be a nilpotent matrix with associated partition \underline{d} . According to the formula (vi) from 5.12.2, we have for all $n - \text{rk}(M^i) = \sum_{j \leq i} d_j^*$. However, the rank is lower semi-continuous: there exists a neighborhood U of M where all matrices $N \in U$ satisfies $\text{rk}(N) \geq \text{rk}(M)$. If M is in the closure of $O(\underline{\delta})$, this neighborhood intersects $O(\underline{\delta})$: thus, let $N \in U \cap O(\underline{\delta})$. Then $n - \text{rk}(N^i) \leq n - \text{rk}(M^i)$ for all i , meaning $\underline{\delta}^* \leq \underline{d}^*$ and therefore $\underline{\delta} \leq \underline{d}$.

Corollary 9.3.2.1. *Let $\underline{d}, \underline{\delta}$ be partitions of n . Then,*

$$\underline{d} \preceq \underline{\delta} \Rightarrow \underline{d} \leq \underline{\delta}.$$

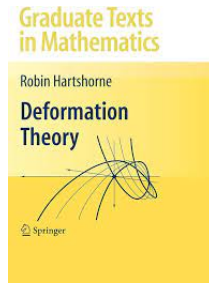
Let us demonstrate the reciprocal implication.

9.3.3 A Nilpotent Matrix Deformation

Following Lemma 9.3.1.1, we simply need to demonstrate the implication in the elementary case. Thus, let $\underline{d} \leq \underline{\delta}$ and let us show that $\underline{d} \preceq \underline{\delta}$. It therefore exists indices $i < j$ such that

$$(\delta_1, \dots, \delta_n) = (d_1, \dots, d_{i-1}, d_i + 1, \dots, d_j - 1, \dots, d_n).$$

We consider $J_{\underline{d}}$ which we want to show is in the closure of $O(\underline{\delta})$, therefore, we want to demonstrate that $J_{\underline{d}}$ is a limit of matrices from $O(\underline{\delta})$.



As \underline{d} and $\underline{\delta}$ only differ at indices i and j , we can assume without loss of generality that we have only two indices. We must therefore show that $J_{(d_i, d_j)}$ is in the closure of $O((d_i - 1, d_j + 1))$. Let us set for example $N(x) = J_{(d_i, d_j)} + xE_{d_i+d_j, d_i}$. This is a triangular block matrix of size $d_i + d_j$ and rank $d_i + d_j - 2$ with $d_i > d_j$. Its type is characterized by its nilpotency index which is $d_i - 1$ (5.12.0.4) for non-zero x so that $N(x)$ is of type $d_i - 1, d_j + 1$. Thus, $N(0) = \lim_{x \rightarrow 0} N(x) \in \overline{O(\underline{\delta})}$ and $\underline{d} \preceq \underline{\delta}$. Hence, recalling 9.3.2.1

$$\underline{d} \preceq \underline{\delta} \iff \underline{d} \leq \underline{\delta}$$

We have therefore proved the theorem 9.3.0.1 in the nilpotent case.

Remark(s) 9.3.3.1. *It is for this argument sequence (and the one in the following paragraph) that the knowledgeable reader wanting to generalize to the Zariski topology of general fields will use the assumption that the field is infinite.*

Let us move to the general case.

9.4 Closure of an Arbitrary Orbit



All work has been done to reduce the general case to the nilpotent case. Let's explain. We consider two n -types $\underline{P}, \underline{Q}$ and study the inclusion $O(\underline{P}) \subset \overline{O(\underline{Q})}$. In other words, we consider a sequence of matrices A_m in $O(\underline{Q})$ converging towards $A_\infty \in O(\underline{P})$. We then freely use the notations and results from 8.3.1.

By the continuity of the characteristic polynomial, it already ensures that $\chi_{A_m}, m \in \overline{\mathbf{N}}$ is a constant polynomial χ whose set of complex roots we denote by Λ . It follows that the characteristic spaces of A_m have a constant dimension d_λ : the multiplicity order of the root λ of χ .

Then (8.3.1.1), we have

$$\lim A_m = A_\infty \text{ if and only if for all } \lambda \in \Lambda, \quad \lim A_{m,\lambda} = A_{\infty,\lambda}$$

But, for each λ , the matrix $A_{m,\lambda} - \lambda \text{Id} \in M_{n,\mathbf{C}}$ is nilpotent and its n -type is (8.3.1.2) is

$$\underline{\delta}_\lambda = 1, \dots, 1, (X - \lambda)^{v_\lambda(\mathbf{Q})}, \quad i = d_\lambda, \dots, 1 \text{ if } n < \infty$$

and

$$\underline{d}_\lambda = 1, \dots, 1, (X - \lambda)^{v_\lambda(P)}, \quad i = d, \dots, 1 \text{ otherwise}$$

where the 1s are repeated $d_\lambda - v_\lambda(\chi)$ times in all cases. But according to the characterization of nilpotent orbits - necessary condition - (9.3.0.1), the existence of this sequence of matrices leads to

(i) $\text{For all } \lambda \in \Lambda, \underline{d}_\lambda \leq \underline{\delta}_\lambda$

Conversely, assuming this condition is satisfied. We denote p_λ the spectral projectors of A_∞ of type \underline{P} . Following the sufficient part of the characterization of nilpotent orbits (9.3.0.1), for every λ there exist nilpotent matrices $N_{m,\lambda}$ that converge to $N_{\infty,\lambda} = A_{\infty,\lambda} - \lambda p_\lambda$. By setting $A_m = \sum_\lambda (N_{m,\lambda} + \lambda p_\lambda)$, we have $\lim A_m = A_\infty$. Thus,

$$\underline{P} \preceq \underline{Q} \iff \text{for all } \lambda \in \Lambda, \underline{d}_\lambda \leq \underline{\delta}_\lambda.$$

Moreover, for two polynomials P, Q whose roots are in Λ , we have

$$P|Q \iff \text{for all } v_\lambda(P) \leq v_\lambda(Q)$$

The condition (i) can therefore be rewritten as

$$\text{for all } i = 1, \dots, n, \text{ we have } \prod_{j \leq i} P_j | \prod_{j \leq i} Q_j$$

This concludes the proof of theorem 9.2.0.2.



9.5 Additional Exercises

Exercise(s) 9.5.0.1. Let \underline{Q} be an n -type and $\chi = \prod Q_i$ the corresponding characteristic polynomial.

1. Show that $O(\underline{\chi}_{red})$ (cf. 6.3.4.1) is the only closed orbit contained in $\overline{O(\underline{Q})}$. Deduce that closed orbits are semi-simple orbits and that $\chi_{red} = (\chi_n, \dots, \chi_1)$ is a minimal type for \preceq .
2. Show that the closure of $O(\underline{\chi}_{red})$ is the set of matrices A such that $\chi_1(A) = 0$ and $\chi_A = \chi$.
3. Generally, show that minimal n -types are of the form $\underline{\chi}_{red}$ for χ monic of degree n . Can you prove this result directly?
4. Conversely, show that maximal n -types are of the form $(1, \dots, 1, \chi)$. Deduce that maximal orbits are those of companion matrices $C(\chi)$.

5. Show that the closure of $O(\mathbf{C}(\chi))$ is the set of matrices A whose $\chi_A = \chi$.

Exercise(s) 9.5.0.2. Let \mathbf{k} be a subfield of \mathbf{C} . Here we consider only n -types \mathbf{k} -rational \underline{d} , i.e. verifying $P_i \in \mathbf{k}[T], i = 1, \dots, n$. We denote $O_{\mathbf{k}}(\underline{d})$ the conjugacy class of $\mathbf{C}(\underline{d})$ under $\mathrm{GL}_n(\mathbf{k})$. Show in this case $O_{\mathbf{k}}(\underline{P}) = O_{\mathbf{C}}(\underline{P}) \cap M_n(\mathbf{k})$. Using 8.3.1.3 and the main theorem 9.2.0.2, show $\overline{O_{\mathbf{k}}(\underline{Q})} = \cup_{\underline{P} \leq \underline{Q}} O_{\mathbf{k}}(\underline{P})$.

Part II

Useful general algebra

Chapter 10

Finiteness Properties of Modules



David Hilbert



Emmy Noether

10.1 Introduction

The notion of Noetherian ring inevitably leads back to Hilbert's foundational paper from 1890 [18] with its three major theorems, the first being the Basis Theorem 10.3.3.1 in the case of polynomial rings. However, as a student rightly pointed out to me, talking only about this (tremendous) paper¹ is unfair. Indeed, it was Emmy Noether who developed the general vision as early as 1920 ([25]).

10.2 Integrality

In addition to the importance of finite-dimensional vector spaces, let us show more generally the importance of finitely generated modules through a few examples.

¹The other two theorems in the article are none other than the Nullstellensatz and the Syzygy Theorem!

10.2.1 Principle of Extension of Algebraic Identities

This principle, extremely useful, is based on a tautology. Let $P \in \mathbf{Z}[T_1, \dots, T_n]$ and $I_i, 1 = 1, \dots, n$ be infinite sets of a field of characteristic zero k . Then, if P is zero on $\prod I_i$, for any ring R and any $(r_i) \in R^n$, we have $P(r_1, \dots, r_n) = 0$. Indeed, we observe that we then have $\mathbf{Z}[T_1, \dots, T_n] \subset k[T_1, \dots, T_n]$ and we reduce by induction to the fact that a polynomial in one variable not identically zero has only a finite number of roots.

Corollary 10.2.1.1. *Let $A \in M_n(\mathbf{R})$ and $\chi_A(T) = \det(T\text{Id} - A)$. Then, $\chi_A(A) = 0$.*

Proof. The matrix equation $\chi_A(A) = 0$ is a system of n^2 polynomial equations with integer coefficients (the coefficients $\chi_A(A) = 0$ where A is the generic matrix $A = [T_{i,j}]$). But these polynomials are zero on $M_n(\mathbf{C})$ according to the usual Cayley-Hamilton theorem. We conclude by setting $I_{i,j} = \mathbf{C}$ thanks to the previous discussion. \square

10.2.2 An Application of Cayley-Hamilton

Proposition 10.2.2.1 (Determinant Trick). *Let f be an endomorphism of a finitely generated R -module M . There exists a monic polynomial $P \in R[T]$ that annihilates f . If additionally $f(M) \subset IM$, it can be assumed that the coefficients of f with index $< \deg(P)$ belongs to I .*

Proof. Let $m_i, 1 \leq i \leq n$ be a finite family of generators of M and consider a matrix $A = [a_{i,j}]$ of f , i.e. for each j , write (in a non-unique way)

$$f(m_j) = \sum_i a_{i,j} m_i.$$

Note that if $f(M) \subset IM$, we can assume $a_{i,j} \in I$. It is then enough to set $P = \det(T\text{Id} - A)$ and invoke, for example, Cayley-Hamilton (10.2.1.1) for $A \in M_n(\mathbf{R})$. \square

By applying the proposition to $f = \text{Id}_M$, we obtain the famous Nakayama Lemma which is very important in advanced commutative algebra.

Corollary 10.2.2.2 (Nakayama). *Let M be a finitely generated module and I an ideal such that $M = IM$. Then, there exists $i \in I$ such that $(1 + i)M = 0$. In particular, if $1 + i$ is invertible (e.g., if i is nilpotent), then $M = 0$.*

10.2.3 Rings of Integers

Let R' be an R -algebra (in other words, consider a ring morphism $R \rightarrow R'$). An element $r' \in R'$ is said to be integral over R if it is annihilated by a monic polynomial with coefficients in R .

Theorem 10.2.3.1. *The subset of R' of elements which integral over R forms a subring of R' .*

Proof. 0 and 1 are integral. We must therefore prove that the difference and the product of two integral elements r' and r'' are integral. Let $M = R[r', r'']$ be the ring of polynomial expressions in r' and r'' with coefficients in R . If r' and r'' are annihilated by monic polynomials of degrees n' and n'' , the family $r'^i r''^j \mid 0 \leq i \leq n', 0 \leq j \leq n''$ generates M and contains $r' - r''$ and $r' r''$. But if $\rho \in M$, the homothety of ratio ρ defines an endomorphism h_ρ of M and thus (10.2.2.1) there exists a monic $P \in R[T]$ such that $P(h_\rho) = h_{P(\rho)} = 0$. Applying to $1 \in M$, we obtain $P(\rho) = 0$ so that all elements of M are integral over R . \square

Corollary 10.2.3.2. *Let k be a subfield of a field k' . Then the subset of elements of k' that are algebraic over k forms a subfield of k' .*

Proof. Following 10.2.3.1 applied to $R = k$, it suffices to show that the inverse of a non-null algebraic element $r' \in k'$ is still nonzero. Suppose therefore P is a unitary annihilator of r' . But then, $T^{\deg(P)} P(1/T)$ is a non-null annihilator of $1/r'$. \square

Remark(s) 10.2.3.3. *With a slight abuse, one often simply say that a complex number which is algebraic over \mathbf{Q} is algebraic, the non algebraic complex numbers being the transcendental ones. A simple countability argument shows that a randomly chosen complex number is almost surely (for the Lebesgue measure) transcendental. For instance, both e (due to C. Hermite, 1873) and π (F. Lindemann, 1883) are transcendental.*

Exercise(s) 10.2.3.4. 1. Show that a rational number is integral over \mathbf{Z} if and only if it is an integer.
2. Show that the minimal degree monic polynomial $P \in \mathbf{Q}[T]$ that annihilates $\exp(\frac{2i\pi}{n})$ has integer coefficients.

10.3 Noetherian Modules

The image of a family of generators of a module through a morphism generates the image module. Thus, every quotient of a finitely generated module is still finitely generated. However, while a submodule of a

finitely generated R module is still finitely generated when R is a field, this is generally not the case (cf 3.2.4). However, it is the case in a Noetherian setting.

Lemma 10.3.0.1. *Let M be an R module. The following properties are equivalent.*

1. *Every submodule of M is finitely generated.*
2. *Every increasing sequence of submodules eventually stabilizes.*
3. *Every non-empty family of submodules of M has a maximal element for inclusion.*

Proof. $1 \Rightarrow 2$. Let M_i be an increasing sequence of submodules. Then, $\cup M_i$ is a submodule of M , thus finitely generated. Choose a finite family of generators: for n large enough, they all belong to M_n and therefore $M_i = M_n$ if $i \geq n$.

$2 \Rightarrow 3$. Let \mathcal{F} be a non-empty family of submodules M without any maximal element (proof by contraposition). We construct a strictly increasing sequence of elements of $\mathcal{F} \neq \emptyset$ by induction by choosing M_0 one of its elements arbitrarily then by induction, assuming the sequence built for $i \leq n$, we observe that M_n is not maximal thus there exists M_{n+1} in \mathcal{F} which strictly contains M_n .

$3 \Rightarrow 1$. Thus, let N be a submodule of M and let \mathcal{F} be the family of its finitely generated submodules. As $\{0\} \in \mathcal{F}$, this family is non-empty. Let N' be a maximal element. It is finitely generated contained in N by construction. Conversely, let $n \in N$. The module $Rn + N'$ is in \mathcal{F} and contains the maximal element N' : therefore, it is equal to it, so that $n \in N'$. We thus have $N' = N$ and therefore N is finitely generated. \square

Definition 10.3.0.2. *1. A module that satisfies the previously mentioned equivalent conditions is said to be Noetherian.*

2. *A ring that is Noetherian as a module over itself is said to be a Noetherian ring.*

Thus, a ring R is Noetherian if it satisfies one of the following three equivalent propositions:

1. Every ideal is finitely generated.
2. Any increasing sequence of ideals eventually stabilizes.
3. Every non-empty family of ideals has a maximal element for inclusion.

Example(s) 10.3.0.3. *Submodules of Noetherian modules are Noetherian (tautological), as are the quotients of Noetherian modules (easy exercise). Fields, principal rings, and quotient rings of Noetherian*

rings are Noetherian. However, a subring of a Noetherian ring is generally not Noetherian (for example, a polynomial ring over a field with an infinity of variables is not Noetherian, whereas it is a subring of its field of fractions which is!).

10.3.1 Stability under exact sequences

Proposition 10.3.1.1. *Consider an exact sequence of modules*

$$0 \rightarrow M_1 \xrightarrow{j} M_2 \xrightarrow{p} M_3 \rightarrow 0.$$

Then M_2 is Noetherian if and only if M_1 and M_3 are.

Proof. The direct part has already been observed in the previous example. Conversely, assume M_1 and M_3 are Noetherian, and let M'_2 be a submodule of M_2 . We have an exact sequence

$$0 \rightarrow j^{-1}(M'_2) \rightarrow M'_2 \rightarrow p(M'_2) \rightarrow 0.$$

But $j^{-1}(M'_2)$ and $p(M'_2)$ are finitely generated as submodules of M_1 and M_3 . Therefore, one can choose a finite family of generators for $p(M'_2)$ of the form $p(g'_{2,i})$ and a finite family of generators $g_{1,k}$ for $j^{-1}(M'_2)$. The finite family $j(g_{1,k}), g'_{2,i}$ of M'_2 generates it. \square

In particular, if R is Noetherian, then R^n is a Noetherian module, and thus so is any quotient. This leads to the following important corollary.

Corollary 10.3.1.2. *The Noetherian modules over a Noetherian ring are exactly the finitely generated modules.*

10.3.2 Existence of Decomposition into Irreducibles in Noetherian Domains

We assume in this section that R is a domain. Recall that $r \in R$ is called irreducible if it is non-zero and non-invertible and its only divisors are either invertible or associated with it. In other words, $r \in R^*$ is irreducible if the equation $r = r_1 r_2$ implies r_1 or r_2 is invertible.

Lemma 10.3.2.1. *Every nonzero and non-invertible element in a Noetherian domain R is a product of irreducible elements.*

Proof. Note that whether r is irreducible only depends on (r) , i.e., it is invariant by multiplication by an invertible. Then, let \mathcal{F} be the set of proper and non-null principal ideals of R whose one of the generators is not a product of irreducibles. If \mathcal{F} were non-empty, it would have a maximal element (r) for inclusion. But r is not irreducible because otherwise $(r) \notin \mathcal{F}$, so r is written $r_1 r_2$ with r_1 and r_2 non-invertible. Thus $(r) \subsetneq (r_i)$. By maximality, $(r_i) \notin \mathcal{F}$ so that each r_i is a product of irreducibles, and so is their product r . A contradiction. \square

Thus, the existence of decomposition into irreducibles is trivial. It is the uniqueness up to multiplication by an invertible (and order aside) that is important (as we will see, this is exactly the content of Euclid's lemma in factorial rings). For example, according to the above, the ring $\mathbf{R}[T, Y]/(T^2 - Y^3)$ is Noetherian, obviously integral (exercise). Yet, the element $T^2 = Y^3$ of the quotient has two decompositions (non-equivalent) because both T and Y are irreducible in the quotient and not associated (exercise).

10.3.3 Hilbert's Basis Theorem

Theorem 10.3.3.1. *Let R be a Noetherian ring.*

1. *The polynomial ring $R[T]$ is Noetherian.*
2. *Every finitely generated R -algebra is a Noetherian ring.*

Proof. The second point is an immediate consequence of the first (by induction, any polynomial ring over R with n variables is Noetherian, and thus so is any quotient). Let's consider the first point.

Let I be an ideal of $R[T]$ and $I^* = I - \{0\}$. If P is a non-null polynomial, denote $\text{dom}(P)$ its highest degree non-null coefficient. The formula $\text{dom}(T^n P) = \text{dom}(P)$ ensures that $\{0\} \cup \text{dom}(I^*)$ is an ideal of R (exercise). It thus has a finite number of generators of the form $\text{dom}(P_i), P_i \in I^*$ which can be assumed to be of the same degree $d \geq 0$ according to the previous formula. An immediate induction then shows $I \cap R_{\geq d}[T] = \langle P_i \rangle$. But $I \cap R_{\leq d}[T]$ is a sub- R -module of $R_{< d}[T] \simeq R^d$: therefore, it is a Noetherian module like R^d (10.3.1.2). One can thus take a finite number of generators Q_j (as an R -module) and the finite family (P_i, Q_j) generates I . \square

We have in fact reused the argument of Euclidean division used to show that $\mathbf{k}[T]$ is principal, the problem being that one can only divide in $\mathbf{k}[T]$ if the leading coefficient of the polynomial is an invertible of R^\times . This is the reason for introducing the ideals of leading coefficients of I .

10.4 Additional Exercises

Exercise(s) 10.4.0.1. *Let G be a finite group operating (on the left) on a ring R . Assume that the cardinality n of G is invertible in R and denote R^G the subring of R of elements invariant by G . Denote $\pi : R \rightarrow R$ the application $x \mapsto \frac{1}{n} \sum_{g \in G} gx$.*

1. Show that p is a projector of image \mathbf{R}^G .
2. Show that p is \mathbf{R}^G linear.
3. Show that if \mathbf{R} is Noetherian, \mathbf{R}^G is Noetherian.

Exercise(s) 10.4.0.2. Let P be a polynomial with integer coefficients P without rational root, d its degree and $x \in \mathbf{R}$ a real root of P . Let $(p, q) \in \mathbf{Z} \times \mathbf{N}^*$.

1. Show $d > 1$.
2. Show $|P(\frac{p}{q})| \geq \frac{1}{q^d}$.
3. Show there exists $C > 0$ such that if $\frac{p}{q} \in [x - 1, x + 1]$ then

$$\left| x - \frac{p}{q} \right| \geq \frac{C}{q^d}.$$

4. Show that $\ell = \sum_{n \geq 0} 10^{-n!}$ is transcendental [Hint : what can you say about the periodicity of a decimal expansion of a rational number ?].

Exercise(s) 10.4.0.3. Let G be an abelian group which is of finite type.

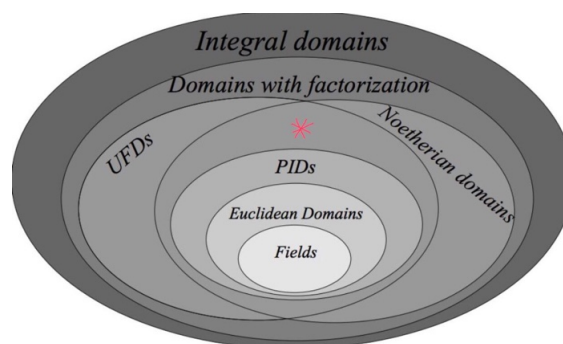
- Prove that there exists an exact sequence $\mathbf{Z}^m \rightarrow \mathbf{Z}^n \rightarrow \mathbf{F} \rightarrow 0$.
- Using the method of chapter 5 and 4.3.2.6, deduce that G is isomorphic to a group of the form $\mathbf{Z}^r \oplus \bigoplus_{i=1}^n \mathbf{Z}/d_i\mathbf{Z}$ with $dn | \dots | d_1$.
- Prove that a finite subgroup of \mathbf{k}^* (where \mathbf{k} is a field) is cyclic.

Exercise(s) 10.4.0.4. TBD

Chapter 11

Reminder on Unique Factorization

Domains



11.1 Introduction

In this chapter, R denotes a *domain*, i.e. an *integral ring* (commutative with unity as usual). Let \mathbf{k} be its field of fractions. Thus, we have a notion of divisibility defining a partial order relation on $R^* = R - \{0\}$. We are interested in rings where a decomposition irreducible factors exists and is reasonably unique. We will constantly use the fact that two elements $a, b \in R$ are equal up to an invertible multiple $u \in R^\times$ if and only if the ideals they generate are equal: they are then said to be *associated* and we write $a \sim b$. This defines an equivalence relation on R compatible with multiplication so that R/\sim is equipped with an associative multiplication with a unit, the class of 1: this is what is called a monoid (commutative with unit). We observe (and use) that the divisibility relation is well defined on the quotient R/\sim . Finally, even the traditional notion of GCD (in principal ideal domain) is only defined up to this equivalence relation, we'll often write equality in this context : for instance, we write $\text{GCD}(a, b) = 1$ for coprime a, b instead writing $\text{GCD} \sim 1$.

11.2 Characterization

Recall that $r \in R$ is said to be irreducible if it is non-zero and non-invertible and if the equation $r = r_1 r_2$ implies r_1 or r_2 is invertible. In other words, $r \in R$ is said to be irreducible if it is non-zero and non-invertible and if its only divisors are either invertible or associated with it.

11.2.1 Uniqueness Criterion

We know that positive irreducible integers are precisely prime numbers. Generally, we only have one implication have

Lemma 11.2.1.1. *Let $r \in R^*$. If the ideal (r) is prime then r is irreducible.*

Proof. If $r = r_1 r_2$, the product $r_1 r_2$ is zero in $R/(r)$ which by definition is integral. Hence, the class $(r_1 \bmod r)$ for example is zero so that $r = \rho_1 r$ and $r = \rho_1 r r_2$. Simplifying by r (integrality), r_2 is invertible. \square

The converse is the the so called Euclid, which in a certain sense is true only in unique factorization domains (see 11.2.1.5 for a precise statement).

Definition 11.2.1.2 (Euclid's Property). *We say (by abuse) that Euclid's lemma is true in R if the ideal generated by an irreducible is prime, that is if any irreducible dividing a product divides one of the factors.*

The following lemma is well known

Lemma 11.2.1.3. *Let r, r_1, r_2 be non zero elements of a principal ideal domain R .*

1. (Gauss lemma) *If $r|r_1 r_2$ and $\text{GCD}(r, r_1) = 1$, then $r|r_2$.*
2. (Euclid Property) *If r is irreducible and $r|r_1 r_2$ then $r|r_1$ or $r|r_2$.*

Proof. For the first point, write (Bézout's theorem) $ur + vr_1 = 1$ for some $u, v \in R$. We get $r_2 = urr_2 + vr_1 r_2$ and therefore $r|r_2$ because r divides each summand.

For the second point, let $d = \text{GCD}(r, r_1)$. Because $d|r$ and r irreducible, d invertible or $d \sim r$. In the second case, we have have done because $r \sim d|r_1$ by definition. In the first case, we apply Gauss lemma and we get $r|r_2$. \square

Definition 11.2.1.4. Let R be a domain and $r, r' \in R$.

1. We will say a decomposition

$$r = u \prod_{i=1}^n p_i$$

with $u \in R^\times$ and p_i irreducible is unique if for any other such decomposition

$$r = u' \prod_{i=1}^{n'} p'_i,$$

we have $r = n'$ and, with renumbering, $p_i \sim p'_i$ for every i . We also say that the (classes of) the p_i 's are unique up to order.

2. A domain is said to be a unique factorization domain (UFD) if every non-zero element has a unique decomposition into irreducible elements in the preceding sense.

The link with what precedes is

Lemma 11.2.1.5 (Uniqueness Lemma). Assume every non-invertible element of R admits a decomposition into irreducible elements. Then, these decompositions are unique if and only if Euclid's lemma is true in R .

Proof. Assume we have uniqueness and let r be irreducible (thus non-zero) with a decomposition $r = r_- r_+$. Let's write the decompositions into irreducibles

$$r_\varepsilon = u_\varepsilon \prod_{j=1}^{n_\varepsilon} p_{\varepsilon,j}$$

getting

$$r = u_- u_+ \prod_{\varepsilon,j} p_{\varepsilon,j}$$

with $\varepsilon = \pm$ say. Thus, we have two decompositions of r into irreducibles, one having of length 1, the other of length $n_- + n_+$. Thus, by uniqueness, $1 = n_- + n_+$ and r is (associated to) one of the $p_{\varepsilon,j}$ (and even the only one) so that it divides r_ε .

Conversely, if Euclid's lemma is true, we prove the uniqueness by n straightforward recursion on the sum of the lengths of two possible decompositions of the same non-zero element. \square

By invoking the existence of decompositions in the Noetherian case (10.3.2.1 and Euclid Property for principal ideal domain (11.2.1.3)), we get

Corollary 11.2.1.6. *An integral Noetherian domain is UFD if and only if it satisfies Euclid's Property. In particular, principal ideal domain are UFD.*

11.3 Transfer

We will demonstrate the transfer theorem from the factorality to polynomial rings

Theorem 11.3.0.1. *If R is UFD, then $R[T]$ is UFD.*

We must therefore demonstrate the uniqueness of decompositions (thus Euclid's lemma) and their uniqueness. For this, we will compare the notion of irreducibles in $R[X]$ and $\mathbf{k}[X]$ using the notion of content (due to Gauss). We will use the equality $(R[T])^\times = R^\times$ which is true for any domain R (just because in this case we have $\deg(PQ) = \deg(P) + \deg(Q)$, see exercise 11.5.0.1 for the general case).

11.3.1 GCD, LCM in UFD

Let (r_i) be a finite family of elements of R which we will assume are not identically zero. Recall that an element $r \in R^*$ is a GCD of the r_i if it is maximal among the common divisors to the r_i . Two GCDs of the same family, when they exist, are of course associated, which is why we speak of the GCD. Therefore, we can consider the GCD, LCM as elements of the monoid R/\sim . Considering maximal common multiples, we obtain the notion of LCM. As with integers, we have

Lemma 11.3.1.1. *If R is UFD, the GCD and the LCM of the (r_i) exist.*

Proof. Consider decompositions into irreducible factors of each of the $r_i \neq 0$ and let q_j be a family of irreducibles not associated with each other so that all these factors are associated with exactly one of the p_i . We can then write uniquely

$$r_i = u_i \prod_j q_j^{v_{i,j}}, \quad v_{i,j} \geq 0 \text{ and } u_i \in R^\times.$$

We then define

$$\text{GCD}(r_i) = \prod_j q_j^{\min_i(v_{i,j})} \text{ and } \text{LCM}(r_i) = \prod_j q_j^{\max_i(v_{i,j})}$$

which are verified to be suitable. □

Note that GCD and LCM are homogeneous of weight 1 for multiplication by R^* .

Exercise(s) 11.3.1.2. *Show that if R is principal, the $\text{GCD}(r_i)$ is a generator of the ideal generated by the (r_i) . Provide a characterization of the LCM in terms of ideals.*

11.3.2 Content

In the remainder of this chapter section, R denotes an UFD domain.

Definition 11.3.2.1. Let $P \in R[T]$ be nonzero. We define the content $c(P)$ of P as the GCD $\in (R/\sim)$ of its coefficients¹. A polynomial with content $c(P) \sim 1$ is said to be primitive.

For example, monic polynomials of $R[T]$ are primitive. The content is homogeneous of weight 1 under multiplication by nonzero element like the GCD.

Theorem 11.3.2.2 (Gauss). Let P, Q be nonzero polynomials of $R[T]$. Then, $c(PQ) \sim c(P)c(Q)$.

Proof. By homogeneity, we may assume P, Q are primitive and we must demonstrate that PQ is primitive. Otherwise, let p be an irreducible of R dividing $c(PQ)$. Since R is UFD, it satisfies Euclid's lemma and the quotient $\bar{R} = R/(p)$ is integral. The coefficient reduction morphism $R \rightarrow \bar{R}$ induces a ring morphism $R[T] \rightarrow \bar{R}[T]$ such that $0 = \overline{PQ} = \bar{P} \cdot \bar{Q}$. Since $\bar{R}[T]$ is integral like \bar{R} , for example $\bar{P} = 0$, i.e. $p|c(P)$, a contradiction because $c(P) \sim 1$. \square

Corollary 11.3.2.3. The irreducibles of $R[T]$ are

1. The irreducibles of R ;
2. Primitive polynomials of $R[T]$ that are irreducible in $\mathbf{k}[X]$.

Proof. Recall the equality $(R[T])^* = R^\times$. The first point follows immediately for reasons of degree.

If P is irreducible in $R[T]$ of degree > 0 , it is certainly primitive according to the first point.

Suppose it is the product of two polynomials $\tilde{P}_1, \tilde{P}_2 \in \mathbf{k}[T]$. By reducing to a common denominator $d_i \in R^*$ for the coefficients of \tilde{P}_i , we can write $\tilde{P}_i = P_i/d_i$ with $P_i \in R[X]$. We then have

$$(*) \quad d_1 d_2 P = P_1 P_2$$

so that $d_1 d_2 = d_1 d_2 c(P) = c(P_1)c(P_2)$ (homogeneity and multiplicativity of content). Replacing in (*), we get

$$P = P_1/c(P_1)P_2/c(P_2)$$

with $P_i/c(P_i) \in R[T]$ by definition of content. As P is irreducible in $R[T]$, we deduce for example $P_1/c(P_1)$ is invertible, thus of degree zero, and therefore the same for \tilde{P}_1 which is proportional to it by a scalar. Hence the irreducibility in $\mathbf{k}[T]$.

The converse is tautological (who can do more can do less) \square

¹Let's emphasize that $c(P)$ belongs to R/\sim , i.e. is only defined up to multiplication by a unit.

11.3.3 The Transfer Theorem

Theorem 11.3.3.1. *If R is UFD, then so is $R[T]$.*

Proof. Because the defining properties of UFD are invariant under multiplication by a unit of R^\times , for simplicity we simply write by an equality an equality up to R^\times .

Existence of decomposition. Let $P \in R[X]$ be non-zero. If P is a constant $r \in R^*$, we write the decomposition $r = \prod p_i$ into irreducible factors in R and invoke (11.3.2.3).

If P is of degree > 0 , by factoring out a GCD of its coefficients, we can assume P is primitive. As in the proof of 11.3.2.3, a common denominator argument then allows us to write its decomposition in the principal therefore UFD $\mathbf{k}[X]$

$$P = \prod P_i/d_i$$

with $P_i \in R[T]$ irreducible in $\mathbf{k}[T]$ and $d_i \in R^*$. By taking the contents, we have $c(P) = \prod d_i$ and $P = \prod P_i/c(P_i)$ which is the sought decomposition.

Uniqueness of decomposition in $R[T]$. Let's demonstrate that $R[T]$ satisfies Euclid's lemma (11.2.1.2). Suppose then P irreducible divides the product of $P_1, P_2 \in R[T]$. If P is of degree > 0 , it is primitive and irreducible in $\mathbf{k}[T]$ according to (11.3.2.3). As $\mathbf{k}[T]$ is UFD since principal, $P|P_1$ for example (in $\mathbf{k}[T]$) and a common denominator argument allows once more to write $dP_1 = Q_1 \cdot P$ with $d \in R^*, Q_1 \in R[T]$. By taking the contents we again have $dc(P_1) = c(Q_1)$ and therefore $P_1 = c(P_1)Q_1/c(Q_1)P$ and thus P divides P_1 in $R[T]$. \square

For example, a polynomial ring in n variables over a field, a principal ring more generally, is UFD. But beware, this remarkable stability of factoriality does not pass to quotients as does the property of being Noetherian. The knowledgeable reader will relate this to the notion of non-singularity in geometry.

Exercise(s) 11.3.3.2. *Show that the ring $\mathbf{R}[X, Y]/(X^2 - Y^3)$ is integral, Noetherian but not UFD.*

11.4 Irreducibility of the Cyclotomic Polynomial Over \mathbf{Q}

From now on, in the rest of this chapter, $k = \mathbf{Q}$ and $\Omega = \mathbf{C}$.

We can take here $\zeta_n = \exp\left(\frac{2\text{Id}\pi}{n}\right)$ so that the primitive n -th roots of unity (in \mathbf{C}) are the complex numbers of the form $\zeta_n^m = \exp\left(\frac{2\text{Id}\pi m}{n}\right)$, where $m \in (\mathbf{Z}/n\mathbf{Z})^*$.

Definition 11.4.0.1. *We define the n -th cyclotomic polynomial*

$$\Phi_n(X) = \prod_{m \in (\mathbf{Z}/n\mathbf{Z})^*} \left(X - \exp\left(\frac{2\text{Id}\pi m}{n}\right) \right).$$

We will show that Φ_n is irreducible and has integer coefficients.

Lemma 11.4.0.2. *We have $\Phi_n(X) \in \mathbf{Z}[X]$.*

Proof. Then, every n -th root of unity has an order d that divides n : it is a primitive d -th root of 1. Conversely, if ζ is a primitive d -th root of 1 with $d|n$, it is an n -th root of 1. We deduce that the set of n -th roots of 1 is the disjoint union parameterized by the divisors d of n of the primitive d -th roots. As

$$X^n - 1 = \prod_{\zeta \in \mu_n} (X - \zeta),$$

we deduce the formula

$$(i) \quad X^n - 1 = \prod_{d|n} \Phi_d(X).$$

Starting from $\Phi_1(X) = X - 1 \in \mathbf{Z}[X]$, we assume by induction on d that Φ_d has integer coefficients according to whatever $d < n$. We just have to recall that the quotient of an integer coefficient polynomial by a monic integer coefficients polynomial is an integer coefficient polynomial (5.2.0.1) to conclude this is also true for $d = n$. \square

But we have in our case the transfert theorem

Lemma 11.4.0.3 (Gauss). *Let $P \in \mathbf{Z}[X]$ be a non-constant polynomial.*

- i) If P is irreducible in $\mathbf{Z}[X]$, it is irreducible in $\mathbf{Q}[X]$.*
- ii) If P is monic, then the monic irreducible factors of the factorization of P in $\mathbf{Q}[X]$ have integer coefficients.*

Proof. It is just an immediate consequence of (11.3.3.1) with $R = \mathbf{Z}$. \square

Definition 11.4.0.4. A complex number is said to be an *algebraic integer* if it is the root of a monic polynomial with integer coefficients.

When the context is clear, we will simply say “integer” instead of algebraic integer.

For example, ζ_n is an integer, but $1/2$ is not (cf. Exercise 11.4.0.5).

The consistency of the terminology is ensured by the following result.

Exercise(s) 11.4.0.5. *Show that $x \in \mathbf{Q}$ is an integer over \mathbf{Z} if and only if it is in \mathbf{Z} .*

Gauss's Lemma 11.4.0.3 immediately gives the following result.

Corollary 11.4.0.6. *The minimal polynomial of an integer element has integer coefficients.*

Then:

Theorem 11.4.0.7. *The cyclotomic polynomial Φ_n is irreducible over \mathbf{Q} .*

The proof, due to Gauss, is very clever.

Proof. Let P be the minimal polynomial of ζ_n . It suffices to prove $\Phi_n | P$, or that all primitive roots of unity cancel P .

Let p be a prime not dividing n and let ζ be a root of P . Then ζ is necessarily a primitive root because $P | \Phi_n$. The key is the following lemma.

Lemma 11.4.0.8. *ζ^p is a root of P .*

Proof. Suppose, by contradiction, the opposite. Write

$$X^n - 1 = P(X)S(X)$$

with $S(X) \in \mathbf{Q}[X]$. Since ζ_n is an integer, we have $P(X) \in \mathbf{Z}[X]$ according to Corollary 11.4.0.6. $P(X)$ being moreover monic, $S(X) \in \mathbf{Z}[X]$. Since $P(\zeta^p)$ is assumed to be non-zero, we have $S(\zeta^p) = 0$. Thus, the polynomials $P(X)$ and $Q(X) = S(X^p)$ have a common complex root. Their GCD (calculated over \mathbf{Q}) is therefore non-constant, so that P divides Q in $\mathbf{Q}[X]$ (irreducibility of P) and also in $\mathbf{Z}[X]$ since P is moreover monic. Reduce modulo p . We obtain

$$\overline{Q}(X) = \overline{S}(X^p) = (\overline{S}(X))^p$$

using the Frobenius morphism. Since by hypothesis $n \neq 0$ in \mathbf{F}_p , $X^n - 1$ and its derivative nX^{n-1} have no common root in $\overline{\mathbf{F}}_p$, so that $X^n - 1$ and \overline{P} have no common factor in $\mathbf{F}_p[X]$. Let Π be an irreducible factor of \overline{P} . As it divides \overline{S}^p , it divides \overline{S} , so that $\Pi^2 | X^n - 1$ in $\mathbf{F}_p[X]$. We obtain a contradiction since \overline{P} is separable. \square

We can now finish the proof of Theorem 11.4.0.7.

Let then ζ be a root of P and ζ' be any root of Φ_n . We write $\zeta' = \zeta^m$ with $\text{GCD}(m, n) = 1$ (because ζ' is primitive). By decomposing m into a product of prime factors, a repeated application of the lemma gives that ζ' is a root of P and therefore $\Phi_n | P$. \square

11.5 Additional exercises

Exercise(s) 11.5.0.1. Let R be a ring.

1. If r is nilpotent, show that $1 + r \in R^\times$.
2. Show that a nilpotent element belongs to any prime ideal of R .

Let r be a non nilpotent element and define the equivalence relation on R by $r_1 \equiv r_2$ if and only if there exist $n \geq 0$ such that $r^n r_1 = r_2$. $R[1/r]$ denote the quotient space and R_{nil} the set of nilpotent elements of R .

3. Show that there exists a unique ring structure such that the quotient map $R \rightarrow R[1/r]$ is a morphism.
4. Prove that $R[1/r]$ is nonzero.
5. Prove that there exists a prime ideal not containing r .
6. Prove that the intersection of prime ideals of R is the set of nilpotent elements of R .
7. Prove $(R[T])^\times = R^\times + TR_{nil}[T]$.

Exercise(s) 11.5.0.2. Let n be a positive integer and z_1, \dots, z_n be complex numbers. Define $P_m(T) = \prod_i (T - z_i^m)$ for $m \geq 0$ and let $V = \mathbf{C}[T]/(P_1(T))$. Finally, let $a \in \text{End}_{\mathbf{C}}(V)$ be the homothety with ratio T defined by $a(Q) = TQ$ for all $Q \in V$ and let $\chi_a(X) = \det(X\text{Id} - a)$ be its characteristic polynomial.

1. Show the formula $\chi_a(X) = P_1(X)$.
2. Show that the z_i are the eigenvalues of a .
3. Show that the dimension of each eigenspace is 1.
4. Show that a is diagonalizable if and only if the z_i are pairwise distinct.
5. Show for all $m \geq 0$ the formula $\chi_{a^m}(X) = P_m(X)$.

Now suppose that $0 < |z_i| \leq 1$ for all i and that $P_1 \in \mathbf{Z}[T]$.

6. Show that the $P_m(T)$ have integer coefficients.
7. Show that the set $\{P_m, m \geq 0\}$ is finite.
8. Conclude that the z_i are roots of unity.

Part III

Metrics on Real and Complex Vector Spaces

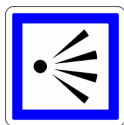
Chapter 12

Euclidean Spaces



Euclid by Raphael

12.1 Perspective



In this section we generalize the warm-up results of chapter 2 to a general Euclidean space emphasizing that most of the results come from this dimension ≤ 2 case. This is consistent with the spatial elementary geometry taught in highschool where most of the time proving theorem needs to reduce to suitable planes or to project to them.

In this chapter, $(E, \langle \cdot, \cdot \rangle)$ will denote an Euclidean space (2.2.0.1). A usual, we write v^2 for $\langle v, v \rangle$ and we have by simple bilinearity

$$(v + w)^2 = v^2 + 2\langle v, w \rangle + w^2$$

proving the usual Pythagoras theorem: $(v + w)^2 = v^2 + w^2$ if and only if v and w are orthogonal.

12.2 Basics on Euclidean Geometry

Most of this section consists in reminders (orthogonality, Gram-Schmidt algorithm, orthogonal and supplementary spaces). Proofs are included for convenience and reference but can certainly be skipped by most of the readers. As announced in § 2, the proof will use two tools: the orthogonal decomposition (12.2.3) allowing to reduce to the dimension ≤ 2 case which is either trivial (dimension 1) or elementary as explained in the warm-up chapter 2.

12.2.1 Examples

Let us first give some examples even we know (and will recall) that two Euclidean spaces of the same dimension are isometric.

Example(s) 12.2.1.1. • *The restriction of a scalar product $E, \langle \cdot, \cdot \rangle$ to a subspace is a scalar product: any such finite dimensional vector subspace has the canonical structure of a Euclidean space, with which it is implicitly equipped.*

- *If (X, μ) is a (positively) measured space, then a scalar product on $L^2(X, \mu; \mathbf{R})$ is defined by*

$$\langle f, g \rangle = \int fg d\mu.$$

Therefore, any of its finite dimensional subspace is an Euclidean space.

- *If $M \in M_{p,q}(\mathbf{R})$, then $\langle M, N \rangle = \text{tr}({}^tMN)$ is a scalar product. For this, consider $M = (a_{ij})$ and compute the diagonal terms of ${}^tMM = (b_{ij})$:*

$$b_{jj} = \sum_k a_{kj}a_{kj} = \sum_k a_{kj}^2$$

and

$$\text{tr}({}^tMM) = \sum_{i,j} a_{ij}^2$$

hence $\langle M, M \rangle > 0$ unless $M = 0$.

12.2.2 Euclidean Norm

Theorem 12.2.2.1. *The mapping $v \mapsto \|v\|$ is a norm called the Euclidean norm.*

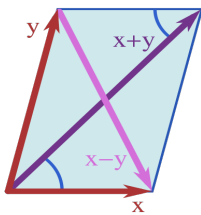
Proof. As in (2.2.1.1), the proof is a consequence of the general Cauchy-Schwartz lemma 12.2.2.2 below. □

Proposition 12.2.2.2 (Cauchy-Schwartz). *Let $v, w \in E$.*

1. *One has $\langle v, w \rangle \leq \|v\|\|w\|$ with equality if and only if v, w are positively colinear.*
2. *One has $|\langle v, w \rangle| \leq \|v\|\|w\|$ with equality if and only if v, w are colinear.*

Proof. This is trivial if $\dim(E) \leq 1$. If not, v, w always belong to a dimension Euclidean 2 to which we apply the Cauchy-Schwartz inequality in dimension 2 (2.2.1.1). \square

This norms benefits the median equality¹, or Apollonius theorem



$$(i) \quad \forall x, y \in E, \quad \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

The Euclidean norm is in fact characterized by the identity of the median:

Exercise(s) 12.2.2.3. *Show that a normed space of finite dimension is Euclidean if and only if for all $x, y \in E$, we have*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

12.2.3 Dual of an Euclidean space, Orthogonal

Proposition 12.2.3.1. *Let F be a subspace of E .*

1. *E admits orthonormal basis.*
2. *The dual of an Euclidean space E is canonically isomorphic to its dual thanks to the isomorphism $v \mapsto (w \mapsto \langle v, w \rangle)$.*
3. *This isomorphism identifies $F^\perp \subset E^*$ with the usual Euclidean orthogonal $\{v \in E \mid \langle v, F \rangle = \{0\}\}$, still denoted F^\perp .*
4. *One has the orthogonal decomposition $F \oplus F^\perp = E$.*

Proof. The proof of (1) is a simple induction, starting in dimension 1 where $v/\|v\|$ is an orthonormal basis for any $v \neq 0$. In dimension $n > 1$, choose a unite vector e_n and observe that the linear form

¹**Exercise:** Explain the denomination "median equality".

$\varphi : w \mapsto (w \mapsto \langle e_n, w \rangle)$ is not zero (because it is positive on e_n). Then, by induction there exists an orthonormal basis on the Euclidean hyperplane $\text{Ker}(\varphi)$ whose together with e_n is an orthonormal basis of E .

The morphism of (2) is an injection $E \hookrightarrow E^*$ and therefore an isomorphism for dimension reasons.

A form $w \mapsto \langle v, w \rangle$ belongs to F^\perp if and only if $\langle v, F \rangle = \{0\}$ proving (3) and therefore the dimension formula $\dim F^\perp + \dim F = \dim E$.

If $v \in F \cap F^\perp$, one has $\|v\|^2 = \langle v, v \rangle = 0$ and therefore $v = 0$. Together with the dimension formula, this gives (4). □

Example(s) 12.2.3.2. In the case of the standard Euclidean \mathbf{R}^n , the identification between $\mathbf{R}^n = (\mathbf{R}^n)^*$ allows the definition of the gradient $\nabla_x(f) \in \mathbf{R}^n$ of a function $f : \mathbf{R}^n \rightarrow \mathbf{R}$ differentiable at a point x by the relation

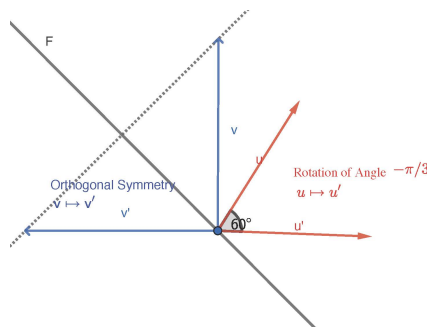
$$\forall v \in \mathbf{R}^n, \langle v, \nabla_x(f) \rangle = df(x).v.$$

Definition 12.2.3.3. For any subspace F of E , we define the orthogonal projection to F (resp. the orthogonal symmetry) with respect to F as $\text{Id}_F \oplus 0_{F^\perp}$ (resp. $\text{Id}_F \oplus (-\text{Id}_{F^\perp})$).

Exercise(s) 12.2.3.4. Let $u \in \mathbf{R}^n$. The Householder matrix is defined as the $n \times n$ matrix H_u given by

$$H_u = I - 2 \frac{u^t u}{\|u\|^2}, \quad \text{if } u \neq 0,$$

and $H_u = I$ if u is the zero vector. Prove that H_u is the orthogonal symmetry of the standard Euclidean \mathbf{R}^n with -1 eigenspace u^\perp .



Exercise(s) 12.2.3.5. Show that the matrix norm of a linear projection in an Euclidean space is ≤ 1 . Show that this norm is equal to 1 if and only if the projection is an orthogonal projection.

12.2.4 Cross product in an oriented Euclidean space

Let E be an oriented Euclidean vector space. Let \mathcal{B} be an orthonormal basis defining the orientation of E (choose for any orthonormal basis and if it is non direct, change one of its vector to its opposite). If \mathcal{B}' is another positively oriented orthonormal basis, the the base change matrix $\text{Mat}(\mathcal{B}, \mathcal{B}')$ is of determinant 1.

In particular, for any vectors $v_1, \dots, v_n \in E$, one has $\det_{\mathcal{B}}(v_1, \dots, v_n) = \det_{\mathcal{B}'}(v_1, \dots, v_n)$ and this value does not depend on \mathcal{B} but on the orientation.

Definition 12.2.4.1. For any $v_1, \dots, v_n \in E$, we denote by $\det(v_1, \dots, v_n)$ the determinant $\det_{\mathcal{B}}(v_1, \dots, v_n)$ where \mathcal{B} is an arbitrary positively oriented orthonormal basis of E .

Proposition 12.2.4.2. Let $v_1, \dots, v_{n-1} \in E$.

1. There exists a unique vector (the cross product of the v_i 's) $v_1 \times \dots \times v_{n-1} \in E$ such that

$$\forall v \in E, \langle v_1 \times \dots \times v_{n-1}, v \rangle = \det(v_1, \dots, v_{n-1}, v).$$

2. The cross-product map $(v_i) \mapsto v_1 \times \dots \times v_{n-1}$ is skew-linear and $v_1 \times \dots \times v_{n-1}$ is orthogonal to the v_i 's.

3. If (v_1, \dots, v_{n-1}) is free, then (v_1, \dots, v_n) is positively oriented. If moreover (v_1, \dots, v_{n-1}) is orthonormal, so is (v_1, \dots, v_n) .

Proof. Recall (12.2) that the dual of an Euclidean space E is canonically isomorphic to its dual thanks to the isomorphism $v \mapsto (w \mapsto \langle v, w \rangle)$ which proves (1).

The skew-linearity of the cross product follows from the skew linearity of the determinant. Moreover, if $i < n$, one has

$$0 = \langle v_1 \times \dots \times v_{n-1}, v_i \rangle = \det(v_1, \dots, v_{n-1}, v_i) = 0$$

proving (2).

If (v_1, \dots, v_{n-1}) is free, the linear form

$$v \mapsto \det(v_1, \dots, v_{n-1}, v)$$

is nonzero and therefore $v_1 \times \dots \times v_{n-1}$ is nonzero. In particular

$$\det(v_1, \dots, v_{n-1}, v_1 \times \dots \times v_{n-1}) = \langle v_1 \times \dots \times v_{n-1}, v_1 \times \dots \times v_{n-1} \rangle > 0$$

proving that $(v_1, \dots, v_{n-1}, v_1 \times \dots \times v_{n-1})$ is a positively oriented basis. If moreover (v_1, \dots, v_{n-1}) is orthonormal, let v_n be the unique normed vector such that $\mathcal{B} = (v_1, \dots, v_{n-1}, v_n)$ is a positively oriented

orthonormal basis. By construction, $\det(v_1, \dots, v_{n-1}, v)$ is the last coordinate $\langle v_n, v \rangle$ of v with respect to \mathcal{B} , and, by definition, so is $\langle v_1 \times \dots \times v_{n-1}, v \rangle$ proving $v_1 \times \dots \times v_{n-1} = v_n$ and (3). \square

12.2.5 Orthogonalization

The following algorithm is the generalization of (2.2.3).

Proposition 12.2.5.1 (Gram-Schmidt Algorithm.). *Let v_1, \dots, v_d be a free family in the Euclidean space E . Then, there exists a unique orthonormal family $\varepsilon_1, \dots, \varepsilon_d$ such that*

1. $\text{Span}\langle v_1, \dots, v_i \rangle = \text{Span}\langle \varepsilon_1, \dots, \varepsilon_i \rangle$ for $i = 1, \dots, d$.
2. $\langle v_i, \varepsilon_i \rangle > 0$ for $i = 1, \dots, d$.

We'll give two ways of thinking the proof.

Proof. Let us observe that the independence of the v_i implies that each $H_i = \text{Span}\langle v_1, \dots, v_i \rangle$ is of dimension i and therefore is a hyperplane in H_{i+1} and that $v_{i+1} \notin H_i$ which moreover carries the natural orientation of its defining basis (v_1, \dots, v_i) .

Geometrical proof. By assumption, one has ε_1 is the only normed vector positively colinear to $v_1 \in H_1$. Assuming $\varepsilon_1, \dots, \varepsilon_i$ have been (uniquely) constructed, they form an orthonormal basis of H_i . Up to sign, there is a unique normed vector in H_{i+1} to the hyperplane H_i (use either the general dimension formula of an orthogonal, or more simply the canonical isomorphism from H_{i+1} to its dual induced by its Euclidean structure). Let e_{i+1} one of the two. Because $v_{i+1} \notin H_i$, one has $\langle v_{i+1}, e_{i+1} \rangle \neq 0$ and it remains to define $\varepsilon_{i+1} = \text{sign}(\langle v_{i+1}, e_{i+1} \rangle)e_{i+1}$.

In other words, (ε_i) is characterized by the fact that each $(\varepsilon_1, \dots, \varepsilon_i)$ is the (unique) a positively oriented orthonormal basis of H_i .

Computational algorithm. Let us recall that $v_{i+1} \notin H_i$ for $i < d$ (with $H_0 = \{0\}$).

Existence. We first define

$$u_1 = v_1 \neq 0, \quad \varepsilon_1 = \frac{u_1}{\|u_1\|},$$

Assuming that u_1, \dots, u_i and an orthonormal suitable system $\varepsilon_1, \dots, \varepsilon_i$ of H_i have been defined, one defines inductively for $i < d$

$$u_{i+1} = v_{i+1} - \langle v_{i+1}, \varepsilon_1 \rangle \varepsilon_1 - \dots - \langle v_{i+1}, \varepsilon_i \rangle \varepsilon_i \neq 0, \quad \varepsilon_{i+1} = \frac{u_{i+1}}{\|u_{i+1}\|}$$

By construction, the family $(\varepsilon_i)_{i \leq d}$ is orthonormal and $v_{i+1} \in \text{Span}(\varepsilon_1, \dots, \varepsilon_{i+1})$ for all $i < d$. We get $H_{i+1} \subset \text{Span}(\varepsilon_1, \dots, \varepsilon_{i+1})$ hence

$$H_{i+1} = \text{Span}(\varepsilon_1, \dots, \varepsilon_{i+1})$$

for dimension reasons. Finally,

$$\langle \varepsilon_{i+1}, v_{i+1} \rangle = \langle \varepsilon_{i+1}, u_{i+1} \rangle = \left\langle \frac{u_{i+1}}{\|u_{i+1}\|}, u_{i+1} \right\rangle = \|u_{i+1}\| > 0$$

hence our recursive algorithm (which is unfortunately numerically unstable). \square

Gram-Schmidt orthogonalization also implies that any subspace F of E has an orthonormal basis whose first $\dim F$ vectors form a basis of F . It could have been used earlier to prove for instance the dimension formula for the orthogonal of a subspace F .

Corollary 12.2.5.2 (Iwasawa or QR Decomposition). *Every matrix $M \in GL_n(\mathbf{R})$ uniquely decomposes into a product $M = QR$ of square matrices with Q orthogonal and R upper triangular with positive diagonal coefficients.*

Proof. Existence. For $j = 1, \dots, n$, let

$$v_j = (m_{i,j})_{i \leq n} \in \mathbf{R}^n$$

be the (independent) column vectors of M and let

$$\varepsilon_j = (q_{i,j})_{i \leq n} \in \mathbf{R}^n$$

the Gram-Schmidt orthonormalization of (v_j) . Finally, let us define

$$r_{l,j} = \langle v_j, \varepsilon_l \rangle$$

By construction $Q = (q_{i,j})$ is orthogonal. Moreover, because

$$v_j \in \text{Span}(\varepsilon_1, \dots, \varepsilon_j)$$

one has

$$v_j = \sum_{l \leq j} \langle v_j, \varepsilon_l \rangle \varepsilon_l = \sum_l r_{l,j} \varepsilon_l \text{ with } r_{l,j} = 0 \text{ if } l > j$$

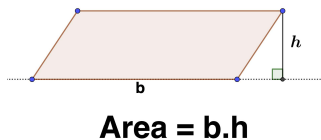
or R upper triangular and $m_{i,j} = \sum_l q_{i,l} r_{l,j}$ meaning precisely $M = QR$. Moreover, $r_{i,i} = \langle v_i, \varepsilon_i \rangle > 0$, hence the existence.

Uniqueness. With obvious notations, if $QR = Q'R'$, we get $T = R'R^{-1}$ both triangular and orthogonal. In particular $T^{-1} = {}^t T$ implying that T is both upper and lower triangular hence diagonal. Because it is orthogonal, its coefficients are equal to ± 1 . But the diagonal coefficients are $r'_{i,i}/r_{i,i} > 0$ and therefore are equal to 1 proving $T = \text{Id}$ hence the uniqueness. \square

Remark(s) 12.2.5.3. *Transposing the QR decomposition, we get the so called LU decomposition of an (invertible) matrix as a product of a lower triangular matrix by an orthogonal one (unitary in the complex case, cf. (17.2.4.2)). This provides another numerical algorithm to invert a matrix, unfortunately numerically unstable like the Gram-Schmidt algorithm. We strongly encourage to implement this decomposition*

in a computer, for example using the SAGE software. One can refine this algorithm using the so-called Householder's matrices (12.7.0.10).

Exercise(s) 12.2.5.4. With the notations above, show that the map $M \mapsto (Q, R)$ is continuous. Deduce that $GL_n(\mathbf{R})$ is homeomorphic (not isomorphic as a group!) to $O_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n+1)}{2}}$ (see the polar decomposition above for another such homeomorphism).



Corollary 12.2.5.5 (Hadamard Inequality). *The Euclidean volume of a parallelepiped is less than the product of the lengths of its sides with equality if and only if it is rectangular.*

Proof. Let v_1, \dots, v_n be a family of n free vectors of \mathbf{R}^n and M the matrix $[v_1, \dots, v_n]$. We want to show $|\det(M)| \leq \prod \|v_i\|$. Keeping the previous notations, we have

$$|\det(M)| = |\det(R)| = \left| \prod (v_i, \varepsilon_i) \right| \stackrel{\text{Cauchy-Schwarz}}{\leq} \prod \|v_i\| \|\varepsilon_i\| = \prod \|v_i\|.$$

In case of equality, equality in Cauchy-Schwarz implies that each v_i is (positively) proportional to ε_i . \square

12.2.6 Gram matrices

The Gram matrix of a finite family $v_i \in E$ is the symmetric real matrix $\text{Gram}(v_i) = (\langle v_i, v_j \rangle)$. We'll denote its determinant by $\text{gram}(x_i)$. It's important to keep in mind that $\text{Gram}(v_i) = \text{Id}$ if and only if v_i is orthonormal. We'll repeatedly use the straightforward formula

$$(ii) \quad {}^tXY = \sum x_i y_i$$

where X, Y are the column vector of coordinates x_i, y_i respectively. The following corollary is well-known and is here for reference (compare with 17.2.5.3 and more generally § 19).

Corollary 12.2.6.1. *Let $v_i, w_j \in E$ be finite families and assume there exists relations $w_j = \sum_i p_{i,j} v_i$. Let $P = (p_{i,j})$ be the corresponding (possibly rectangular) matrix. Finally, let $x = \sum x_i v_i, y = \sum y_i v_i \in E$ and X, Y the column vectors of the x_i, y_i 's. Then, one has*

1. $\langle x, y \rangle = {}^tX \text{Gram}(v_i) Y$
2. $\text{Gram}(w_i) = {}^tP \text{Gram}(v_i) P$.
3. $\text{gram}(v_i) = 0$ if the v_i 's are not independent and is > 0 else.

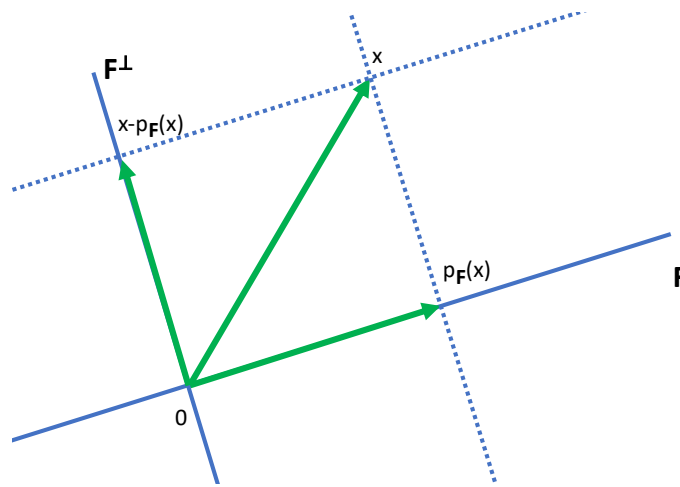
Proof. The first item is a straightforward consequence of 12.2.6.ii.

The second is a direct consequence of the first item.

For the third, assume first that $\sum_i y_i v_i = 0$ with at least one nonzero y_i . Thanks to the first point, we get ${}^t X \text{Gram}(v_i) Y$ for any column X and $Y \neq 0$ the column of the y_i 's. Therefore $Y \in \text{Ker}(\text{Gram}(v_i))$ proving that $\text{gram}(v_i) = 0$. If the v_i 's are independant, let w_i be it's Gram-Schmidt orthonormalization and P be the corresponding (invertible) base change matrix whose Gram matrix is the identity. Thanks to the first item, we get $1 = \text{gram}(v_i) \det(P)^2$. \square

12.2.7 Minimization of distance

It's a higschool result that the distance from a point M to a line ℓ is the length MH where H is the orthogonal projection of M on ℓ .



Proposition 12.2.7.1. *Let $f_i, i = 1, \dots, d$ is a basis of a subspace F of E and $x \in E$. Let p_F be the orthogonal projection to F and $d(x, F) = \inf_{y \in F} \|x - y\|$ the distance form x to F .*

1. *The projection $p_F(x)$ is the only point $y \in F$ such that $d(x, F) = \|x - y\|$.*
2. $d(x, F)^2 = \|x - p_F(x)\|^2 = \frac{\text{gram}(x, f_i)}{\text{gram}(f_i)}$.

Proof. By definition, $x - p_F(x) \in F^\perp$. Therefore, for any $y \in F$, one has

$$\|x - y\|^2 = \|x - p(x)\|^2 \geq \|x - p(x)\|^2 + \|p(x) - y\|^2$$

proving that $y = p(x)$ is the unique solution of the first inequality.

For the second item, observe that $\text{gram}(x, f_i)$ is linear in x and zero if $x \in F$ (because the first column of the Gram matrix is then a linear combination of the others). Therefore,

$$\text{gram}(x, f_i) = \text{gram}(x - p_F(x), f_i)$$

But, $x - p_F(x) \in F^\perp$ implying that the the first row of the Gram matrix is

$$(\|x - p_F(x)\|^2, 0, \dots, 0).$$

Expanding the Gram determinant along this line gives

$$\text{gram}(x, f_i) = d(x, F)^2 \text{gram}(f_i)$$

thanks to the first item. We conclude using 12.2.6.1. \square

The proposition above can be generalized for F only assumed to be closed and convex (16.4.0.1).

Remark(s) 12.2.7.2. *With the notation of the proof of 12.2.5.1, we have the formula for the Gram-Schmidt process*

$$u_{i+1} = v_{i+1} - p_{H_i}(v_{i+1}) \quad \varepsilon_{i+1} = \frac{u_{i+1}}{\|u_{i+1}\|}$$

for any $i < d$.

Exercise(s) 12.2.7.3. *Give at least two different ways to compute $\inf_{a,b \in \mathbf{R}} \int_0^1 (t^2 + at + b)^2 dt$.*

12.3 Geodesic distance on the Euclidean sphere

Let $S_n \in \mathbf{R}^{n+1}$ be the Euclidean sphere of unit vectors and let $x, y \in S_n$. We are interested in length of "smooth paths" $\gamma : [a, b] \rightarrow S_n$ drawn on the Euclidean sphere between x, y and we define $\theta_{xy} = \arccos\langle x, y \rangle$.

If $n = 1$, we identify the plane to the complex plane and write thanks to the lifting theorem $\gamma(t) = \exp(\mathbf{i}\theta(t))$. We have

$$\exp(\mathbf{i}\theta(a)) = x, \exp(\mathbf{i}\theta(b)) = y, \langle x, y \rangle = \text{Re}(\bar{x}y) = \cos(\theta(b) - \theta(a))$$

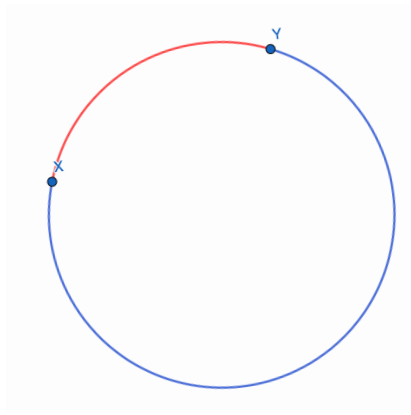
The length of γ is therefore

$$\int_a^b |\theta'| \geq |\theta(b) - \theta(a)|$$

whose class (up to sign and mod 2π) is the non oriented angle between x and y . It follows that the smallest possible length is $\arccos\langle x, y \rangle$ with equality if and only if γ is the usual parametrization of the smallest arc of circle (the red part of the figure below) between x and y (of course, if $x = -y$ they are two such arcs.)

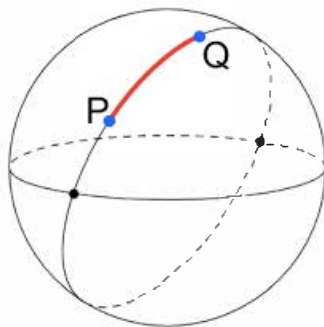
Exercise(s) 12.3.0.1. *Using that \exp is a group morphism that we have the triangle inequality $\theta_{x,z} \leq \theta_{x,y} + \theta_{y,z}$ for any points x, y, z in the unit plane circle. Deduce that $(x, y) \mapsto \theta_{x,y}$ is a distance.*

²At least piecewise C^1 allowing to define the length of γ by $\int \|\gamma'(t)\|$



Proposition 12.3.0.2. For any $n \geq 1$, the map $(x, y) \mapsto \theta_{x,y} = \arccos \langle x, y \rangle$ is a distance.

Proof. The triangle inequality is the only point to check. If we knew that suitable arc of great circle are shortest paths (see 12.7.0.11), the result would follow from the above computation formally. Let us give a (classical³) direct proof.



We know that the Gram determinant $\text{gram}(x, y, z) = \det \begin{pmatrix} 1 & \cos \theta_{xy} & \cos \theta_{xz} \\ \cos \theta_{xy} & 1 & \cos \theta_{yz} \\ \cos \theta_{xz} & \cos \theta_{yz} & 1 \end{pmatrix}$ is ≥ 0 meaning precisely (by direct computation of the determinant) that

$$1 - \cos^2 \theta_{xy} - \cos^2 \theta_{yz} - \cos^2 \theta_{xz} + 2 \cos \theta_{xy} \cos \theta_{yz} \cos \theta_{xz} \geq 0.$$

Adding $\cos^2 \theta_{xz} - \cos^2 \theta_{xz} \geq 0$ we get

$$(1 - \cos^2 \theta_{xy})(1 - \cos^2 \theta_{yz}) - (\cos \theta_{xy} \cos \theta_{yz} - \cos \theta_{xz})^2 \geq 0,$$

or equivalently:

$$\sin^2 \theta_{xy} \sin^2 \theta_{yz} \geq (\cos \theta_{xy} \cos \theta_{yz} - \cos \theta_{xz})^2.$$

Because $\theta_{xy}, \theta_{yz} \in [0, \pi]$, their sinus are ≥ 0 .

³Our exposition is the one of <https://math.stackexchange.com/q/1925049>

Taking square roots of the last inequality we get

$$\sin \theta_{xy} \sin \theta_{yz} \geq |\cos \theta_{xy} \cos \theta_{yz} - \cos \theta_{xz}| \geq \cos \theta_{xy} \cos \theta_{yz} - \cos \theta_{xz}$$

$$\sin \theta_{xy} \sin \theta_{yz} \geq \cos \theta_{xy} \cos \theta_{yz} - \cos \theta_{xz}$$

Using the trigonometrical identity:

$$\cos(\theta_{xy} + \theta_{yz}) = \cos \theta_{xy} \cos \theta_{yz} - \sin \theta_{xy} \sin \theta_{yz}$$

we obtain

$$\cos \theta_{xz} \geq \cos(\theta_{xy} + \theta_{yz})$$

hence the triangle identity because arccos is a decreasing function.

$$\theta_{xz} \leq \theta_{xy} + \theta_{yz}.$$

□

12.4 Adjoint morphism

Proposition 12.4.0.1. *Let $\mathcal{B} = (e_i)$ be a basis of E and f be an endomorphism of E . There exists a unique endomorphism f^* of E , called the adjoint of f such that*

1. For all $x, y \in E$,

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle.$$

2. One has

$$\text{Mat}(\mathcal{B}, f^*) = \text{Gram}(e_i)^{-1} ({}^t \text{Mat}(\mathcal{B}, f)) \text{Gram}(e_i)$$

In particular, f and f^* have the same rank

3. If moreover \mathcal{B} is orthonormal, we have

$$\text{Mat}(\mathcal{B}, f^*) = {}^t \text{Mat}(\mathcal{B}, f).$$

Proof. Let us denote $G = \text{Gram}(e_i)$ and $A = \text{Mat}(\mathcal{B}, f)$. We write the sought identity in terms of matrices taking into account $\langle x, y \rangle = {}^t XGY$ (12.2.6.1)

$${}^t (AX)GY = {}^t X{}^t AGY = {}^t XGG^{-1}{}^t AGY = {}^t XG(G^{-1}{}^t AG)Y$$

which allow to define f^* by the equality $\text{Mat}(\mathcal{B}, f^*) = G^{-1}{}^t AG$. All the items follow immediately. □

For instance, isometries f of E are isomorphisms such that $f^{-1} = f^*$. Usual properties of transposition give the usual formulas (linearity of adjunction, $(f \circ g)^* = g^* \circ f^*$, $\text{Id}^* = \text{Id}$). Note that, like in this Euclidean case, f and f^* are similar (5.5.0.3).

12.5 Comparison

Starting with some endomorphism f of E , we would like to compare f^* with $f, f^{-1}, -f$. We would like to understand when we have an equality between two of these morphisms, giving 6 case.

1. $f^* = f^{-1}$ or f is an isometry (see 12.6.2.3).
2. $f^* = f$ or f self-adjoint (see 12.6.3.1).
3. $f^* = -f$ or f skew-adjoint (see 12.6.5.1).

In all these case, f commutes with f^* .

12.6 Real Normal Endomorphisms



We want to give a nice form of normal real endomorphisms in a suitable orthonormal basis. Let us recall (12.4.0.1) that the matrix in an orthonormal basis of the adjoint of an endomorphism is its transpose. Be aware, this is not true if the basis is not assumed to be orthonormal.

Definition 12.6.0.1. An endomorphism $f \in \text{End}(E)$ (resp. a matrix $M \in M_n(\mathbf{C})$) is normal if $f \circ f^* = f^* \circ f$ (resp. if $M^t M = {}^t M M$).

12.6.1 Reduction of Normal Real endomorphisms

Therefore, after the choice of some orthonormal basis, we will be interested in normal matrices. This problem relies on two statements :

1. The orthogonal of a stable subspace of a normal endomorphism f is stable by f and its restriction to each of these spaces is normal (12.6.1.1). This will allow to reduce to the dimension 2 case.
2. The computation of normal matrices in $M_2(\mathbf{R})$.

This illustrates, once again, that everything boils down to plane geometry!

The tool to reduce to dimension ≤ 2 is the behavior of orthogonals with respect to normal endomorphisms:

Proposition 12.6.1.1. Let $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ be a real block square matrix that commutes with its transpose. Then, $C = 0$. In other words, the orthogonal of a space stable by a normal endomorphism is stable.

three-dimensional Euclidean space. Similarly, in even dimension, it is of the form $\text{diag } R_{\theta_i}$, hence without necessarily an invariant 'axis' as in odd dimension.

Corollary 12.6.2.5. *The continuous $\exp : \mathcal{A}_n(\mathbf{R}) \rightarrow \text{SO}_n(\mathbf{R})$ is surjective, where $\mathcal{A}_n(\mathbf{R})$ denotes the set of skew-symmetric matrices. In particular, $\text{SO}_n(\mathbf{R})$ is path-connected.*

Proof. If A is skew-symmetric, then

$${}^t(\exp(A)) = \exp({}^tA) = \exp(-A) = (\exp(A))^{-1},$$

thus $\exp(A) \in \text{O}_n(\mathbf{R})$. Because $\text{tr}(A) = \text{tr}({}^tA) = -\text{tr}(A)$, one has moreover $\text{tr}(A) = 0$. Therefore, $\det(\exp(A)) = \exp(\text{tr } A) = 1$ proving $\exp(\mathcal{A}_n(\mathbf{R})) \subset \text{SO}_n(\mathbf{R})$.

Because $\exp(\text{O}^{-1}A\text{O}) = \text{O}^{-1}\exp(A)\text{O}$, the reduction theorem (12.6.2.3) shows that the surjectivity statement is equivalent to prove that for any $\theta \in \mathbf{R}$, one has $R_\theta \in \exp(\mathcal{A}_2(\mathbf{R}))$. Once again we are reduced to the dimension 2 statement.

Let $\tilde{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. We have $\tilde{J}^2 = -I$ and therefore

$$\exp(\theta\tilde{J}) = \sum_{k \geq 0} (-1)^k \frac{\theta^{2k}}{k!} \text{Id} + \sum_{k \geq 0} (-1)^k \frac{\theta^{2k+1}}{k!} \tilde{J} = \cos(\theta) \cdot \text{Id} + \sin(\theta) \cdot \tilde{J} = R_\theta.$$

Remark(s) 12.6.2.6. *The previous theorem is a generalization of the usual reduction of rotations in an Euclidean 3-space : for any rotation r , there exists an orthonormal basis such where its matrix is $\text{diag}(1, R_\theta)$ where $\theta \bmod \text{frm}-\pi$ is defined up to sign and is characterized by $\text{tr}(r) = 1 + 2 \cos(\theta)$. Like in the plane situation, to get a well-defined $\theta \bmod \text{frm}-\pi$, we need some orientation. If $r = \text{Id}$, there is no ambiguity. Assume further that E is oriented (2.2.3.1) and $r \neq \text{Id}$. In this case, the axis $\text{Ker}(r - \text{Id})$ is a line. We orient this axis by (arbitrary) changing one of its two unit vector as the first vector e_1 of our basis. Then, there is a unique orientation of the orthogonal plane e_1^\perp compatible with the chosen orientation of E . The restriction of r to this invariant plane is R_θ where $\theta \bmod 2\pi$ is well-defined (2.2.4).*

□

12.6.3 Application to Real Self-adjoint endomorphisms

where ${}^t(\xi_i) = OX$

Assume $\lambda_i > 0$ for all i . Then, $\langle X, X \rangle_S \geq \inf(\lambda_i) \sum \xi_i^2 \geq 0$ and $\langle X, X \rangle_S = 0$ only if $\sum \xi_i^2 = 0$, that is to say of $OX = 0$ and therefore $X = 0$ because O is invertible being orthogonal.

Conversely, assume that S is positive definite and define $X_i = O^{-1}e_i$. Then, $\langle X_i, X_i \rangle_S = \lambda_i > 0$. \square

Notice that another way to understand the proof is that the bilinear form is simply given by the formula $\sum \lambda_i \xi_i^2$ in the coordinates ξ_i of an orthonormal basis of eigenvectors of S .

Exercise(s) 12.6.3.5. Let A, B two positive symmetric matrices. Prove $0 \leq \text{tr}(AB) \leq \text{tr}(A) \text{tr}(B)$.

Corollary 12.6.3.6 (Simultaneous Reduction). Let $S, S' \in M_n(\mathbf{R})$ be two real symmetric matrices with S positive definite. Then, there exists an invertible matrix $\Pi \in \text{GL}_n(\mathbf{R})$ and a diagonal matrix $\Delta \in \text{Diag}_n(\mathbf{R})$ such that

$${}^t\Pi S \Pi = \text{Id} \quad \text{and} \quad {}^t\Pi S' \Pi = \Delta.$$

Proof. We define an Euclidean structure. Choose an orthonormal basis of \mathbf{R}^n for the Euclidean form q_S . If P is the matrix changing its coordinates, we have ${}^tPSP = \text{Id}$. According to the theorem of reduction 12.6.3.1 applied to the symmetric matrix ${}^tPS'P$, hence defining a normal endomorphism for the canonical scalar product of \mathbf{R}^n , there exists $O \in O_n(\mathbf{R})$ such that

$${}^tO{}^tPSPO = \Delta \quad \text{and hence} \quad {}^tOO = \text{Id}.$$

It suffices to define $\Pi = PO$. \square

Remark(s) 12.6.3.7. Note, Δ has nothing to do with the eigenvalues of S' !

Exercise(s) 12.6.3.8. Let S_1, S_2 be symmetric positive definite matrices.

1. Prove $\det(\text{Id} + S_1) \leq 1 + \det(S_1)$.
2. Prove $\det(S_1 + S_2) \leq \det(S_1) + \det(S_2)$.
3. Generalize the precedent inequality if S_1, S_2 are only assumed to be symmetric positive.

12.6.4 Ellipsoid

Let us give a useful consequence of 12.6.3.4.

Definition 12.6.4.1. Let S be a positive definitive of $M_n(\mathbf{R})$. The ellipsoid associated to S is the (convex compact) unit ball $\mathcal{E}_S = \{X \in \mathbf{R}^n \mid \|X\|_S \leq 1\}$. The (orthogonal) lines generated by the eigenvalues of S are called the axis⁵ of \mathcal{E}_S .



Lemma 12.6.4.2. Let $\mathcal{E}_S = \{X \in \mathbf{R}^n, {}^tXSX \leq 1\}$ where $S \in \mathcal{S}_n^{++}(\mathbf{R})$. Then $\text{vol } \mathcal{E}_S = \mu(S) \text{vol } \mathcal{E}_I > 0$ where $\mu : \mathcal{S}_n^{++}(\mathbf{R}) \mapsto (\det S)^{-1/2}$ where vol is the Euclidean volume (defined by the Lebesgue measure for instance).

Proof. Let us first observe that $\text{vol } \mathcal{E}_{Id} > 0$ because its interior is nonempty. Then according to the reduction theorem, there exists $O \in O_n(\mathbf{R})$ such that ${}^tOSO = D$ is a diagonal matrix whose diagonal coefficients $\lambda_1, \dots, \lambda_n$ are all strictly positive. Consider D' the diagonal matrix whose diagonal terms are $1/\sqrt{\lambda_j}$, and denote $R = OD'O^{-1}$, which is invertible and symmetric. Therefore, we have $RSR = Id_n$. Moreover

$$\mathcal{E}_S = \{X \in \mathbf{R}^n, {}^tXR^{-2}X \leq 1\} = \{X \in \mathbf{R}^n, {}^t(R^{-1}X)(R^{-1}X) \leq 1\} = \{X \in \mathbf{R}^n, R^{-1}(X) \in \mathcal{E}_{Id}\} = R(\mathcal{E}_{Id}).$$

By the change of variables formula, we obtain the desired result:

$$\text{vol } \mathcal{E}_S = \det R \text{vol } \mathcal{E}_{Id} = \mu(S) \text{vol } \mathcal{E}_{Id}.$$

□

Remark(s) 12.6.4.3. Notice that \mathcal{E}_{Id} is the standard Euclidean ball. The interested reader will easily prove the well-known volume formula $\text{vol } \mathcal{E}_{Id} = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$ where Γ is the Euler Γ function (*exercise*).

12.6.5 Application to Real Skew-adjoint Endomorphisms

Theorem 12.6.5.1 (Skew-adjoint morphisms Reduction). *The skew-adjoint endomorphisms of an Euclidean space are the endomorphisms u whose matrix in a suitable orthonormal basis is block diagonal where the non-zero blocks are of the form $r \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $r \in \mathbf{R}^*$. Moreover, the scaling coefficients r are uniquely defined up to order.*

⁵O course, this notion is really meaningful only when all the eigenvalues are distinct

8. Show the inequality $\lambda_1(S - S') \leq \lambda_k(S) - \lambda_k(S') \leq \lambda_n(S - S')$.
9. Deduce that the functions $S \mapsto \lambda_k(S)$ are 1-Lipschitz, with the space of symmetric matrices equipped with the operator norm.

Exercise(s) 12.7.0.3. Let $A \in M_n(\mathbf{R})$ be a symmetric matrix with eigenvalues $\alpha_1 \leq \dots \leq \alpha_n$ and $P \in M_{m,n}$ the matrix of an orthogonal projection onto a subspace of dimension m . Let B be the induced matrix $B = PA^tP$ with eigenvalues $\beta_1 \leq \dots \leq \beta_j \leq \dots \leq \beta_m$. Using the results of (12.7.0.2), prove the Cauchy interlacing theorem states: for all $j \leq m$, $\alpha_j \leq \beta_j \leq \alpha_{n-m+j}$.

Exercise(s) 12.7.0.4. Compute the dimension of the space of symmetric matrices in $M_n(\mathbf{R})$. Deduce that the maximal dimension of a linear subspace $M_n(\mathbf{R})$ whose all matrices are nilpotent is $\frac{n(n-1)}{2}$.

Exercise(s) 12.7.0.5. Let $M = \begin{pmatrix} A & B \\ {}^tB & C \end{pmatrix}$ be a symmetric positive definite matrix ($B \in M_{p,q}(\mathbf{R})$ is eventually rectangular). We want to prove $\det(M) \leq \det(A)\det(B)$ with equality if and only if $B = 0$.

1. Prove that A, B are positive definite and $\det(M) > 0$.
2. Assume that $A = \text{Id}_p$, $B = \text{Id}_q$ and let $N = M - \text{Id}_{p+q}$. Prove that if λ is an eigenvalue of N , so is N and that the corresponding eigenspaces have the same dimension. Deduce the required statement in this case.
3. Prove the required statement in the general case.

Exercise(s) 12.7.0.6. Let S_1, S_2 be positive definite symmetric matrices and real $\alpha_1, \alpha_2 \geq 0$ such that $\alpha_1 + \alpha_2 = 1$. After justifying the existence of the integral, prove the formula

$$\int_{\mathbf{R}^n} \exp(-\langle x, x \rangle_S) dx = \left(\frac{\pi}{\det(S)} \right)^{\frac{n}{2}}$$

Using Hölder inequality, prove $\det(\alpha_1 S_1 + \alpha_2 S_2) \geq \det(S_1)^{\alpha_1} \det(S_2)^{\alpha_2}$ with equality if and only if $\alpha_1 \alpha_2 = 0$. Can you generalize if the matrices are only symmetric.

Exercise(s) 12.7.0.7. Let λ_i be a strict increasing sequence of positive real numbers. We recall that the Stone-Weierstrass theorem implies that real polynomial functions are dense in $(\mathcal{E}, \|\cdot\|_2) = L^2([0, 1], d\lambda, \mathbf{R})$ where $d\lambda$ is the Lebesgue measure. Let $\Pi_n = \text{Span}(t^{\lambda_0}, \dots, t^{\lambda_n}) \subset \mathcal{E}$ and q any non negative integer.

1. Prove that $d(t^q, \Pi_n)^2 = \frac{D(q, \lambda_0, \dots, \lambda_n)}{D(\lambda_0, \dots, \lambda_n)}$ where $D(a_1, \dots, a_n) = \det\left(\frac{1}{1+a_i+a_j}\right)$ is the Cauchy determinant of the positive numbers a_1, \dots, a_n .
2. Prove that $\prod_{i=1}^n \prod_{j=1}^n (a_i + b_j) D(a_1, \dots, a_n)$ is a polynomial in a_1, \dots, a_n .
3. Prove the formula $D(a_1, \dots, a_n) = \frac{\prod_{i>j} (a_i - a_j)^2}{\prod_{i,j=1}^n (a_i + a_j + 1)}$.
4. Prove

For all $q \in \mathbf{N}$, $\lim_n d(t^q, \Pi_n)^2 = 0$ if and only if $\sum \frac{1}{\lambda_i} = \infty$.

What density theorem do we have we proved (L^2 -Müntz theorem)?

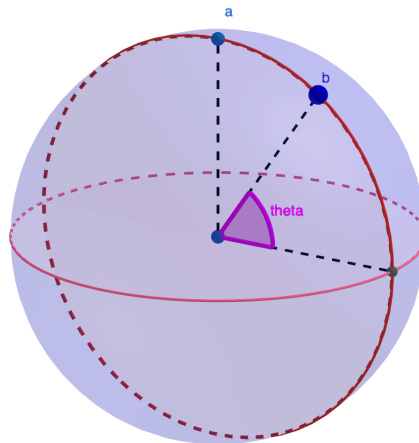
For more advanced versions, see [28] or [1]

Exercise(s) 12.7.0.8. *projection convexe fermé Hilbert. TBD.*

Exercise(s) 12.7.0.9. *QR rectangle. TBD.*

Exercise(s) 12.7.0.10. Let $M \in GL_n(\mathbf{R})$. Prove by induction on n , that there exists Housholder matrices $H_{u_i} \in M_{n-i+1}(\mathbf{R})$ (cf. 12.2.3.4) such that $H_n H_{n-1} \dots H_1 M$ is upper triangular with positive diagonal coefficients and $H_i = \text{diag}(\text{Id}_{i-1}, H_{n-i+1})$. Write a corresponding SAGE code and look at its numerical stability. Write a SAGE code for the QR decomposition using Gram-Schmidt algorithm (12.2.5.2) and experimentally compare these two algorithms in terms of speed and numerical stability.

Exercise(s) 12.7.0.11. Let S be the sphere of unit vector in the Euclidean space \mathbf{R}^3 . By path from a to b we mean any piecewise C^1 map from a compact interval I to \mathbf{R}^3 with image in the sphere starting from a and finishing to b . Recall that the length of γ is $\ell(\gamma) = \int_I \|\gamma'\|$. We fix two such points and we chose coordinates to have $a = (0, 0, 1)$ (the north pole) and $b = (\sin \beta, 0, \cos \beta)$ where $\beta \in [0, \pi]$.



1. γ is defined (spherical coordinates) by

$$t \mapsto (\sin \theta(t) \cos \varphi(t), \sin \theta(t) \sin \varphi(t), \cos \theta(t))$$

with $\theta(t) \in [0, \pi]$, $\varphi(t) \in [0, 2\pi]$. Prove that $\ell(\gamma) \geq \int \sqrt{(\theta'(t))^2} \geq \beta = \arccos \langle a, b \rangle$. What can be said in case of equality ?

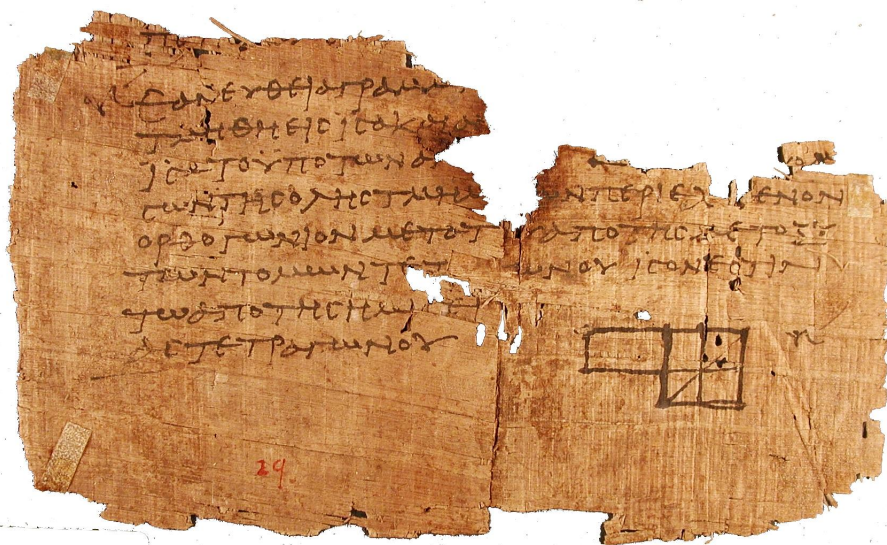
2. Prove that $\ell(\gamma) = 0$ if and only if γ is constant.

⁵Namely continuous, differentiable at all but a finite number points where the derivative is never vanishing and has finite right and left limits.

3. If γ is non injective, prove that there exists a path $\bar{\gamma}$ from a to b such that $\ell(\gamma) \geq \ell(\bar{\gamma})$.
4. Prove that $\ell(\gamma) \geq \arccos \langle a, b \rangle$ with equality if and only if γ is a an injective parametrization of a some arc of great circle joining a and b .
5. Can we release the non vanishing hypothesis of the derivatives ?
6. How to generalize to higher dimensional spheres?

Chapter 13

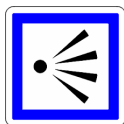
Euclidean Geometry



Postula II.5 of Euclid's

Elements¹

13.1 Perspective



We focus our attention to the properties of the orthogonal group and its links with properties of the general linear group.

¹If a straight line be cut into equal and unequal segments, the rectangle contained by the unequal segments of the whole together with the square on the straight line between the points of section is equal to the square on the half. from a papyrus around 75 A.D.founded in Oxyrhynchus, Egypt

13.2 Topological Properties of the Orthogonal Group

Proposition 13.2.0.1. *The group $O(E)$ is compact and has exactly two connected components that are homeomorphic: $SO(E)$ and $O(E)^- = O(E) \setminus SO(E) = \{u \in SO(E) \mid \det(u) = -1\}$.*

Proof. Consider the map $f : M \in M_n(\mathbf{R}) \mapsto {}^tMM$. We have $O_n(\mathbf{R}) = f^{-1}(I_n)$, therefore $O_n(\mathbf{R})$ is closed in $M_n(\mathbf{R})$. To show that $O(n)$ is bounded, consider the norm induced by $\text{tr}({}^tMM)$. We have $O_n(\mathbf{R}) \subseteq \langle 0, \sqrt{n} \rangle$, proving compactness.

We know that $SO(E)$ is path-connected (12.6.2.5). Moreover, because \det is continuous, we have a partition

$$O(E) = SO(E) \sqcup O^-(E)$$

in two closed sets. If s is any element of $O^-(E)$, the multiplication by s induces an homeomorphism $SO(E) \simeq O(E)$ showing that these closed sets are moreover connected. \square

13.3 Study of \mathcal{S}_n^{++}

We denote \mathcal{S}_n (resp. \mathcal{S}_n^{++}) as the set of real symmetric matrices (resp. positive definite).

Lemma 13.3.0.1. *The spaces \mathcal{S}_n^{++} and \mathcal{S}_n^+ are convex and therefore are path connected).*

Proof. We handle the case of \mathcal{S}_n^{++} . Let $S_0, S_1 \in \mathcal{S}_n^{++}$. We denote, for $s \in [0, 1]$, $S_s = (1-s)S_0 + sS_1$. We have $S_s \in \mathcal{S}_n$, and, for all $X \in \mathbf{R}^n \setminus \{0\}$,

$${}^tXS_sX = (1-s)({}^tXS_0X) + s({}^tXS_1X) > 0.$$

\square

Proposition 13.3.0.2. *The volume application $S \in \mathcal{S}_n^{++} \mapsto \text{vol } \mathcal{E}_S$ is strictly convex.*

Proof. By 12.6.4, we have to prove that $S \mapsto \mu(S) = (\det S)^{-1/2}$ is strictly convex. Let $S_0, S_1 \in \mathcal{S}_n^{++}$ be distinct. We denote, for $s \in [0, 1]$, $S_s = (1-s)S_0 + sS_1$. According to the reduction theorem, there exists $P \in GL_n(\mathbf{R})$ such that ${}^tPS_0P = I_n$ and ${}^tPS_1P = D$ is a diagonal matrix with diagonal coefficients $\lambda_1, \dots, \lambda_n$. If $D = I_n$ then we would have ${}^tPS_0P = {}^tPS_1P$ and $S_0 = S_1$, which contradicts the assumption. Therefore, $D \neq I_n$, and we can assume $\lambda_1 \neq 1$.

Thus,

$$\det S_s = \frac{1}{\det^2 P} \det({}^tPS_sP) = \frac{1}{\det^2 P} \det((1-s)I_n + sD) = \frac{1}{\det^2 P} \prod [(1-s) + s\lambda_j].$$

Let $A_j(s) = (1 - s) + s\lambda_j$ and $u(s) = \mu(S_s)/|\det P|$. This application is differentiable, and

$$u'(s) = \sum_{j=1}^n \frac{-1}{2} \frac{\lambda_j - 1}{A_j(s)^{3/2}} \frac{1}{\prod_{i \neq j} A_i(s)^{1/2}} = \frac{-1}{2} u(s) \sum_{j=1}^n \frac{\lambda_j - 1}{A_j(s)}$$

and

$$u''(s) = \frac{1}{4} u(s) \left(\sum_{j=1}^n \frac{\lambda_j - 1}{A_j(s)} \right)^2 + \frac{1}{2} u(s) \sum_{j=1}^n \left(\frac{\lambda_j - 1}{A_j(s)} \right)^2 \geq \frac{1}{2} u(s) \left(\frac{\lambda_1 - 1}{A_1(s)} \right)^2 > 0.$$

□

Exercise(s) 13.3.0.3. Give another proof of the proposition using the strict concavity of the logarithm. Compare also with (12.7.0.6).

Proposition 13.3.0.4. The application $\exp : \mathcal{S}_n \rightarrow \mathcal{S}_n^{++}$ is a homeomorphism compatible with the transposition.

Proof. The application \exp is continuous, and if $M \in \mathcal{S}_n$, then there exists $O \in O(n)$ such that $O^{-1}MO$ is diagonal. We deduce that $\exp M = O(\exp O^{-1}MO)O^{-1} \in \mathcal{S}_n^{++}$. Conversely, if $M \in \mathcal{S}_n^{++}$, there exists $O \in O(n)$ such that $O^{-1}MO$ is diagonal with strictly positive eigenvalues (12.6.3.1). We can then consider the diagonal matrix N formed by the logarithms of the eigenvalues of M . We have $ONO^{-1} \in \mathcal{S}_n$ and $\exp(ONO^{-1}) = M$.

It remains to see that the map is injective (with a continuous inverse). First, the theorem of reduction allows us to diagonalize a matrix $M \in \mathcal{S}_n$. In this form, $\exp M$ is also diagonal and M and $\exp M$ have the same decomposition into eigenspaces, and the eigenvalues are linked via the numerical exponential. Therefore, if $\exp M = \exp N$, the decomposition into eigenspaces allows us to conclude that $M = N$.

Finally, to see that the inverse application is continuous, it suffices to show that \exp is proper. For this, we equip \mathcal{S}_n with the Euclidean norm $\|M\| = \sqrt{\operatorname{tr}(MM)}$. This norm is obviously invariant under conjugation through conjugation by orthogonal matrices. But, again, real symmetric matrices are orthogonally diagonalizable (12.6.3.1) and therefore $\|M\|^2 = \operatorname{tr} M^2$ is nothing but the sum of the squares of the eigenvalues of M . Consequently, if $\exp M$ stays within a compact set of \mathcal{S}_n^{++} , the eigenvalues of $\exp M$ remain within a compact set of \mathbf{R}_+^* , thus the eigenvalues remain within a compact set of \mathbf{R} , and it follows that M also remains within a compact set of \mathcal{S}_n . Hence, we deduce that \exp is continuous, proper, and injective, thus a homeomorphism onto its image.

The last item is just rewriting the formula ${}^t(\exp(S)) = \exp({}^tS)$. □

Corollary 13.3.0.5. The map $Sq : \mathcal{S}_n^{++} \rightarrow \mathcal{S}_n^{++}$ defined by $Sq(S) = S^2$ is a homeomorphism whose inverse is denoted by $S \mapsto \sqrt{S}$. Moreover $\sqrt{{}^tS} = {}^t\sqrt{S}$.

Proof. Thanks to the preceding result, we can identify Sq to the double map $S \mapsto 2S$ of S_n which is obviously an homeomorphism compatible with the transposition. \square

13.4 Loewner Ellipsoid



Karel Loewner

Theorem 13.4.0.1 (Loewner's Theorem). *If K is a compact subset of \mathbf{R}^n whose interior contains the origin O , then there exists a unique $S \in S^{++}$ such \mathcal{E}_S is the ellipsoid (with center O) of minimal volume containing K .*

Remark(s) 13.4.0.2. *From this, one can formally deduce by polar duality (cf. (16.6.0.5) the John's theorem which asserts the existence of an ellipsoid of maximal volume contained within K . Notice also that the Loewner ellipsoid heavily depends on the interior point O as shown by the example of a plain square as illustrated below.*

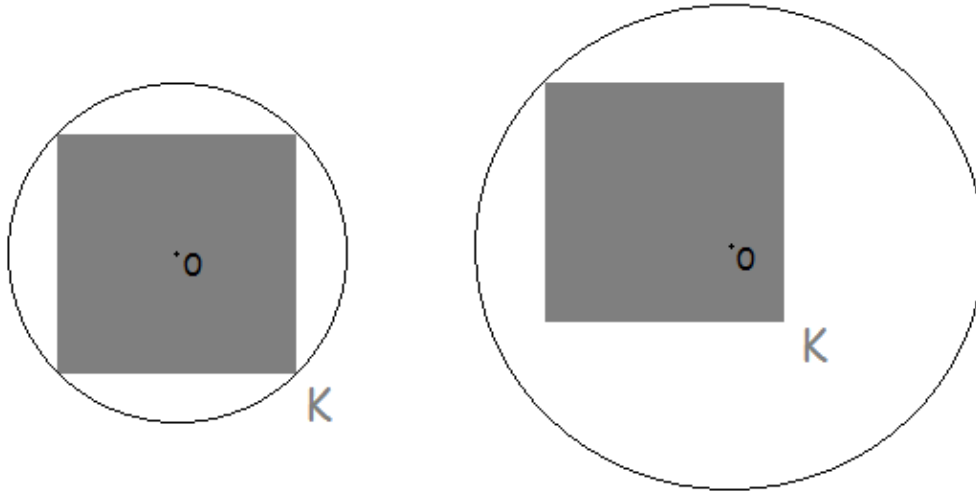
Before demonstrating the theorem, we establish some lemmas.

We can now proceed to the demonstration of the Loewner ellipsoid theorem.

Proof. By assumption, there exist $\rho_1, \rho_2 > 0$ such that $\langle 0, \rho_1 \rangle \subseteq K \subseteq \langle 0, \rho_2 \rangle$.

Consider

$$\mathcal{C} = \{S \in S_n^{++}, K \subseteq \mathcal{E}_S \text{ and } \text{vol}(\mathcal{E}_S) \leq \text{vol}(\langle 0, \rho_2 \rangle)\}.$$



Loewner ellipsoids

is convex (as the volume application is convex (cf. also 12.6.3.6). 13.3.0.2)), non-empty (as $\rho_2^{-1} \text{Id} \in \mathcal{C}$). Let's show that \mathcal{C} is compact. The closure of \mathcal{C} in \mathcal{S}_n is evident since: indeed, to say

$$S \in \mathcal{S}_n^{++} \text{ and } \text{vol}(\mathcal{E}_S) \leq \text{vol}((0, \rho_2)),$$

is to say

$$S \in \mathcal{S}_n^+ \text{ and } \sqrt{\det(S)} \geq \rho_2^{-n}$$

which are closed conditions in \mathcal{S}_n . Furthermore, if x of norm 1, $\rho_1 Sx \in \mathcal{E}_S$ i.e. $q_S(x) \leq \rho_1^{-2}$ thus S is bounded which gives compactness. By continuity the volume application reaches a minimum at at least one point. By strict convexity, this point is unique. \square

13.5 Compact subgroups of $GL_n(\mathbf{R})$

Proposition 13.5.0.1. $SO_n(\mathbf{R})$ (resp. $O_n(\mathbf{R})$) is a maximal compact subgroup of $SL_n(\mathbf{R})$ (resp. of $GL_n(\mathbf{R})$).

Proof. Let $G \subset SL_n(\mathbf{R})$ be a compact subgroup that contains $SO_n(\mathbf{R})$, and let $g \in G \setminus SO_n(\mathbf{R})$. We have $g = OS$ where $O \in SO_n(\mathbf{R})$ since $\det g > 0$ and $S \neq I$ since $g \notin SO_n(\mathbf{R})$. Therefore, $S \in G$. But if v is an eigenvector associated with an eigenvalue λ of S different from 1, then $\text{Log} \|S^n x\|$ tends to infinity, which contradicts the compactness of G . \square

We show that we can improve this result as follows.

Theorem 13.5.0.2. *A compact subgroup G of $GL_n(\mathbf{R})$ is conjugate to a subgroup of $O_n(\mathbf{R})$.*

Proof. First note that if \mathcal{E} is an ellipsoid and if $M \in GL_n(\mathbf{R})$, then $M\mathcal{E}$ is also an ellipsoid. Indeed, if $\mathcal{E} = \{{}^tXSX = 1\}$ with $S \in \mathcal{S}_n^{++}$, then

$$M\mathcal{E} = \{{}^t(M^{-1}X)SM^{-1}X = 1\} = \{{}^tX({}^tM^{-1}SM^{-1})X = 1\}$$

And ${}^tM^{-1}SM^{-1}$ is also positive definite since it is just a change of variables, so $M\mathcal{E}$ is also an ellipsoid.

Let B be the closed unit ball in \mathbf{R}^n . Denote $K = \cup_{g \in G} g(B)$. Then K is compact since G and B are compact, K is invariant by definition, and 0 is an interior point of K since K contains $I(B) = B$. According to Loewner's theorem, there exists a unique ellipsoid \mathcal{E}_S that contains K with minimal volume.

Since G is compact, for every $g \in G$, $|\det g| = 1$. Therefore, $\text{vol } g(\mathcal{E}_S) = \text{vol } \mathcal{E}_S$, and as $K = g(K) \subset g(\mathcal{E}_S)$, we obtain $g(\mathcal{E}_S) = \mathcal{E}_S$. Thus, \mathcal{E}_S is invariant under G . Let T be a square root of S^{-1} . Then $\mathcal{E}_S = T(B)$, and $TGT^{-1} \subset O_n(\mathbf{R})$. \square

Remark(s) 13.5.0.3. *It is easy to see that $O_2(\mathbf{C})$ is not compact, hence neither is $O_n(\mathbf{C})$.*

13.6 Polar Decomposition.

Theorem 13.6.0.1. *The map $\Phi : O(n) \times \mathcal{S}_n^{++} \rightarrow GL_n(\mathbf{R})$ defined by $\Phi(O, S) = OS$ is a homeomorphism of inverse $\Psi : M \mapsto \Psi(M) = (M(\sqrt{{}^tMM})^{-1}, \sqrt{{}^tMM})$.*

Proof. One has ${}^t\sqrt{{}^tMM} = \sqrt{{}^t({}^tMM)} = \sqrt{{}^tMM}$ and therefore

$$(M(\sqrt{{}^tMM})^{-1}){}^t(M(\sqrt{{}^tMM})^{-1}) = M(\sqrt{{}^tMM})^{-1}(\sqrt{{}^tMM})^{-1}{}^tM = M({}^tMM)^{-1}{}^tM = MM^{-1}({}^tM)^{-1}{}^tM = \text{Id}$$

proving $M(\sqrt{{}^tMM})^{-1} \in O_n(\mathbf{R})$ and $\psi(M) \in \mathcal{S}_n^{++} \rightarrow GL_n(\mathbf{R})$. \square

From this, we deduce some results.

Proposition 13.6.0.2. *$GL_n(\mathbf{R})$ has exactly two connected components.*

Proof. By polar decomposition, we have a partition of $GL_n(\mathbf{R})$ in two closed subset

$$GL_n(\mathbf{R}) \approx O_n(\mathbf{R}) \times \mathcal{S}_n^{++} = (SO_n(\mathbf{R}) \times \mathcal{S}_n^{++}) \cup (O_n^-(\mathbf{R}) \times \mathcal{S}_n^{++})$$

each one being connected by (12.6.2.5) and 13.3.0.4 for instance. \square

Exercise(s) 13.6.0.3. Let $D(t)$ be the dilation $\text{Id} + (t - 1)\mathbf{E}_{1,1}$. Show that the map $(t, M) \mapsto D(t)M$ is a homeomorphism from $\mathbf{R}^* \times SL_n(\mathbf{R})$ onto $GL_n(\mathbf{R})$. Using Gaussian elimination, show that $SL_n(\mathbf{R})$ is generated by products of at most n^2 transvections. Conclude that $SL_n(\mathbf{R})$ is connected and then that $GL_n(\mathbf{R})$ has two connected components. What happens over \mathbf{C} ?

Proposition 13.6.0.4. The spaces $GL_n(\mathbf{R})$ and $SL_n(\mathbf{R})$ are respectively homeomorphic to $O_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n+1)}{2}}$ and $SO_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n-1)}{2}}$.

Proof. The application $\exp : \mathcal{S}_n \rightarrow \mathcal{S}_n^{++}$ is a homeomorphism and $\mathcal{S}_n \approx \mathbf{R}^{\frac{n(n+1)}{2}}$, thus

$$GL_n(\mathbf{R}) \approx O_n(\mathbf{R}) \times \mathcal{S}_n^{++} \approx O_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n+1)}{2}}.$$

Similarly, $SL_n(\mathbf{R}) \approx SO_n(\mathbf{R}) \times (\mathcal{S}_n^{++} \cap SL_n(\mathbf{R}))$ and the map $\exp : \mathcal{S}_n \cap \text{tr}^{-1}\{0\} \rightarrow \mathcal{S}_n^{++} \cap SL_n(\mathbf{R})$ is a homeomorphism. \square

13.7 Algebraic Properties of $O_n(\mathbf{R})$

Definition 13.7.0.1. Let us recall that an isometry whose fixed point subspace H is of dimension $n - 1$ is called a reflection: it is the orthogonal symmetry with respect to H . A rotation whose orthogonality conjugate to $\text{diag}(-1, -1, 1, \dots, 1)$ is called a inversion.

Theorem 13.7.0.2. $O(E)$ is generated by reflections. More precisely, if $u \in O(E)$, then u is the product of at most $\dim E - \dim \text{Ker}(\text{Id} - u)$ reflections.

Proof. We first observe that the product of 2 reflections in \mathbf{R}^2 with respect to lines e_1 and e_2 is a rotation by twice the angle between e_1 and e_2 . Therefore, if we write the reduced form of u , each block R_θ accounts for two reflections, while each (-1) counts as one. \square

Exercise(s) 13.7.0.3. Show that $SO(E)$ is generated by inversions.

Let us recall that the center $Z(G)$ of a group is the (commutative normal) subgroup consisting of elements commuting with all elements of G .

Theorem 13.7.0.4. $Z(O(E)) = \{\pm \text{Id}\}$ and, $Z(SO(E)) = \{\text{Id}\}$ if $\dim E$ is odd, $Z(SO(E)) = \{\pm \text{Id}\}$ if $\dim E$ is even and $\dim E \geq 4$, and $Z(SO(E)) = SO(E)$ if $\dim E = 2$.

Proof. Let x be of norm 1. Complete it to an orthonormal basis. The symmetry with respect to x is written

$$\text{Mat}(s_x, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix}.$$

If $u \in O(E)$, then $us_xu^{-1} = s_{u(x)}$, so if $u \in Z(O(E))$, then $s_x = s_{u(x)}$, hence there exists $\lambda_x \in \mathbf{R}$ such that $u(x) = \lambda_x x$. Since $u \in O(E)$, we have $\lambda_x = \pm 1$. This implies that u is a homothety of ratio $\lambda = \pm 1$. Indeed, for x, y independent,

$$u(x + y) = \lambda_{x+y}x + \lambda_{x+y}y = \lambda_x x + \lambda_y y.$$

Regarding the center of $SO(E)$, we reason in the same way. □

Theorem 13.7.0.5. 1. $D(O(E)) = SO(E)$.

2. $D(SO(E)) = SO(E)$ if $\dim E \geq 3$ and $D(SO(E)) = \{\text{Id}\}$ if $\dim E = 2$.

and,

Proof. If $u, v \in O(E)$ then $\det uvu^{-1}v^{-1} = 1$, thus $D(O(E)) \subset SO(E)$. Moreover, even products of reflections generate $SO(E)$. Let's show that products of two reflections are commutators: let x, y be unit vectors. There exists $u \in O(E)$ such that $u(x) = y$. We have $s_y = s_{u(x)} = u \circ s_x \circ u^{-1}$, therefore

$$s_x \circ s_y = s_x \circ u \circ s_x \circ u^{-1} = s_x \circ u \circ s_x^{-1} \circ u^{-1}.$$

□

Theorem 13.7.0.6. $SO(E)$ is simple if $\dim E = 3$ (compare with (13.9.0.2) and (14.3.4.1)).

Proof. Let G be a non-trivial normal subgroup of $SO(E)$. To show that $G = SO(E)$, it suffices to show that G contains a half-turn (angle π rotation). At that point, we will know it contains all by conjugation, and thus that $G = SO(E)$.

If $\theta = \pi$, then we are done.

Let $g \in G$ be non-trivial. As $g \in \text{SO}(E)$, it is a rotation about axis x and angle $\theta \in]0, \pi[$. If $\theta \in]0, \pi/2[$, let N be the first positive integer such that $N\theta \geq \pi/2$. Then, $N\theta \in [\pi/2, \pi[$ and $\cos(N\theta) \leq 0$. Changing g to g^N , one can assume $-1 < \cos(\theta) \leq 0$.

Observe that, for $v \in \text{SO}(E)$, we have $vgv^{-1}g^{-1} \in G$. In particular, if $v = s_y$, where $y \in E \setminus \{0\}$, then $s_y g s_y g^{-1} = s_y \circ s_{g(y)} \in G$.

If y and $g(y)$ are orthogonal, then $s_y \circ s_{g(y)}$ would be a half-turn. To see this, just consider an orthonormal basis containing y and $g(y)$.

To conclude, we therefore look for $y \neq 0$ such that $g(y) \perp y$. Let (x, e_2, e_3) be an orthonormal basis. We have

$$\text{Mat}(g) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

thus if $y = y_1x + y_2e_2 + y_3e_3$, we seek to solve

$$y_1^2 + y_2(y_2 \cos \theta - y_3 \sin \theta) + y_3(y_2 \sin \theta + y_3 \cos \theta) = 0,$$

which means

$$y_1^2 + (y_2^2 + y_3^2) \cos \theta = 0.$$

Thus, a solution exists because $\cos(\theta) \leq 0$. □

13.8 Euclidean Similitude

Euclidean similitudes, or similitudes for short, are endomorphisms u such that there exists a scalar $\lambda = \lambda(u) \in \mathbf{R}^*$ (the similitude ration) such that $\langle u(x), u(y) \rangle = \lambda \cdot \langle x, y \rangle$ (compare with 22.5.0.1). Observe that necessarily $\lambda(u) > 0$.

Matrix-wise, if \mathcal{B} is an orthonormal basis of E , we obtain the following identity:

$${}^t\text{Mat}(u, \mathcal{B}) \text{Mat}(u, \mathcal{B}) = \lambda \text{Id}.$$

Thus, $\det^2 u = \lambda^n$. In particular, if $u \in \text{O}(E)$ and $\lambda > 0$, then $\lambda^{2/n} > 0$ and the similitude ratio of $\lambda^{2/n}u$ is λ .

In particular, we have the exact sequence

$$1 \rightarrow \text{O}(E) \rightarrow \text{GO}(E) \rightarrow \mathbf{R}^{*+} \rightarrow 1.$$

We have the following characterization of Euclidean similarities.

Proposition 13.8.0.1. *Let $u \in \text{GL}(E)$. Then, u is a similitude if and only if u preserves orthogonality, that is*

$$\forall x, y \in E, x \perp y \iff u(x) \perp u(y).$$

Proof. It is straightforward to verify that an Euclidean similitude preserves orthogonality. Conversely, consider an orthonormal basis $\mathcal{B} = (e_1, \dots, e_n)$ of E . Let $\varepsilon_i = u(e_i)$, $i = 1, \dots, n$, which an orthogonal basis by assumption.

Let $\lambda_i = \|\varepsilon_i\|$. It suffices to show that λ_i is independent of i to conclude that u is a similitude. For $i \neq j$, the vectors $e_i + e_j$ and $e_i - e_j$ are orthogonal. Consequently, $u(e_i + e_j) = \varepsilon_i + \varepsilon_j$ and $u(e_i - e_j) = \varepsilon_i - \varepsilon_j$ are also orthogonal and, evaluating their scalar product, we deduce

$$0 = \langle \varepsilon_i + \varepsilon_j, \varepsilon_i - \varepsilon_j \rangle = \|\varepsilon_i\|^2 - \|\varepsilon_j\|^2 = \lambda_i^2 - \lambda_j^2.$$

□

13.9 Additional Exercises

Exercise(s) 13.9.0.1. Let E be an Euclidean space and recall the definition of the norm operator $\|u\| \stackrel{\text{def}}{=} \sum_{x \neq 0} \frac{\|u(x)\|}{\|x\|}$ for $u \in \text{End}(E)$. A point x of a convex set C is called extremal if it cannot be written as the midpoint of two distinct points of this convex set. Define $B = \{u \in \text{End}(E), \|u\| \leq 1\}$ and $G = \text{SO}(n)$ for simplicity. We seek to show that G is the set of extremal point $\text{ext}(B)$ of B .

1. Show by contradiction that every isometry u is extremal in B .

Let us prove the reverse inclusion by contradiction. Let $u \in B$ such that $u \notin G$, and let \mathcal{B} be an orthonormal basis of E . By the polar decomposition, we can write $\text{Mat}(\mathcal{B}, u) = QS$ with $Q \in \text{O}_n(\mathbf{R})$ and $S \in \mathcal{S}_n^{++}(\mathbf{R})$.

2. What can be said about the eigenvalues of S ?
3. Write S as the average of two other well-chosen symmetric matrices and conclude.

Exercise(s) 13.9.0.2. Let $n \geq 5$, and let N be a normal subgroup of $\text{SO}_n(\mathbf{R})$ strictly containing $Z(\text{SO}_n(\mathbf{R}))$. We aim to prove $Z = \text{SO}_n(\mathbf{R})$: in other words, like $\text{SO}_3(\mathbf{R})$, the group $\text{PSO}_n(\mathbf{R}) = \text{SO}_n(\mathbf{R})/Z(\text{SO}_n(\mathbf{R}))$ is simple if $n \geq 5$ (compare with (13.7.0.6) and (14.3.4.1)).

1. Let U be a vector subspace of dimension 3 of \mathbf{R}^n . Define an embedding $\iota_U : \text{O}(U) \hookrightarrow \text{O}_n(\mathbf{R})$ preserving the determinant.
2. Show that it is sufficient to find an element of N different from the identity whose fixed-point space has dimension $\geq n - 3$. Explain why the sought element is then of the form $s = \tau_b \tau_c$ with τ_b, τ_c two orthogonal reflections.
3. Let $b \in E$. Explain why it would be tempting to consider $s = \rho \tau_b \rho^{-1} \tau_b^{-1}$ with $\rho \in N$ and $c = \rho(b)$ but not conclusive.
4. Consider $a \in E$ and let $\sigma = \tau_a \tau_b$. Describe the geometric nature of $s = \rho \sigma \rho^{-1} \sigma^{-1}$? Show that it is sufficient to find a $\rho \in N$ different from $\pm \text{Id}$ having a fixed point.

5. Let $v \in \mathbf{N}$ different from $\pm \text{Id}$ and u an inversion with P equal to the plane $\text{Ker}(u - \text{Id})$. Show that $\rho = vuv^{-1}u^{-1}$ fixes $P^\perp \cap v(P)^\perp$. Conclude.

6. Can you give a simpler proof of the last item in the odd dimension case?

Exercise(s) 13.9.0.3. TBD If C is a convex subset of E , we define its polar dual $C^* = \{x \in E \mid \forall c \in C, \langle x, c \rangle \leq 1\}$.

Exercise(s) 13.9.0.4. TBD Simplicité $SO_3(\mathbf{R})$ et Iwasawa.

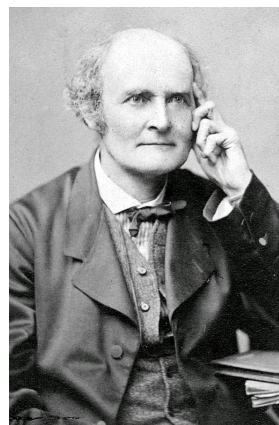
Exercise(s) 13.9.0.5. $SO_3(\mathbf{Q})$ n'est pas simple (regarder les rotations de la forme $\text{Id} + 2^n M$ avec M à coeffs dans $\mathbf{Z}_{(2)}$).

Chapter 14

Quaternion Algebra and Euclidean Geometry in dimension ≤ 4

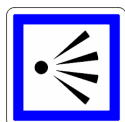


William Hamilton



Arthur Cayley

14.1 Perspective



Although it would have been possible to avoid the use of quaternions, we introduce this important notion to give a geometrical study of orthogonal groups of small dimension.

14.2 Construction

Let \mathbf{H} be the real vector subspace of $M_2(\mathbf{C})$ of matrices of the following form

$$q(a, b) = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}, \quad a, b \in \mathbf{C}$$

Writing $a = t + it_{\mathbf{J}}$, $b = t_{\mathbf{I}} + it_{\mathbf{K}}$, $t, t_i \in \mathbf{R}$ we get

$$q = t\mathbf{1} + t_{\mathbf{I}}\mathbf{I} + t_{\mathbf{J}}\mathbf{J} + t_{\mathbf{K}}\mathbf{K}$$

with

$$\mathbf{1} = \text{Id}, \mathbf{I} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

It follows that these four matrices define a \mathbf{R} -basis of \mathbf{H} which is therefore of dimension 4.

We finally define

$$q^* = {}^t\bar{q} = \begin{pmatrix} \bar{a} & b \\ -\bar{b} & a \end{pmatrix} = q = t\mathbf{1} - t_{\mathbf{I}}\mathbf{I} - t_{\mathbf{J}}\mathbf{J} - t_{\mathbf{K}}\mathbf{K}$$

and

$$\|q\| = \sqrt{|a|^2 + |b|^2} = \sqrt{\det(q)} \in \mathbf{R}.$$

In the following, we will identify \mathbf{R} with $\mathbf{R}\mathbf{1}$ which is clearly in the center of \mathbf{H} and therefore will write t for $t\mathbf{1}$ for any $t \in \mathbf{R}$. By construction, $q \in \mathbf{R}$ if and only if $q = q^*$.

Proposition 14.2.0.1. *With the above notations, we have*

1. \mathbf{H} is a four dimension \mathbf{R} -subalgebra of $M_2(\mathbf{C})$ which is a skew-field with neutral element $\mathbf{1}$: \mathbf{H} a non-commutative field.
2. One has $qq^* = q^*q = \|q\|^2$ and the inverse of $q \in \mathbf{H} \setminus \{0\}$ is $q^*/\|q\|^2$.
3. $\|\cdot\|$ is an Euclidean norm on \mathbf{H} which is compatible with the product. Moreover, the scalar product is given by $\langle q_1, q_2 \rangle = \frac{1}{2}(q_1q_2^* + q_1^*q_2) \in \mathbf{R}$.
4. The conjugation map $q \mapsto q^*$ is an anti-involution of skew-fields, i.e. a linear symmetry switching the order of product $(q_1q_2)^* = q_2^*q_1^*$.
5. The center of \mathbf{H} is \mathbf{R} .

Proof. The formula

$$q(a, b)q(c, d) = q(ac - \bar{b}d, bc + a\bar{d}) \in \mathbf{H}, \quad \forall a, b, c, d \in \mathbf{C}$$

proves that \mathbf{H} is a subalgebra. A direct computation gives

$$(i) \quad \mathbf{I}^2 = \mathbf{J}^2 = \mathbf{K}^2 = \mathbf{I}\mathbf{J}\mathbf{K} = -\mathbf{1}$$

which implies for instance $\mathbf{I}\mathbf{J} = \mathbf{K} = -\mathbf{J}\mathbf{I}$ (and the analogous expressions by cyclic permutations of the indices). More generally, we get

\nearrow	1	I	J	K
1	1	I	J	K
I	I	-1	K	-J
J	J	-K	-1	I
K	K	J	-I	-1

Quaternion Multiplication Table

By direct computation (or Cramer's formula in dimension 2), we get $qq^* = q^*q = N(q)$. Remembering the polarization formula for scalar products, we have proved the first two items.

The formula $\|(a, b)\| = \sqrt{|a|^2 + |b|^2}$ is the definition of the standard norm on \mathbf{C}^2 and the formula $\|q\|^2 = \det(q)$ gives the product compatibility, hence the third item

The fourth item is the usual property of adjoint complex matrices.

For the last one, if q commutes in \mathbf{I} , because $\mathbf{I}, \mathbf{J}, \mathbf{K}$ anti-commutes pairwise, one gets $t_{\mathbf{I}}, t_{\mathbf{J}}, t_{\mathbf{K}} = 0$. \square

14.3 Imaginary Quaternions

By analogy to the complex case, there is two equivalent ways to define a real quaternion q : either $q \in \mathbf{R}\mathbf{1}$, either q is invariant by conjugation. Pursuing this analogy, we'll say that q is imaginary if $q^* = -q$. Because q is a linear symmetry, one therefore has a canonical eigenspace decomposition $q = q_+q_-$ into a sum of a real quaternion $q_+ \in \mathbf{R}\mathbf{1}$ and an imaginary q_- . Notice that the conjugation is an isometry because $\|q^*\| = \sqrt{q^*q^{**}} = \sqrt{q^*q} = \|q\|$ proving that we indeed have an orthogonal decomposition

$$\mathbf{H} = \mathbf{R}\mathbf{1} \oplus \mathbf{H}_-$$

Of course, the space of imaginary quaternions is the 3-dimensional vector space $\mathbf{H}_- = \text{Span}(\mathbf{I}, \mathbf{J}, \mathbf{K})$. It is not completely obvious however that this decomposition is canonical, namely depends only of the field structure. This indeed the case.

Lemma 14.3.0.1. *One has $\mathbf{H}_- = \{q \in \mathbf{H} \mid q^2 \in \mathbf{R}_- \mathbf{1}\}$.*

Proof. With the notations above, one computes

$$q^2 = (t^2 - t_{\mathbf{I}}^2 - t_{\mathbf{J}}^2 - t_{\mathbf{K}}^2) \mathbf{1} + (tt_{\mathbf{I}} \mathbf{I} + tt_{\mathbf{J}} \mathbf{J} + tt_{\mathbf{K}} \mathbf{K}).$$

If $q \in \mathbf{H}_-$, then $t = 0$ and $q^2 = (-t_{\mathbf{I}}^2 - t_{\mathbf{J}}^2 - t_{\mathbf{K}}^2) \mathbf{1} \in \mathbf{R}_- \mathbf{1}$.

Conversely, $q \notin \mathbf{H}_-$, then $t \neq 0$. But, if moreover q^2 is real, we have $tt_i = 0$ for $i > 0$ and therefore $t_i = 0$. This would imply $q^2 = t^2 \mathbf{1} \in]0, \infty[\mathbf{1}$, a contradiction. \square

Corollary 14.3.0.2. *The conjugation is the unique anti-involution σ of the skew-field \mathbf{H} such that $\text{Ker}(\sigma - \text{Id}) = \mathbf{R}\mathbf{1}$. In particular, the norm defined on \mathbf{H} in 14.2.0.1 does only depend on the field structure.*

Proof. By assumption, σ induces the identity on the real numbers. Therefore, σ is a linear symmetry fixing the line $\mathbf{R}\mathbf{1}$.

But, if $q \in \mathbf{H}_-$, we have q^2 real and therefore $\sigma(q^2) = q^2$. Because $q \in \mathbf{H}_-$, one has moreover $q^2 \in \mathbf{R}_- \mathbf{1}$ hence $\sigma(q^2) = q^2 \in \mathbf{R}_- \mathbf{1}$. But σ is an anti-morphism, therefore $\sigma(q^2) = \sigma(q)^2 \in \mathbf{R}_- \mathbf{1}$ hence $\sigma(q) \in \mathbf{H}_-$ by (14.3.0.1). We have proved $\sigma(\mathbf{H}_-) \subset \mathbf{H}_-$. If σ is not the conjugation, the restriction of σ to \mathbf{H}_- is not $-\text{Id}$. Because it is still a symmetry, this restriction would have a nonzero fixed point in \mathbf{H}_- , contradicting $\text{Ker}(\sigma - \text{Id}) = \mathbf{R}\mathbf{1}$. \square

14.3.1 Quaternions and $\text{SO}_3(\mathbf{R})$

Let us recall that \mathbf{H} has a canonical structure of four dimensional space Euclidean space (14.3.0.2) and let $\mathbf{S}_3 = \{q \in \mathbf{H} \mid \|q\| = 1\}$ be its unit sphere of our Euclidean. Because the norm is compatible with the product, It subgroup of the multiplicative group \mathbf{H}^* which is is a compact connected like any non Euclidean sphere (in dimension > 1). We keep in mind that $q^* = q^{-1}$ for $q \in \mathbf{S}^3$.

Proposition 14.3.1.1 (Hamilton, 1844). *Let $q = q_+ + q_- \in \mathbf{S}_3$.*

1. *The map $x \mapsto qxq^{-1} = qxq^*$ induces a rotation $\rho(q) \in \text{SO}(\mathbf{H}_-)$.*
2. *ρ defines a morphism $\mathbf{S}_3 \rightarrow \text{SO}(\mathbf{H}_-)$ of kernel $\{\pm \mathbf{1}\}$.*
3. *If $q \neq \pm \mathbf{1}$, the rotation $\rho(q)$ has axis $\mathbf{R}q_-$ and angle¹ $\theta = \pm 2 \arccos(q_+)$.*
4. *ρ is onto and defines an isomorphism $\mathbf{S}_3/\{\pm \mathbf{1}\} \simeq \text{SO}(\mathbf{H}_-) = \text{SO}_3(\mathbf{R})$.*

Proof. We have $\|qxq^*\| = \|q\|\|x\|\|q^*\| = \|x\|$ because $\|q\| = \|q^*\| = 1$ and therefore $x \mapsto qxq^*$ is a (linear) isometry of \mathbf{H} . Because $\mathbf{R}\mathbf{1}$ is invariant, so is its orthogonal \mathbf{H}_- implying that it induces an isometry $\rho(q) \in \text{O}(\mathbf{H}_-)$. The composite of ρ with the determinant defines a continuous map $\mathbf{S}_3 \rightarrow \{\pm 1\}$ which is constant by connectedness of \mathbf{S}_3 . Because $\rho(\mathbf{1}) = 1$, all the $\rho(q)$ are of determinant 1 hence are rotations. For the second item, ρ being defined by an interior homomorphism, ρ is a morphism. An element q in the kernel of ρ in the kernel commutes with any imaginary element and therefore with any quaternion

¹Recall that $\theta \bmod 2\pi$ is defined up to sign because we do not have chosen an orientation of \mathbf{H}_- (12.6.2.6). The interested reader will compute θ as oriented angle for instance orienting \mathbf{H} by $\mathcal{B}_- = (\mathbf{I}, \mathbf{J}, \mathbf{K})$ and the axis by q_- (exercise).

because $\mathbf{1}$ is central. It follows that q is real and therefore $q = \pm \mathbf{1}$ because it is of norm 1.

For the third item, we first notice that $\rho(q) \neq \text{Id}$ and therefore the axis is well defined. Because, q_+ is in the center, $q_- = q - q_+$ commutes with q and $\rho(q)(q_-) = q_- qq^* = q_-$. Moreover, q_- is non zero because $q \neq \pm \mathbf{1}$ and q_- is a basis of the axis.

Let us compute the trace $\text{tr}(\rho(q)) = 1 + 2 \cos(\theta)$ (12.6.2.6). In the basis $(\mathbf{I}, \mathbf{J}, \mathbf{K})$, the (\mathbf{I}, \mathbf{I}) coefficient is the real factor of \mathbf{I} in $q \mathbf{I} q^*$ which is $t_{\mathbf{I}}^2 - t^2 - t_{\mathbf{J}}^2 - t_{\mathbf{K}}^2$, the other two coefficients being obtained by cyclic permutation (cf/ i). Summing the three terms, we get

$$1 + 2 \cos(\theta) = (t_{\mathbf{I}}^2 - t^2 - t_{\mathbf{J}}^2 - t_{\mathbf{K}}^2 + t_{\mathbf{J}}^2) + (t^2 - t_{\mathbf{K}}^2 - t_{\mathbf{I}}^2 + t_{\mathbf{K}}^2) = 2t^2 = 4a^2 - 1$$

using $1 = \|q\|^2 = t^2 + t_{\mathbf{I}}^2 + t_{\mathbf{J}}^2 + t_{\mathbf{K}}^2$ or equivalently $\cos^2(\theta/2) = a^2$.

For the last item, let $r \neq \pm id$ be a rotation of \mathbf{H}_- of angle θ (defined up to sign). Let e_- be one of two unit vector of its axis and $q = \cos(\theta/2) + \sin(\theta/2)e_- \in \mathbf{S}_3$. Then, the preceding computation shows that $\rho(q)$ and r have same (non oriented) angle and axis and therefore $\rho(q) = r$ or $\rho(q^*) = r$. \square

14.3.2 Cross product and Rodrigues formula, tbd

14.3.3 Quaternions and $\text{SU}_2(\mathbf{R})$

By the very definition of \mathbf{H} we have chosen (14.2.0.1), we have

$$\mathbf{S}_3 = \{q(a, b) \mid \det(q(a, b)) = 1, a, b \in \mathbf{C}\}$$

which is an equality of groups. Because $q(a, b) = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$, $a, b \in \mathbf{C}$ we have $\mathbf{S}_3 = \text{SU}_2(\mathbf{C})$.

Corollary 14.3.3.1. *We have the equality of groups $\mathbf{S}_3 = \text{SU}_2(\mathbf{C})$ and a canonical continuous isomorphism of compact groups $\text{SU}_2(\mathbf{C})/\{\pm \text{Id}\} \simeq \text{SO}_3(\mathbf{R})$.*

Exercise(s) 14.3.3.2. *SU_2 et représentation adjointe. TBD.*

14.3.4 Quaternions and $\text{SO}_4(\mathbf{R})$

Proposition 14.3.4.1 (Cayley, 1855). *Let $(q_1, q_2) \in_3 \times \mathbf{S}_3$.*

1. *The map $x \mapsto q_1 x q_2^*$ induces a rotation $\tilde{\rho}(q) \in \text{SO}(\mathbf{H})$.*
2. *$\tilde{\rho}$ defines a morphism $\mathbf{S}_3 \times \mathbf{S}_3 \rightarrow \text{SO}(\mathbf{H})$ of kernel $\{\pm(\mathbf{1}, \mathbf{1})\}$.*
3. *$\tilde{\rho}$ is onto and defines an isomorphism $(\mathbf{S}_3 \times \mathbf{S}_3)/\{\pm(\mathbf{1}, \mathbf{1})\} \simeq \text{SO}(\mathbf{H}) = \text{SO}_4(\mathbf{R})$. In particular, $\text{SO}(4, \mathbf{R})$ is not simple (compare with (13.7.0.6) and (13.9.0.2)).*

Proof. The proof of the first two items goes exactly as the analogous assertions in (12.6.2.6) and the identity

$$\tilde{\rho}((\tilde{q}_1, \tilde{q}_2)) \circ \tilde{\rho}((q_1, q_2))(q) = \tilde{\rho}((\tilde{q}_1, \tilde{q}_2))(q_1 q q_2^*) = \tilde{q}_1 q_1 q q_2^* \tilde{q}_2^* = (\tilde{q}_1 q_1) q (\tilde{q}_2 q_2)^* = \tilde{\rho}((\tilde{q}_1 q_1, \tilde{q}_2 q_2))(q).$$

The only remaining point is the surjectivity of $\tilde{\rho}$. Let $r \in \text{SO}(\mathbf{H})$ and $q = r(\mathbf{1})$. Because r is an isometry, $q \in \mathbf{S}$ and $(\tilde{\rho}((q^{-1}, \mathbf{1})) \circ r)(\mathbf{1}) = \mathbf{1}$. The image of $\tilde{\rho}$ being a subgroup of $\text{SO}_3(\mathbf{R})$, one can assume $r(\mathbf{1}) = \mathbf{1}$ which implies $r(\mathbf{1}^\perp) = \mathbf{1}^\perp = \mathbf{H}_-$. The restriction \tilde{r} of r to \mathbf{H}_- is an isometry and its determinant is 1 because $\det(r) = 1$ and $r(\mathbf{1}) = \mathbf{1}$. In other words, $\tilde{r} = \rho(\tilde{q})$ for some $\tilde{q} \in \text{S}_3$ implying $r = \tilde{\rho}(\tilde{q}, \tilde{q})$. \square

14.4 Spin

TBD.

14.5 Additionnal Exercises

Exercise(s) 14.5.0.1. *Frobenius TBD*

Exercise(s) 14.5.0.2. *Groupe d'ordres 8. TBD*

Exercise(s) 14.5.0.3. *Rotations isoclines. TBD*

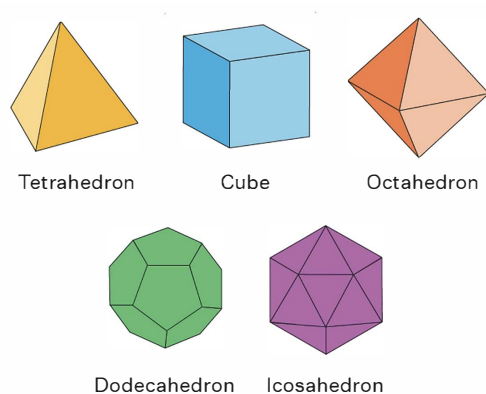
Exercise(s) 14.5.0.4. *Flip SO_4 non int erieur.*

Exercise(s) 14.5.0.5.

$$\mathbf{Rot}_{\mathbf{u}, \theta}(\mathbf{x}) = \cos \theta \mathbf{x} + (1 - \cos \theta)(\mathbf{u} \cdot \mathbf{x})\mathbf{u} + \sin \theta (\mathbf{u} \wedge \mathbf{x}).$$

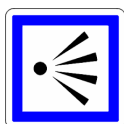
Chapter 15

Finite Groups of Euclidean Isometries in Dimension ≤ 3



The five Platonic Solids

15.1 Perspective

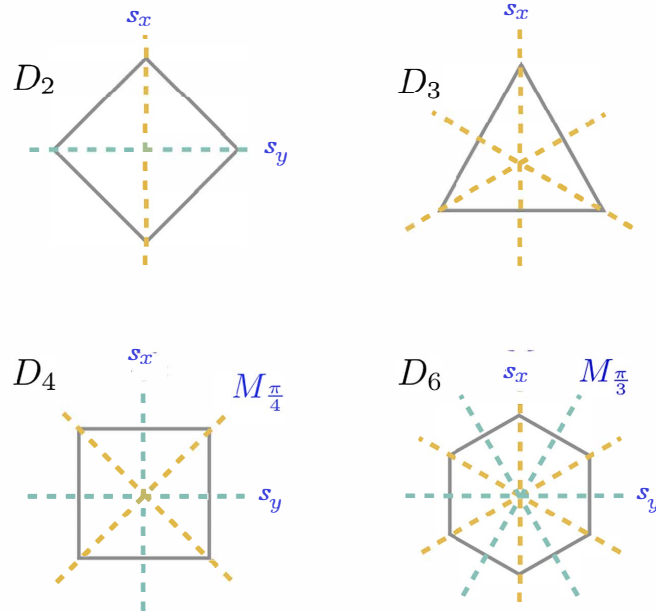


We are interested in classifying the finite subgroups of $O_2(\mathbf{R})$ and $O_3(\mathbf{R})$ and their relationships with regular polytopes.

15.2 Subgroups of $O_2(\mathbf{R})$

We first consider $O_2(\mathbf{R})$ before passing to $SO_3(\mathbf{R})$.

Proposition 15.2.0.1. *Let G be a non-trivial finite subgroup of $O_2(\mathbf{R})$. Then G preserves a regular n -gon. If G consists only of rotations, then G is isomorphic to $\mathbf{Z}/n\mathbf{Z}$. Otherwise, G is isomorphic to a dihedral group D_n .*



Dihedral group

Proof. If G consists only of rotations, then G identifies as a finite subgroup of \mathbf{R}/\mathbf{Z} . We denote $p : \mathbf{R} \rightarrow \mathbf{R}/\mathbf{Z}$ as the canonical projection. Then $p^{-1}(G)$ is a subgroup of \mathbf{R} containing the integers. Since G is finite, $p^{-1}(G)$ is discrete, hence it is $(1/n)\mathbf{Z}$ for some integer $n \geq 1$. Therefore, G is isomorphic to $\mathbf{Z}/n\mathbf{Z}$. We deduce that G preserves a regular n -gon inscribed in the unit disk.

Otherwise, let $SG = G \cap SO_2(\mathbf{R})$. As stated below, SG is isomorphic to $\mathbf{Z}/n\mathbf{Z}$. Moreover, we have the short exact sequence

$$1 \rightarrow SG \rightarrow G \rightarrow \mathbf{Z}/2\mathbf{Z} \rightarrow 1.$$

Thus, G has order $2n$.

Let σ be a symmetry in G . It fixes two opposite points x and $-x$ on the unit circle \mathbf{S}^1 . We then write $\sigma = \sigma_x = \sigma_{-x}$. We denote X as the set of fixed points of all the symmetries in $G \setminus SG$. We have n symmetries thus $2n$ fixed points.

Furthermore, for $g \in G$,

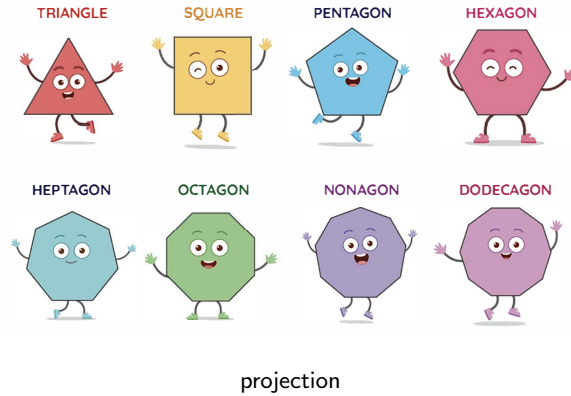
$$g \circ \sigma_x \circ g^{-1}(g(x)) = g(x)$$

hence

$$g \circ \sigma_x \circ g^{-1} = \sigma_{g(x)}$$

and G operates on X . The stabilizer of each point in X is of order two, so each orbit is of cardinality n . Thus, we have two orbits under the action of SG .

We deduce that G is isomorphic to D_n . □



15.3 Subgroups of $SO_3(\mathbf{R})$

Let G be a subgroup of $SO_3(\mathbf{R})$ of order N . Each non-trivial element is a rotation, thus fixes two opposite points on the sphere \mathbf{S}^2 . We denote $X = X_G$ as the set of these points. As above, G operates on X .

The stabilizer of each point x in X fixes the orthogonal plane x^\perp . Its restriction is a finite subgroup of $SO_2(\mathbf{R})$, hence isomorphic to $\mathbf{Z}/r_x\mathbf{Z}$. Its orbit is thus of order $n_x = N/r_x$. Each point x is the fixed point of $(r_x - 1)$ rotations, being different for all other points except its opposite. Thus, we have

$$2N - 2 = \sum_{x \in X} (r_x - 1) = \sum_{j \in X/G} n_j (r_j - 1).$$

We deduce that

$$2 - 2/N = \sum_{j \in X/G} (1 - 1/r_j).$$

Since $r_j \geq 2$ by definition, hence

$$2 > 2 - 2/N = \sum_{j \in X/G} (1 - 1/r_j) \geq |X/G|/2.$$

Consequently, there are at most three orbits. Moreover, the group G does not operate transitively on X . Indeed, we would have $2 - 2/N = 1 - 1/r$, thus $1 = 2/N - 1/r \leq 1/N$, since $r \leq N$!!

Proposition 15.3.0.1. *If there are two orbits, then G is a group of plane rotations, isomorphic to $\mathbf{Z}/N\mathbf{Z}$.*

Proof. We have $2 - 2/N = 2 - (1/r_1 + 1/r_2)$ hence $2/N = 1/r_1 + 1/r_2$. If $N = r_1$ then $r_2 = N$. Thus, X has two elements, and G fixes their orthogonal. Consequently, it operates as a subgroup of $SO_2(\mathbf{R})$ and it is isomorphic to $\mathbf{Z}/N\mathbf{Z}$.

Otherwise, we have $2r_1 \leq N$, hence $(1/r_1) \geq 2/N$ so $r_2 \leq 0$!! \square

The case of three orbits involves several cases. Our equation becomes

$$\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} = 1 + \frac{2}{N},$$

where we choose $r_1 \leq r_2 \leq r_3$.

If $r_1 \geq 3$, then the left term is smaller than 1 while the second is strictly greater. Thus $r_1 = 2$.

If $r_2 = 2$, then $1/r_3 = 2/N$, thus $N = 2r_3$. We then have $r_j = (2, 2, N/2)$ and $n_j = (N/2, N/2, 2)$.

If $r_2 \geq 3$ then $1/r_3 = 1/2 + 2/N - 1/r_2 > 1/2 - 1/3$, so $r_3 < 6$.

- If $r_j = (2, 3, 3)$ then $N = 12$ and $n_j = (6, 4, 4)$.
- If $r_j = (2, 3, 4)$ then $N = 24$ and $n_j = (12, 8, 6)$.
- If $r_j = (2, 3, 5)$ then $N = 60$ and $n_j = (30, 20, 12)$.

If $r_2 \geq 4$, then

$$\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} \leq 1 < 1 + \frac{2}{N},$$

thus we have the complete list.

Proposition 15.3.0.2. *G is a dihedral group D_r in the case*

$$N = 2r \quad r_j = (2, 2, r) \quad n_j = (r, r, 2).$$

Proof. Opposite points behave the same so the third orbit consists of two opposite points. Consequently, G fixes their orthogonal, and the isometries that do not fix these points are symmetries on this plane. Thus, it is the dihedral group. \square

If we are not in one of the previous cases, then no orbit is planar. Indeed, restricting to the plane would bring us back to the already treated cases.

Proposition 15.3.0.3. *If $N = 12$, $r_j = (2, 3, 3)$, and $n_j = (6, 4, 4)$ then G is the group of isometries of a tetrahedron, and is isomorphic to \mathfrak{a}_4 .*

Proof. Let $x \in o(2)$. Its stabilizer is a group of rotations that operates on its orbit, thus on three points. These points form an equilateral triangle. This implies a tetrahedron. Moreover, G operates on these vertices: it is identified as a subgroup of the permutation group of 4 elements \mathfrak{S}_4 . Since the only endomorphism that fixes these four non-coplanar points is the identity. Thus, G is a subgroup of index 2: it is \mathfrak{a}_4 . \square

Proposition 15.3.0.4. *If $N = 24$, $r_j = (2, 3, 4)$, and $n_j = (12, 8, 6)$ then G is the group of isometries of a cube and an octahedron, and is isomorphic to \mathfrak{S}_4 .*

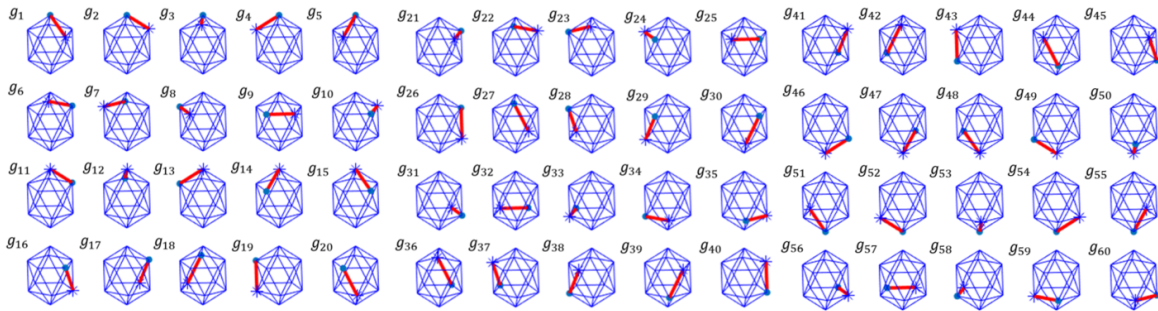
Proof. The stabilizer of a point $x \in o(3)$ operates on $o(2)$, in two orbits. Each forms a square, and these two squares cannot be coplanar. Changing the point of $o(3)$, we see that $o(2)$ are the vertices of a cube whose faces are in the direction of the points of $o(3)$ and the edges of $o(1)$.

Moreover, G operates on the pairs of opposite vertices, hence we have a morphism $\varphi : G \rightarrow \mathfrak{S}_4$. If $\varphi(g)$ is the identity, and g exchanges two vertices, then, since there are only two fixed points, g exchanges at least two other pairs. Consequently, $g = -\text{Id}$, but g is a rotation, hence this is impossible, and φ is injective.

By duality, the convex hull of $o(3)$ is an octahedron with faces centered on $o(2)$. □

We note that the cube contains two “opposite” tetrahedra, considering as edges the diagonals of the faces. The group either preserves these tetrahedra or exchanges them. From this, we can deduce that \mathfrak{a}_4 is a subgroup of index 2 in G , thus normal.

Proposition 15.3.0.5. *If $N = 60$, $r_j = (2, 3, 5)$, and $n_j = (30, 20, 12)$ then G is the group of isometries of a dodecahedron and an icosahedron, and is isomorphic to \mathfrak{a}_5 .*



The 60 rotations of an icosaedron¹

Proof. The three orbits are organized in pairs.

The stabilizer of a point $x \in o(3)$ operates on $o(2)$, in four orbits of five elements each. The closest points in $o(2)$ to x form a pentagon (we cannot have two orbits on the same plane considering another point of $o(3)$). Operating across the entire orbit of x , one can envision a dodecahedron with 20 vertices and 30 edges. By duality, an icosahedron is obtained.

¹We denote the identity by g_1 , then we highlight one edge, and show how each rotation $g_i \in G$ transforms the highlighted edge, cf. C. Esteves, Y. Xu, C. Allec-Blanchette and K. Daniilidis, , 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, doi: 10.1109/ICCV.2019.00165.

We fix a vertex $x \in o(2)$. It corresponds to a common vertex of three pentagons. The stabilizer of x permutes them, as well as the common edges. It also operates on the other vertices of these pentagons, in two orbits. The three segments joining x to one of these orbits intersect at right angles.

This suggests a cube if we consider $(-x)$. The centers of the faces correspond to points of $o(1)$, namely the centers of the edges. It can be seen that each edge is of the same length, and that there are indeed 8 vertices.

For each pair of vertices corresponds two cubes, and each cube admits 8 vertices, thus 4 diagonals. Therefore, we obtain $10 \times 2/4 = 5$ cubes. Equivalently, each edge corresponds to a face of a cube, which has 6 such faces. This gives us $30/6 = 5$ cubes.

Our group G operates on these cubes by permutation. Suppose an element $g \in G$ fixes all these cubes globally. If it is non-trivial, then it must fix its axis of rotation. If its order is 2, it fixes an edge of the dodecahedron, thus a face of a cube. It cannot fix the other cubes then. If its order is 3, then its axis passes through the vertices of two cubes. The cubes induced by contiguous vertices must also be preserved by g , which is impossible. The order cannot be 5 since no element that preserves a cube is of an order multiple of 5.

Therefore, $g = \text{Id}$. Our group is thus a subgroup of \mathfrak{S}_5 with index 2 (for a question of order), it is \mathfrak{a}_5 . \square

15.3.1 Normal Subgroups

If G is a finite subgroup of $\text{SO}_3(\mathbf{R})$, and if H is a subgroup of G , then $X_H \subset X_G$. If $x \in X_H$, $g \in G$, then there exists $h \in H$ such that $h(x) = x$, and $ghg^{-1}(gx) = gx$. Thus, the action of G induced by inner automorphisms on H corresponds with the action of G on the images of X_H .

Furthermore, if H is normal, then X_H is preserved. Thus, G operates on X_H , and each point of X_H has an orbit by H which is a sub-orbit for G . In other words, $o_H(x)$ divides $o_G(x)$, and X_H is a union of orbits of G . Thus, the orbits within X_H of G are decomposed into orbits of H . This easily demonstrates that \mathfrak{a}_5 is simple.

Remark(s) 15.3.1.1. *A convex polytope is an intersection of half-spaces with non-empty interiors. If P is a convex polytope, it can be assumed that the origin is within its interior. If we project the edge of P onto \mathbf{S}^2 , then a triangulation of the sphere in vertices, edges, and faces is obtained. If we denote s as the number of vertices, a as the number of edges, and f as the number of faces, then $s - a + f = 2$. Indeed, each time a vertex with the edges containing it is removed, as many faces as edges are removed, except that the vertex turns into a face: $s - a + f$ remains constant as the number of vertices decreases. When only 4 vertices remain, then the formula can be verified.*

15.4 Appendix: Local Extrema and the Position of a Hypersurface Relative to Its Tangent Plane

If $f : \mathbf{R}^n \rightarrow \mathbf{R}$ is a class \mathcal{C}^2 function, then Schwarz's theorem implies that the matrix of second partial derivatives is symmetric. This allows us to apply the preceding discussion to the calculus of variations.

Morse Lemma. — *Let $f : \mathbf{R}^n \rightarrow \mathbf{R}$ be a class \mathcal{C}^k function, $k \geq 2$ such that $f(0) = D_0f = 0$ and D_0^2f is invertible. Then there exists a neighborhood V of the origin and class \mathcal{C}^{k-2} functions y_1, \dots, y_r and z_1, \dots, z_s defined on V such that $r + s = n$ and, for $x \in V$, one has*

$$f(x) = \sum_{1 \leq j \leq r} y_j^2 - \sum_{1 \leq j \leq s} z_j^2.$$

PROOF. — Consider the integral remainder development of f near the origin. We have

$$f(x) = \int_0^1 (1-t) D_{tx}^2 f(x, x) dt = {}^tX \cdot \int_0^1 (1-t) D_{tx}^2 f dt \cdot X.$$

We denote

$$A(x) = \left(\int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(tx) dt \right)_{i,j} \quad \text{and} \quad A_0 = A(0).$$

Then we use the following lemma.

Lemma 15.4.0.1. *If $A_0 \in \mathcal{S}_n(\mathbf{R}) \cap \text{GL}_n(\mathbf{R})$, there exists a neighborhood U of A_0 in $\mathcal{S}_n(\mathbf{R})$ and a smooth function $\psi : U \rightarrow \text{M}_n(\mathbf{R})$ such that $\psi(A_0) = I$ and, for any $A \in U$,*

$${}^t\psi(A) \cdot A_0 \cdot \psi(A) = A.$$

Thus, if x is sufficiently close to the origin, then $f(x) = {}^t_x \psi(A(x)) A_0 \psi(A(x)) x$. We set $\psi_1(x) = \psi(A(x)) x$, and we obtain

$$f(x) = {}^t\psi_1(x) A_0 \psi_1(x).$$

There exists a basis P in which A_0 is diagonal with r values on the diagonal equal to 1 and s equal to -1 . We denote J this matrix and set $\psi_2 = P \cdot \psi_1$. It follows $f(x) = {}^t\psi_2(x) J \psi_2(x)$. If we call y_1, \dots, y_r and z_1, \dots, z_s the coordinates of ψ_2 , we obtain the sought form.

This establishes the Morse lemma modulo Lemma 15.4.0.1.

PROOF OF LEMMA 15.4.0.1. — We consider the function $h : \text{M}_n(\mathbf{R}) \rightarrow \mathcal{S}_n(\mathbf{R})$ defined by $h(M) = {}^tMA_0M$. We compute the differential at the identity of h .

$$h(I + M) = {}^t(I + M)A_0(I + M) = A_0 + ({}^tMA_0 + A_0M) + {}^tMA_0M = A_0 + ({}^tMA_0 + A_0M) + O(\|M\|^2).$$

Thus $D_1h(M) = {}^tMA_0 + A_0M$. This function is not invertible since $\dim \mathcal{S}_n(\mathbf{R}) < \dim M_n(\mathbf{R})$. However, the kernel consists of the matrices M for which A_0M is antisymmetric. We consider the space E of matrices M for which A_0M is symmetric. This space is complementary to $\text{Ker } D_1h$, and the restriction of D_1h to E now becomes invertible (injective since A_0 is invertible and source and target spaces are of the same dimension).

The local inversion theorem applied to $h|_E$ shows that there exist neighborhoods U of A_0 and V of I and a smooth diffeomorphism $\psi : U \rightarrow V$ that inverts $h|_E$.

Corollary 15.4.0.2. *Under these assumptions, 0 is a strict local maximum of f if and only if D^2f is negatively defined, and is a strict local minimum of f if and only if D^2f is positively defined.*

Corollary 15.4.0.3. *In \mathbf{R}^{n+1} we study the hypersurface \mathcal{S} defined by $x_{n+1} = F(x_1, \dots, x_n)$ where F is a class $\mathcal{C}^k(\mathbf{R}^n)$ function, $k \geq 2$, and D^2F is non-degenerate. The tangent plane at the point $p = (x_0, F(x_0))$ locally separates \mathcal{S} from a half-space if and only if $D_{x_0}^2F$ is defined.*

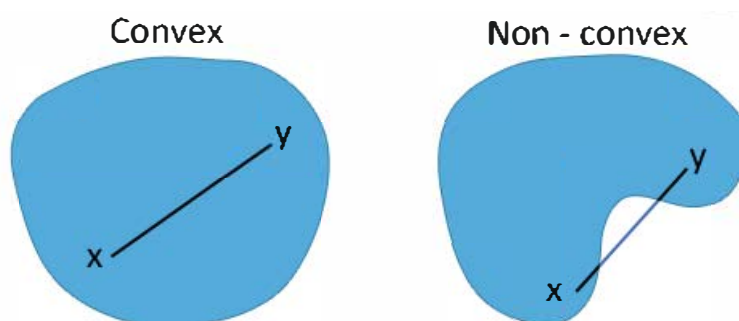
Proof. We consider the function $g(x) = F(x) - (F(x_0) + D_{x_0}F(x - x_0))$. This function satisfies the conditions of the Morse lemma. We deduce that

$$F(x) = F(x_0) + D_{x_0}F(x - x_0) + \sum_{1 \leq j \leq r} y_j^2 - \sum_{1 \leq j \leq s} z_j^2.$$

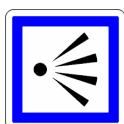
By a change of variables (by translation), we may assume that $x = F(x) = 0$. We consider a basis of $\text{Ker } D_0F$ which we complete into a basis of \mathbf{R}^{n+1} . In this basis, $D_0F = 0$. The Morse lemma allows us to conclude. \square

Chapter 16

Convexity in Euclidean Spaces



16.1 Perspective



In this chapter, E will denote a dimension $d < \infty$ real vector space. The reader will check that results with a ∞ label, like all results of section 16.2, remain valid even if $d = \infty$. We have chosen to give most of the proofs in the Euclidean case even they are not dependent of any scalar product in order to have a more geometrical intuition. The interested reader will give alternative proofs not using any scalar product in these cases.

16.2 Generalities

16.2.1 Definitions

Definition 16.2.1.1. *A subset of E is convex if for any $c, c' \in C$ the line segment $[c, c'] = \{tc + (1-t)c', t \in [0, 1]\}$ is contained in C .*

An alternative equivalent way to phrase convexity would be to demand that for every line $\ell \subset E$ the intersection $C \cap \ell$ be connected.

Remark(s) 16.2.1.2. *An immediate consequence of the definition is the following. For any family $(C_i)_{i \in I}$ of convex sets, the intersection $\bigcap_{i \in I} C_i$ is convex.*

Proposition 16.2.1.3 (∞). *A set $C \subseteq E$ is convex if and only if for all $n \geq 1, c_1, \dots, c_n \in C$ and $\lambda_1, \dots, \lambda_n \geq 0$ with $\sum_{i=1}^n \lambda_i = 1$, one has $\sum_{i=1}^n \lambda_i c_i \in C$.*

Proof. “ \Leftarrow ”: Obvious with $n = 2$.

“ \Rightarrow ”: Induction on n . For $n = 1$ the statement is trivial. For $n \geq 2$ let $c_i \in C$ for all i . (Simply omit those points whose coefficient is zero.) We need to show that

$$\sum_{i=1}^n \lambda_i c_i \in C.$$

Define $\lambda = \sum_{i=1}^{n-1} \lambda_i$ and for $1 \leq i \leq n-1$ set $\mu_i = \lambda_i/\lambda$. Observe that $\mu_i \geq 0$ and $\sum_{i=1}^{n-1} \mu_i = 1$. By the inductive hypothesis, $c' := \sum_{i=1}^{n-1} \mu_i c_i \in C$ and thus by convexity of C also $\lambda c' + (1-\lambda)c_n \in C$. We conclude by noting that $\lambda c' + (1-\lambda)c_n = \sum_{i=1}^n \lambda_i c_i$. \square

16.2.2 Convex hull

Definition 16.2.2.1. *Let P be a subset of E .*

1. *The convex hull $\text{conv}(P)$ of P is the intersection of all convex supersets of P .*
2. *A convex combination (of length $\leq n$) of P is any element of E of the form $\sum_{i=1}^n \lambda_i p_i$ with $n \geq 1$, $p_i \in P$, $\lambda_i \in \mathbf{R}^+$ such that $\sum_{i=1}^n \lambda_i = 1$.*

By (16.2.1.2), $\text{conv}(P)$ is the smallest convex subset of E containing P . Hopefully, it can be computed quite explicitly thanks the following result.

Proposition 16.2.2.2 (∞). *For any $P \subset E$, the convex hull $\text{conv}(P)$ is the set of all convex combinations of P .*

Proof. “ \subseteq ”: Consider a convex set $C \supseteq P$. By Proposition 3.3 (only-if direction) the right-hand side is contained in C . As C was arbitrary, the claim follows.

“ \supseteq ”: Denote the set on the right-hand side by R . Clearly $R \supseteq P$. We show that R forms a convex set. Let $p = \sum_{i=1}^n \lambda_i p_i$ and $q = \sum_{i=1}^n \mu_i p_i$ be two convex combinations. (We may suppose that both p and q are expressed over the same p_i by possibly adding some terms with a coefficient of zero.) Then for $\lambda \in [0, 1]$ we have $\lambda p + (1 - \lambda)q = \sum_{i=1}^n (\lambda \lambda_i + (1 - \lambda)\mu_i)p_i \in R$ as $\lambda \lambda_i + (1 - \lambda)\mu_i \geq 0$ for all $1 \leq i \leq n$ and $\sum_{i=1}^n (\lambda \lambda_i + (1 - \lambda)\mu_i) = \lambda + (1 - \lambda) = 1$. \square

Theorem 16.2.2.3 (Carathéodory). *Let P be a subset of E . Then, $\text{conv}(P)$ is the set of convex combination of P of length at most $d + 1$. In particular, if P is compact, then so is $\text{conv}(P)$.*

Proof. If the assertion is false, there exists $n \geq d + 1$ and an convex combination c of P of length $\leq n + 1$ which is not a convex combination of length $\leq d + 1$. Let us chose such an element with $n \geq d + 2$ minimal. By minimality, c can be written

$$c = \sum_{i=0}^n \lambda_i p_i$$

with $p_i \in P, \lambda_i > 0$ and $\sum \lambda_i = 1$. But, the n -vectors $p_i - p_0, i > 0$ are linked because $n > d = \dim(E) = d$: there exists real numbers $(\mu_i, i > 0)$ such that $\sum_{i>0} \mu_i (p_i - p_0) = 0$ with positive coefficient μ_j , or, setting $\mu_0 = -\sum_{i>0} \mu_i$, such that

$$\sum_{i=0}^n \mu_i p_i = 0.$$

For any non negative real t , one gets

$$c = \sum_{i=0}^n (\lambda_i - t\mu_i) p_i.$$

If t is small enough, all coefficients $\lambda_i - t\mu_i$ are positive like λ_i . Let τ be the largest real number such that all the $\lambda_i - t\mu_i$ are ≥ 0 for $t \in [0, \tau]$. Because $\lambda_j - t\mu_j < 0$ if t is large enough, τ is well-defined (and $\tau > 0$ by the remark above). By construction, we have $\lambda_i - \tau\mu_i \geq 0$ for all $i, \sum (\lambda_i - \tau\mu_i) = 1$ and there exists k such that $\lambda_k - \tau\mu_k = 0$ implying that

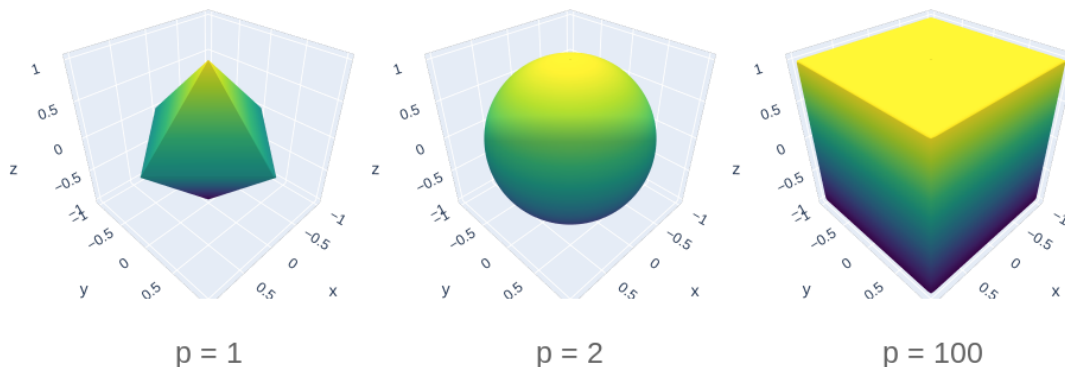
$$c = \sum_{i=0, i \neq k}^n (\lambda_i - \tau\mu_i) p_i$$

is a convex combination of length $\leq n$, a contradiction. If P is compact, let $\Delta \subset [0, 1]^{d+1}$ be the compact of elements with sum equal to 1. Then, $\text{conv}(P)$ is the continuous image of the application $\Delta \times P^{d+1} \rightarrow \text{conv}(P)$ defined by $((\lambda_i), (p_i)) \mapsto \sum \lambda_i p_i$ and is therefore compact. \square

Exercise(s) 16.2.2.4. *Let $L^2(\mathbf{N})$ be the vector space of sequence (u_n) such that $\sum u_n^2 < \infty$ with the usual L^2 -norm $\|(u_n)\| = \sqrt{\sum u_n^2}$. Let $e_k \in \ell_2$ be the sequence $(e_k)_n = \delta_{k,n}$. Show that $\{e_n/(n+1), n \geq 0\} \cup \{0\}$ but that its convex hull is non closed in $L^2(\mathbf{N})$.*

16.2.3 Topology of a convex set

In this section, E can be any normed space, no matter its dimension.



Unit ball¹ for $\| \cdot \|_p$

Let U be a neighborhood of 0 in E which is contained in C and let $x \in E$. If α is large enough, $x/\alpha \in C$. We can therefore define the gauge function of C as follows.

Definition 16.2.3.1 (Gauge function). *Let C be a closed convex bounded set with $0 \in \overset{\circ}{C}$ ² of E . The gauge function of C is the map $\rho_C : E \rightarrow \mathbf{R}^+$ defined by $\rho_C(x) = \inf\{\alpha > 0 \mid x/\alpha \in C\}$.*

Remark(s) 16.2.3.2.

- If $\rho_C(x) = 0$ then $x = 0$. In this case, one can choose a sequence (α_n) with 0 limit $x/\alpha_n \in C$. But C being bounded, this forces x to be 0 .
- If $x \neq 0$, the inf is a min because C is closed: $x/\rho_C(x) \in C$. Moreover, if $t \geq \rho_C(x)$, one has $x/t = \rho_C(x)/t(x/\rho_C(x)) + (1-t)0 \in C$. Hence, $x/t \in C$ if and only if $t \geq \rho_C(x)$.
- In particular, if $1 \geq \rho_C(x)$, we get $x \in C$. The converse is obvious : $C = \{x \in E \mid \rho_C(x) \leq 1\}$.
- For $t > 0$, one has $tx/t\alpha = x/\alpha$ proving that ρ_C is positively homogeneous: $\forall x \in E, \forall t \geq 0, \rho_C(tx) = t\rho_C(x)$.
- The convex closed set $C = \overline{B(0, r)}$, $r > 0$ is bounded (compact if $\dim(E) < \infty$) and 0 is an interior point. In this case, we have $\rho_C(x) = \|x\|/r$.

¹See by "Graphing the p-Norm Unit Ball in 3 Dimensions" by Kayden Mimmack, 2019.

²A convex body if $\dim(E) < \infty$, see (16.6.0.1).

Proposition 16.2.3.3 (∞). *The gauge function ρ_C satisfies the following properties:*

1. $\forall x, y \in E, \rho_C(x + y) \leq \rho_C(x) + \rho_C(y)$.
2. ρ_C is Lipschitz hence continuous.
3. $\{x \in E \mid \rho_C(x) < 1\} = \overset{\circ}{C}$ and $\{x \in E \mid \rho_C(x) = 1\} = \partial C$.

Proof. One can assume x and y are nonzero and let $\bar{x} = \frac{x}{\rho_C(x)}$, $\bar{y} = \frac{y}{\rho_C(y)}$.

By homogeneity of the gauge function, one gets $\rho_C(\bar{x}) = \rho_C(\bar{y}) = 1$ hence $\bar{x}, \bar{y} \in C$ (16.2.3.2).

$$\alpha = \frac{\rho_C(x)}{\rho_C(x) + \rho_C(y)}, \quad \bar{z} = \alpha\bar{x} + (1 - \alpha)\bar{y}.$$

By convexity, $\bar{z} \in C$ implying $\rho_C(\bar{z}) \leq 1$ (16.2.3.2). Reducing to the same denominator, we get

$$\bar{z} = \frac{x + y}{\rho_C(x) + \rho_C(y)}.$$

and by homogeneity we obtain $\rho_C(x + y) \leq \rho_C(x) + \rho_C(y)$ wanted in (1).

Let $r > 0$ such that $\overline{B(0, r)} \subset C$ is included in C and $0 \neq y \in E$. Then $r \frac{y}{\|y\|} \in C$ and therefore $\rho_C(y) \leq \frac{\|y\|}{r}$ by (16.2.3.2) giving the continuity at the origin.

Moreover, by (1), as the usual trick to prove the continuity of a norm, we get

$$\rho_C(x + y) \leq \rho_C(x) + \rho_C(y), \quad \rho_C(x) \leq \rho_C(x + y) + \rho_C(-y),$$

hence

$$|\rho_C(x + y) - \rho_C(x)| \leq \max(\rho_C(y), \rho_C(-y)) \leq \frac{\|y\|}{r}$$

and (2) is proved.

The continuity of ρ_C that $\{x \in E \mid \rho_C(x) < 1\}$ is open in E . If $x \in \overset{\circ}{C}$ $\rho_C(x) = 1$ with $\rho_C(x)$ would exist, for $0 < \varepsilon \ll 1$, we would have $x/(1 + \varepsilon)^{-1} = x + \varepsilon x \in C$ hence $1 = \rho_C(x) \leq (1 + \varepsilon)^{-1} < 1$, a contradiction proving

$$\overset{\circ}{C} = \{x \in E \mid \rho_C(x) < 1\}$$

hence (3) because C is closed in E . □

Remark(s) 16.2.3.4. *If C is moreover symmetrical, i.e. is invariant by $-\text{Id}$, one has $\rho(x) = \rho(-x)$ and the gauge function is a norm.*

Corollary 16.2.3.5 (∞). *Let C be a bounded convex set with a non-empty interior in a normed vector space E . Then there exists a homeomorphism $f : E \rightarrow E$ that maps C° to the open ball $B(0, 1)$ and C to the closed ball $\overline{B(0, 1)}$. In particular, in finite dimension, all convex bodies are homeomorphic.*

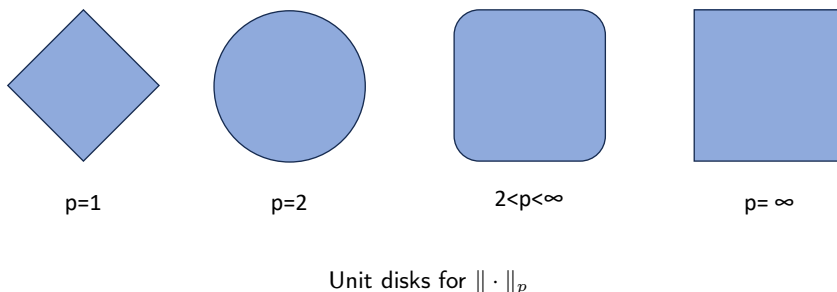
Proof. We define a map $f : E \rightarrow E$ by

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ \frac{\|x\|}{\rho_C(x)}x & \text{otherwise.} \end{cases}$$

One has $f(\overset{\circ}{C} = B(0, 1))$ by (16.2.3.3) and $f(C) = \overline{B(0, 1)}$ by (16.2.3.2). The function f is clearly continuous at all $x \neq 0$. By assumption, there exists $R > 0$ such that $C \subset B(0, R)$ and therefore $\forall x \in E, \rho_C(x) \geq \frac{\|x\|}{R}$ (16.2.3.2) giving the continuity at 0. But the map $g : E \rightarrow E$ defined by

$$g(x) = \begin{cases} 0 & \text{if } x = 0, \\ \frac{\rho_C(x)}{\|x\|}x & \text{otherwise.} \end{cases}$$

is set-theoretically the inverse of f . Just as for f , g is continuous away from 0 and it is continuous at 0 because ρ_C is Lipschitz (16.2.3.3). □



Exercise(s) 16.2.3.6. Show that in general, one cannot find a differentiable homeomorphisms of E mapping a convex into another one.

16.3 The Farkas Lemma

Let E be a vector space (with no further assumption for instance on its dimension or topology). For any non negative integer m and $\underline{\alpha} = (\alpha_1, \dots, \alpha_m) \in (E^*)^m$, let us define $C(\underline{\alpha}) \subset E$ the cone of vertex the origin 0 by

$$C(\underline{\alpha}) = \{ x \in E \mid \alpha_1(x) \leq 0 \cdots \alpha_m(x) \leq 0 \}.$$

Let us give the very elegant proof by David Bartl ([2]) of the Farkas Lemma, a key result in linear programming, a counterpart in this context of the duality result of 7.6.0.2.

Theorem 16.3.0.1 (∞). For any linear form any linear form $\gamma \in E^*$, one has

$$(1) \quad C(\underline{\alpha}) \subset C(\gamma)$$

if and only if

$$(2) \quad \exists t_1, \dots, t_m \in \mathbf{R}^+ \mid \gamma = t_1 \alpha_1 + \dots + t_m \alpha_m$$

Proof. (2) \Rightarrow (1) is trivial. We prove the (2) \Rightarrow (1) part by induction on m . If $m = 0$, one has $C(\underline{\alpha}) = E = C(\gamma)$ implying $\gamma = 0$ which is indeed a non negative (empty!) combination of the α_i 's.

Let us assume that the assertion has been proved for $m \geq 0$ and assume that $C(\underline{\alpha}) \subset C(\gamma)$ for $\underline{\alpha} \in (E^*)^{m+1}$.

If $C((\alpha_1, \dots, \alpha_m)) \subset C(\gamma)$ we are done by induction hypothesis.

If not, there exists

$$(3) \quad \xi \in C((\alpha_1, \dots, \alpha_m)) \mid \gamma(\xi) > 0$$

In particular, $\xi \notin C(\underline{\alpha})$ and therefore $\alpha_{m+1}(\xi) > 0$. Changing ξ into $\xi/\alpha_{m+1}(\xi)$, we may and do assume $\alpha_{m+1}(\xi) = 1$. For any $\varphi \in E^*$, we set

$$\tilde{\varphi}(x) = \varphi(x - \alpha_{m+1}(x)\xi).$$

We have by construction $\tilde{\alpha}_{m+1} = 0$ and therefore

$$x \in C(\tilde{\alpha}_1, \dots, \tilde{\alpha}_m) \implies x - \alpha_{m+1}(x)\xi \in C(\underline{\alpha}) \subset C(\gamma).$$

In other words,

$$C(\tilde{\alpha}_1, \dots, \tilde{\alpha}_m) \subset C(\tilde{\gamma})$$

implying by induction hypothesis

$$\exists t_1, \dots, t_m \in \mathbf{R}^+ \mid \tilde{\gamma} = t_1 \tilde{\alpha}_1 + \dots + t_m \tilde{\alpha}_m$$

or

$$\gamma = t_1 \alpha_1 + \dots + t_{m+1} \alpha_{m+1}$$

with $t_{m+1} = \gamma(\xi) - t_1 \alpha_1(\xi) - \dots - t_m \alpha_m(\xi) \geq \gamma(\xi) > 0$ by (3).

□

Remark(s) 16.3.0.2. *There exists numerous versions and generalizations of Farkas' lemma which are discussed in [2]. The proof given in this paper is general enough to recover all the principal versions!*

Corollary 16.3.0.3. *With the notations above, assume further that E is finite dimensional (or more E is a topological vector space and $\underline{\alpha} \in (E^*)^m$ is made of continuous linear forms). Then, the space of non negative combination of the α_i 's is closed in E^* (in the infinite dimensional case, in the topological dual with the weak point-wise convergence topology).*

16.4 Projection to a Closed Convex Set

In this section 16.4, E is an Euclidean space³ and C is a nonempty closed convex subset.

Recall that the distance of a point to C is defined by

$$d(x, C) := \inf\{\|x - y\| \mid y \in C\}$$

For closed convex sets, an important consequence is the following projection property.

Theorem 16.4.0.1. *With the notations above*

1. *For each $x \in E$ there exists a unique $w \in C$ such that*

$$\|x - w\| = d(x, C)$$

w is called the projection of x to C and is denoted by $p_C(x)$.

2. *We have the obtuse angle⁴ property : $w = p_C(x)$ if and only if*

$$\langle x - w, u - w \rangle \leq 0 \quad \forall u \in C$$

3. *If $x \notin C$, the affine hyperplane H_x through $p_C(x)$ and orthogonal to $e = x - p_C(x)$ is a supporting hyperplane, i.e. such C is contained in one of the two half-spaces delimited by H . In this case, each point of $C \cap H$ is in the boundary $\partial C = C \setminus \overset{\circ}{C}$ of C in E .*

4. *The projection p_C is 1-Lipschitz and is therefore continuous.*

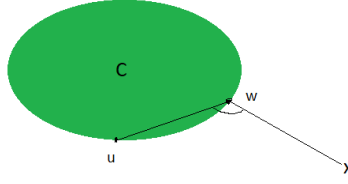
Proof. Uniqueness. The median equality for $x = x - w_1$, $y = x - w_2$ reads

$$4\|x - \frac{w_1 + w_2}{2}\|^2 + \|w_1 - w_2\|^2 = 2\|x - w_1\|^2 + 2\|x - w_2\|^2 = 4d(x, C)^2.$$

Since C is convex, $\frac{w_1 + w_2}{2} \in C$ we have $4\|x - \frac{w_1 + w_2}{2}\|^2 \geq 4d(x, C)^2$ and therefore $\|w_1 - w_2\|^2 \leq 0$, namely $w_1 = w_2$.

³Or, more generally, a real Hilbert space.

⁴More geometrically, $w = p_C(x)$ if and only if $\widehat{(x - w, u - w)} \geq \frac{\pi}{2}$ in the plane $\text{Span}(x - w, u - w)$.



Existence. By definition of $d(x, C)$, there exists $w_k \in C$ such that

$$d(x, C) \leq \|x - w_k\| < d(x, C) + \frac{1}{k}$$

The same mediane equality argument with $x = x - w_p$, $y = x - w_q$ gives

$$\begin{aligned} 4d(x, C)^2 + \|w_q - w_p\|^2 &\leq 4\|x - \frac{w_p + w_q}{2}\|^2 + \|w_q - w_p\|^2 \\ &= 2\|x - w_p\|^2 + 2\|x - w_q\|^2 \\ &\leq 4d(x, C)^2 + \frac{2}{p^2} + \frac{2}{q^2} \end{aligned}$$

hence $\|w_q - w_p\|^2 \leq \frac{2}{p^2} + \frac{2}{q^2}$: the sequence (w_p) is a Cauchy sequence and then converges to $w \in C$ because C is closed in a complete space. Considering the limit of

$$d(x, C) \leq \|x - w_k\| < d(x, C) + \frac{1}{k}$$

we have $d(x, C) = \|x - w\|$. Now suppose $w_1 \neq w_2 \in C$ satisfy

$$\|x - w_1\| = \|x - w_2\| = d(x, C)$$

Then we have

$$\|w_1 - w_2\|^2 = 2\|x - w_1\|^2 - 2\|x - \frac{w_1 + w_2}{2}\|^2$$

Since C is convex, $\frac{w_1 + w_2}{2} \in C$. This gives

$$\|x - \frac{w_1 + w_2}{2}\|^2 < \|x - w_1\|^2 = d(x, C)^2$$

But since C is convex, $\frac{w_1 + w_2}{2} \in C$. This is a contradiction.

Suppose $w = p_C(x)$. Let $u \in C$, $\lambda \in (0, 1)$. Since C is convex, $\lambda u + (1 - \lambda)w \in C$. Then

$$\|x - w\|^2 = d(x, C)^2 \leq \|x - (\lambda u + (1 - \lambda)w)\|^2 = \|x - w\|^2 - 2\lambda\langle x - w, u - w \rangle + \lambda^2\|u - w\|^2$$

That is

$$2\langle x - w, u - w \rangle \leq \lambda\|u - w\|^2$$

Letting $\lambda \rightarrow 0^+$ we have

$$\langle x - w, u - w \rangle \leq 0$$

Conversely suppose

$$\langle x - w, u - w \rangle \leq 0 \quad \forall u \in C$$

Then

$$\|x - u\|^2 = \|x - w\|^2 + 2\langle x - w, w - u \rangle + \|w - u\|^2 \geq \|x - w\|^2$$

Hence $\|x - w\| \leq \|x - u\|$ for all $u \in C$ and $w = p_C(x)$.

The obtuse angle property says precisely that the affine $\{u \in E \mid \langle x - p_C(x), u - p_C(x) \rangle = 0\}$ is an hyperplane ($x - p_C(x) \neq 0$ because $x \notin C$ and this hyperplane is precisely H_x .

Let H be a supporting hyperplane of C of equation $\langle c, e \rangle = \delta$ for some nonzero e . If $\langle c, e \rangle \geq \delta$ for any $c \in C$ and $c_0 \notin \partial C$, then for $\varepsilon > 0$ small enough $c_0 - \varepsilon e \in C$ and therefore, $\langle c_0 - \varepsilon e, e \rangle \geq \delta$ implying $\langle c_0, e \rangle > 0$ and therefore $c_0 \notin H$.

For the last item (where the convex assumption is non necessary), we simply write for $x, y \in E$ and $c \in C$ the triangle inequality

$$\|x - p_C(x)\| \leq \|x - c\| \leq \|x - y\| + \|y - c\|.$$

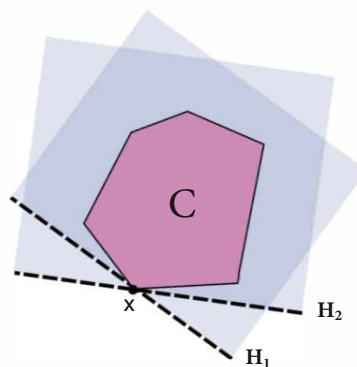
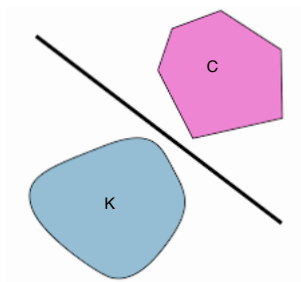
With $c = p_C(y)$, we get $\|x - p_C(x)\| - \|x - p_C(y)\| \leq \|x - y\|$ and by symmetry of x, y , we get

$$\|x - p_C(x)\| - \|x - p_C(y)\| \leq \|x - y\|.$$

□

Corollary 16.4.0.2 (Separation lemma). *Let K be a compact closed subset of E disjoint from C .*

1. *There exists $e \in E \setminus \{0\}, \alpha \in \mathbf{R}$ such that for any $(c, k) \in C \times K$, one has $\langle c, e \rangle > \delta$ and $\langle k, e \rangle < \delta$. Geometrically, each of two half-spaces delimited by the affine hyperplane $H = \{x \mid \langle x, e \rangle = \delta\}$ contains exactly one of the convex set C or K .*
2. *Each point of the boundary ∂C belongs to some supporting hyperplane.*



Proof. The map

$$\begin{cases} K & \rightarrow &]0, \infty[\\ k & \mapsto & d(k, C) = \|k - p_C(k)\| \end{cases}$$

is continuous on the compact K : let us chose a minimizing point $k_0 \in K$ and denote $c_0 = p_C(k_0)$. By construction, one also has $p_K(c_0) = k_0$. By the obtuse angle property of both p_C and p_K , one has

$$\langle k_0 - c_0, c - c_0 \rangle \leq 0 \text{ and } \langle c_0 - k_0, k - k_0 \rangle \leq 0$$

i.e.

$$\langle k, c_0 - k_0 \rangle \leq \langle k_0, c_0 - k_0 \rangle = \alpha \text{ and } \beta = \langle c_0, c_0 - k_0 \rangle \leq \langle c, c_0 - k_0 \rangle.$$

But $\beta - \alpha = \langle c_0 - k_0, c_0 - k_0 \rangle = \|c_0 - k_0\|^2 > 0$ because $C \cap K = \emptyset$ and $e = c_0 - k_0 \neq 0$ for the same reason. One can chose for instance $\delta = \frac{\alpha + \beta}{2}$ (corresponding to the normal hyperplane to e passing through the middle of $[c_0, k_0]$) proving (1). □

16.5 Krein-Milman Theorem

In this section, K denotes a compact convex subset of an Euclidean space E of dimension d .

Definition 16.5.0.1. *A point of a convex set is extremal if it is not an interior point of any of its line segments.*

For instance, the boundary sphere of an Euclidean ball is the set of its extremal points , the extremal points of a plain convex polygon are its vertices (**exercise**)...

Exercise(s) 16.5.0.2. *Let Δ_d the standard d -simplex, convex hull of the elements of the standard basis of \mathcal{B} of \mathbf{R}^d and the origin 0 . Prove that the set of extremal points of Δ_d is $\{0\} \cup \mathcal{B}$.*

In all these cases, the convex hull of the extremal points is precisely the convex itself. This is a general fact.

Let us start with a useful lemma.

Lemma 16.5.0.3. *Let $k \in K$.*

1. *If $x \in \partial K$, then x belongs to some supporting hyperplane H .*
2. *Conversely, let H be a supporting hyperplane of K containing k . Then k is an extremal point of $K \cap H$ is an extremal point of K .*

Proof. Let us chose a sequence $x_i \notin K$ such that $\lim x_i = k$. Let be the sequence of unit vectors

$$e_i = \frac{x_i - p_K(x_i)}{\|x_i - p_K(x_i)\|}$$

and H_i the hyperplane orthogonal to e_i passing through $p_K(x_i)$. Extracting a subsequence if necessary, one can assume that e_i converges to some unit vector e (compactness of a sphere). By continuity of p_K , one has $\lim p_C(x_i) = k$ and the normal hyperplane to e passing through k is the sought supporting hyperplane (check!).

After a translation if necessary, one can assume that $k = 0$. Let $\varphi \in H^*$ be a (linear) equation of the (linear) hyperplane H . One has for instance $\varphi(K) \subset \mathbf{R}^+$ because H is a supporting hyperplane. If $0 \in]k_0, k_1[\subset K$, there exists $t \in]0, 1[$ such that $0 = tk_0 + (1-t)k_1$. Applying φ , we get $0 = t\varphi(k_0) + (1-t)\varphi(k_1)$ with moreover $\varphi(k_0) \geq 0$ and $\varphi(k_1) \geq 0$. Because $t(1-t) > 0$, this implies $\varphi(k_0) = \varphi(k_1) = 0$ hence $0 \in]k_0, k_1[\subset K \cap H$, a contradiction because 0 is extremal in $K \cap H$. \square

Observe that the converse is false (draw an example).

Theorem 16.5.0.4 (Krein-Milman). *Any compact convex K of an Euclidean space is the convex hull of its extremal points.*

Proof. We make an induction on $d = \dim(E)$, starting with the trivial case $d = 0$. Assume $d > 0$ and the theorem proved in dimension $d - 1$ and let $k \in K$ which can be assumed not extremal. If $k \in \partial K$, we apply induction hypothesis to $K \cap H$ for H a supporting hyperplane of K passing through k using the above lemma 16.5.0.3. If not, k is an interior point because K is closed in E . Then k is an interior point of some small line segment I contained in K . Because K is compact, $K \cap \mathbf{R}I$ is a line segment $[k_0, k_1]$ and $k_0, k_1 \in \partial K$ by construction. And we apply the previous argument. \square

Remark(s) 16.5.0.5 (∞). *The Krein-Milman theorem is true in a much more general context. One can show, using Zorn's lemma however, that the convex hull of the extremal points of a compact subset is dense in any compact of a locally convex topological vector space (for instance any metric vector space), cf. [10].*

16.6 Polar Dual of a Convex Body

Here, E is again a d -dimensional Euclidean space (see remark 16.6.0.6 below).

Definition 16.6.0.1. *A convex body K of E is a compact convex K such that 0 is an interior point⁵. The polar dual of a compact body K is $K^* = \{y \in E \mid \forall x \in K \langle x, y \rangle \leq 1\}$.*

⁵Of course, the choice of the interior point is not so crucial. Choosing 0 is just convenient for the duality result 16.6.0.3.

Example(s) 16.6.0.2. *The ellipsoid defined by $S \in \mathcal{S}_d^{++}$ is a convex body (it is the (closed) unit ball for the associated Euclidean norm $\|\cdot\|_S$).*

Notice that the assertion " 0 is a nonempty interior of C in E " is very mild for any convex subset C in the following sense. Let $E' = \vec{C}$. I claim that 0 belongs to the interior of C in E' . Indeed, if $c_1, \dots, c_n \in E$ is a basis of E' defining a linear homeomorphism $\iota : \mathbf{R}^n \simeq E'$, the image of $\iota([0, 1]^n)$ is an open subset neighborhood of 0 contained in C because $0 \in C$ and ι is open.

Lemma 16.6.0.3. *The polar dual define a decreasing involution of the of convex bodies.*

Proof. The polar dual of a convex body K is obviously closed, convex and contains 0 . Let us chose $r > 0$ such that $\overline{B(0, r)} \subset K$. Then, if $0 \neq x \in K^*$, we have $rx/\|x\| \in K$ and $\langle rx/\|x\|, x \rangle = 1$ or $\|x\| \leq 1/r$: the polar dual K^* is closed, bounded in finite dimension and therefore compact.

Let $K \subset K'$ and y in the polar dual of K' . Therefore, for any $x \in K'$, we have $\langle x, y \rangle \leq 1$ and therefore this true for any $x \in K$, i.e. $y \in (K')^*$.

Assume now $x \in K$ and let $y \in K^*$. By definition, for any $\xi \in K$, we have $\langle \xi, y \rangle \leq 1$ and therefore for $\xi = x$, we get by symmetry of the scalar product $\forall y \in K^* \langle y, x \rangle \leq 1$ meaning $x \in (K^*)^*$ or $K \subset (K^*)^*$.

Assume finally $x \notin K$. By the separation lemma 16.4.0.2, there exists $e \in E \setminus \{0\}, \delta \in \mathbf{R}$ such that $\langle x, e \rangle > \delta$ and for any $k \in K$, $\langle k, e \rangle < \delta$. But $\delta > 0$ because $0 \in K$ allowing to define $y = e/\delta$ which belongs to K^* because for any $k \in K$, $\langle k, y \rangle < 1$. Because $\langle x, y \rangle > 1$, this implies $x \notin (K^*)^*$. \square

Example(s) 16.6.0.4. *Let us give a few examples of computation of dual polar.*

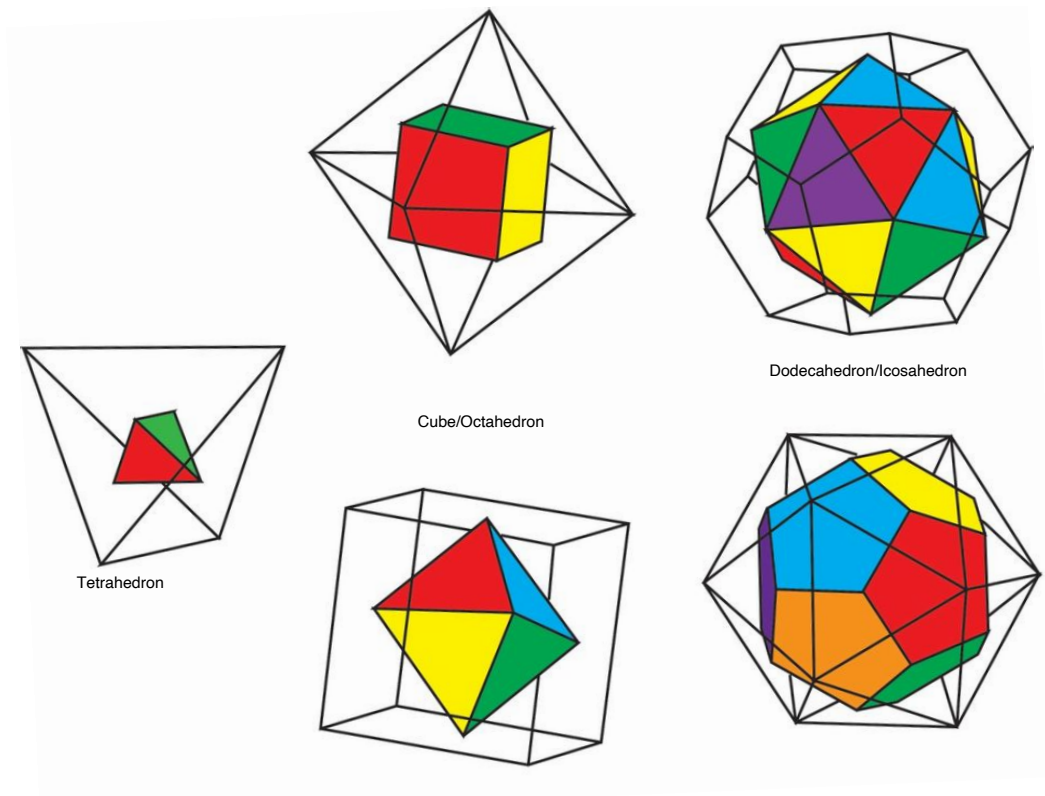
1. For any $S \in \mathcal{S}_d^{++}$, the polar dual of the ellipsoid (12.6.4) \mathcal{E}_S of \mathbf{R}^d is $\mathcal{E}_{S^{-1}}$. Indeed, using the polar involution, one just has to prove $\mathcal{E}_{S^{-1}} \subset (\mathcal{E}_S)^*$. Let $X \in \mathcal{E}_{S^{-1}}$, i.e. $\|X\|_{S^{-1}} \leq 1$. The Cauchy-Schwartz (12.2.2.2) inequality for the Euclidean norm $\|\cdot\|_S$ gives for any $Y \in \mathcal{E}_S$

$$\langle X, Y \rangle \leq \|X\|_{S^{-1}} \|Y\|_S \leq 1$$

hence $X \in (\mathcal{E}_S)^*$.

2. The set of Platonic solids is self-dual (up to rescaling) according o the picture below (16.7.0.1).

The computation of the polar dual of an ellipsoid and the formal properties of polar duality give immediately



Duality of Platonic Solids

Corollary 16.6.0.5 (John's ellipsoid). *Let K be a convex body. Then, the polar dual of the Loewner ellipsoid of K^* (13.4.0.1) is the unique ellipsoid of minimal volume contained in K . By duality,*

Remark(s) 16.6.0.6. *The lemma 16.6.0.3 is valid in a much more general setting but the polar dual has to be defined in the topological dual endowed with the so called $*$ -weak topology ([10]). With these appropriate definitions, the proofs of the statements are analogous to the previous one.*

16.7 Additional Exercises

Exercise(s) 16.7.0.1. *TBD*

Exercise(s) 16.7.0.2. *Let n be an integer. We denote S_n the set of permutations of $\{1, \dots, n\}$. For $\sigma \in S_n$, we define the permutation matrix $M_\sigma = (a_{ij}) \in M_n(\mathbf{R})$ by*

$$a_{ij} = \delta_{\sigma(i)j}.$$

A matrix $M = (a_{ij}) \in M_n(\mathbf{R})$ is called bistochastic if $a_{ij} \geq 0$ for all i, j and

$$\sum_{j=1}^n a_{ij} = \sum_{i=1}^n a_{ji} = 1 \quad \forall i = 1, \dots, n.$$

We denote B_n the set of bistochastic matrices of $M_n(\mathbf{R})$.

1. Show that B_n is compact and convex.
2. Show that a permutation matrix is an extreme point of B_n .
3. Let F_n be the vector space of matrices $M = (a_{ij}) \in M_n(\mathbf{R})$ such that

$$\sum_{j=1}^n a_{ij} = \sum_{i=1}^n a_{ji} = 0.$$

Show that $\dim F_n = (n - 1)^2$.

4. Let M be an extreme point of B_n . Show that M has at most $2n - 1$ non-zero entries.
5. Deduce that M is a permutation matrix.
6. Let $M \in B_n$ be a bistochastic matrix. Show that there exist $\lambda_1, \dots, \lambda_k > 0$ and permutation matrices M_1, \dots, M_k such that

$$\sum_{i=1}^k \lambda_i = 1 \quad \text{and} \quad M = \sum_{i=1}^k \lambda_i M_i \quad \text{and} \quad k \leq (n - 1)^2 + 1.$$

7. Let H be a subgroup of finite index n of G . Let L_i be the elements of the set G/H (left translates of H) and R_i the elements of $H \setminus G$ (right translates of H). Prove that there exists n elements $g_i \in G$ and $\sigma \in S_n$ such that $L_i = g_i H$ and $R_i = H g_{\sigma(i)}$ [Consider the matrix $\text{Card}(G_i \cap H_j)$].

Exercise(s) 16.7.0.3. Helly TBD

Exercise(s) 16.7.0.4. Dualité des cônes TBD

Exercise(s) 16.7.0.5. Polytopes TBD

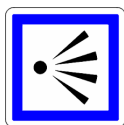
Chapter 17

Complex Hermitian Spaces



Charles Hermite

17.1 Perspective



We give basic results on Hermitian Geometry by analogy with the Real Euclidean Geometry. To emphasize the analogy, we will keep the same plan like the one of the Euclidean case and give the statements in their suitable adapted form. But we will only give detailed proofs for statements requiring new ideas or different calculations letting to the reader to verify that the other proofs are straightforward adaptations of those in the Real case. We should have give a unified and short presentation but we think that repetition, self mathematical practice and progressive generalizations are often at the earth of mathematical education.

17.2 Basics on Hermitian Geometry

Definition 17.2.0.1. Let V be a finite-dimensional \mathbf{C} -vector space of dimension d .

- A sesquilinear map on V is a map

$$\begin{cases} V \times V & \rightarrow & \mathbf{C} \\ (v, w) & \mapsto & \langle v, w \rangle \end{cases}$$

on V is a map which is

1. additive in each variable
 2. skew-linear in the first variable: $\langle \lambda v, w \rangle = \bar{\lambda} \langle v, w \rangle$
 3. linear in the second
- An Hermitian map on V is a sesquilinear map on V which is skew-symmetric: $\langle w, v \rangle = \overline{\langle v, w \rangle}$. In particular $\langle v, v \rangle \in \mathbf{R}$.
 - An Hermitian scalar product on V is a Hermitian map which is positive definite: $\langle \lambda v, v \rangle \in \mathbf{R}^{*+}$ unless $v = 0$. We will then simply say that V is an Hermitian space.

In this chapter, $(V, \langle \cdot, \cdot \rangle)$ will denote an Hermitian space (17.2.0.1).

17.2.1 Examples

Example(s) 17.2.1.1. • The restriction of an Hermitian scalar product $V, \langle \cdot, \cdot \rangle$ to a (complex) subspace is a scalar product: any such finite dimensional vector subspace has the canonical structure of an Hermitian space, with which it is implicitly equipped. In particular, \mathbf{C}^n has a standard Hermitian structure defined by

$$\langle z, z' \rangle = {}^t \bar{z} z' = \sum \bar{z}_i z'_i.$$

- If (X, μ) is a measured space, then a scalar product on $L^2(X, \mu; \mathbf{C})$ is defined by

$$\langle f, g \rangle = \int \bar{f} g d\mu.$$

Therefore, any of its finite dimensional subspace is an Hermitian space.

- If $M \in M_{p,q}(\mathbf{C})$, then $\langle M, N \rangle = \text{tr}({}^t \bar{M} N)$ is a Hermitian scalar product. For this, consider $M = (a_{ij})$ and compute the diagonal terms of ${}^t \bar{M} M = (b_{ij})$:

$$b_{jj} = \sum_k \bar{a}_{kj} a_{kj} = \sum_k |a_{kj}|^2$$

and

$$\text{tr}(\overline{M}M) = \sum_{i,j} |a_{ij}|^2$$

hence $\langle M, M \rangle > 0$ unless $M = 0$.

17.2.2 Hermitian Norm

The following remark is obvious but important and we set $\|v\| = \sqrt{\langle v, v \rangle}$.

Lemma 17.2.2.1. *The real part of a Hermitian inner product is an inner product on the underlying \mathbf{R} -vector space. In particular, it is a normed vector space with a norm satisfying $\|\lambda v\| = |\lambda| \|v\| = |\lambda| \sqrt{\langle v, v \rangle}$.*

As usual, we write v^2 for $\langle v, v \rangle$ and we have by simple bilinearity

$$(v + w)^2 = v^2 + 2\text{Re}\langle v, w \rangle + w^2$$



proving the usual Pythagoras theorem: if and only if $\text{Re}\langle v, w \rangle = 0$, the latter condition being formally different from the Euclidean case.

By 12.2.2.2, we deduce the complex Cauchy-Schwarz inequality :

Proposition 17.2.2.2 (Complex Cauchy-Schwarz Inequality). *Let V be an Hermitian vector space. We have $\text{Re}\langle v, w \rangle \leq \|v\| \|w\|$ with equality if and only if v, w are positively linked. In particular $\|\cdot\|$ is a norm.*

Remark(s) 17.2.2.3. *The previous bilinearity relation immediately give the median formula*

(i)
$$\forall v, w \in V, \|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$

With this underlying Euclidean structure, the unit sphere is $\{z \in \mathbf{C}^n \mid \sum |z_i|^2 = 1\}$ and the distance on the sphere (12.3.0.2) reads in this case $d(z, z') = \arccos \text{Re}\langle z, z' \rangle$.

17.2.3 Dual of an Hermitian space, Orthogonal

The skew symmetry needs to be a little bit carefully for adapting (12.2.3). Let us denote by \overline{V} the abelian group V with twisted vector space structure $(\lambda, v) \mapsto \lambda * v = \overline{\lambda}v$. Of course, a basis of V defines a basis of

\bar{V} and therefore an isomorphism between V and \bar{V} : these two spaces have the same dimension. But with such a generality, the reader will convince himself that there is not any canonical isomorphism $V \simeq \bar{V}$. The partial map

$$\begin{cases} \bar{V} & \rightarrow & V^* \\ v & \mapsto & (w \mapsto \langle v, w \rangle) \end{cases}$$

is a *linear* isomorphism identifying the dual V^* with \bar{V} . Like in the Euclidean case, for every subspace F of V , this isomorphism identifies F^\perp with the usual Hermitian orthogonal $F^\perp = \{v \in V \mid \langle v, F \rangle = \{0\}\}$ and we get as before dimension formula and the orthogonal (Hermitian) decomposition

$$\dim(F^\perp) = d - \dim(F) \text{ and } F \oplus F^\perp = V.$$

In particular, the orthogonal projection p_F onto F is well defined (and \mathbf{C} -linear).

17.2.4 Orthogonalization

Like in 12.2.5.1, we have

Proposition 17.2.4.1 (Gram-Schmidt Algorithm.). *Let e_1, \dots, e_d be a free family in the Euclidean space E . Then, there exists a unique orthonormal family $\varepsilon_1, \dots, \varepsilon_d$ such that*

- $\text{Span}(e_1, \dots, e_i) = \text{Span}(\varepsilon_1, \dots, \varepsilon_i)$ for $i = 1, \dots, d$.
- $(e_i, \varepsilon_i) \in \mathbf{R}^{*+}$ for $i = 1, \dots, d$.

In particular, V admits an orthonormal basis.

And, like in the Euclidean case, we get

Corollary 17.2.4.2. *Every real matrix $M \in \text{GL}_n(\mathbf{C})$ uniquely decomposes into a product $M = QR$ with Q unitary and R upper triangular with diagonal coefficients > 0 .*

and its corollary (see 12.2.5.5)

Corollary 17.2.4.3 (Complex Hadamard Inequality). *Let A_i be the columns of a complex square matrix A . Then, $\det(A) \leq \prod \|A_i\|$ with equality if and only the columns are perpendicular.*

17.2.5 Gram matrix, Minimization of distances

Let us first adapt the notion of symmetric matrix in this Hermitian context.

Definition 17.2.5.1. A complex matrix H is said to be Hermitian if ${}^t\bar{H} = H$.

Exercise(s) 17.2.5.2. Prove that the space of dimension n Hermitian matrix is a real vector space but not a complex vector space and give its dimension.

The Gram matrix of a finite family $v_i \in V$ is the Hermitian matrix $\text{Gram}(v_i) = (\langle v_i, v_j \rangle)$. We'll denote again its determinant by $\text{gram}(x_i)$. Like in 12.2.6.1 we get the matrix computation of the Hermitian product

$$(ii) \quad \langle x, y \rangle = {}^t\bar{X} \text{Gram}(v_i) Y$$

where X, Y are the (complex) column vector of the x_i, y_i respectively. The determinant of an Hermitian matrix being real because it is its own conjugate, $\text{gram}(x_i)$ is a real number. The following corollary is the straightforward adaptation of (12.2.6.1) in the Euclidean case (compare also with § 19 and § 20).

Corollary 17.2.5.3. Let $v_i, w_j \in V$ be finite families and assume there exists relations $w_j = \sum_i p_{i,j} v_i$. Let $P = (p_{i,j})$ be the corresponding (possibly rectangular) complex matrix. Finally, let $x = \sum x_i, y = \sum y_i v_i \in V$ and X, Y the column vectors of the x_i, y_i 's. Then, one has

1. $\langle x, y \rangle = {}^t\bar{X} \text{Gram}(v_i) Y$
2. $\text{Gram}(w_i) = {}^t\bar{P} \text{Gram}(v_i) P$.
3. $\text{gram}(v_i) = 0$ if the v_i 's are not independent and is > 0 else.

and we get complex version of (12.2.7.1)

Proposition 17.2.5.4. Let $f_i, i = 1, \dots, d$ is a basis of a subspace F of E and $x \in E$. Let p_F be the orthogonal projection to F and $d(x, F) = \inf_{y \in F} \|x - y\|$ the distance from x to F .

1. The projection $p_F(x)$ is the only point $y \in F$ such that $d(x, F) = \|x - y\|$.
2. $d(x, F)^2 = \|x - p_F(x)\|^2 = \frac{\text{gram}(x, f_i)}{\text{gram}(f_i)}$.

17.3 Adjoint morphism

Proposition 17.3.0.1. *Let $\mathcal{B} = (e_i)$ be a basis of V and f be an endomorphism of V . There exists a unique endomorphism f^* of V , called the adjoint of f such that*

1. *For all $x, y \in V$,*

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle.$$

2. *One has*

$$\text{Mat}(\mathcal{B}, f^*) = \overline{\text{Gram}(e_i)}^{-1} {}^t \overline{\text{Mat}(\mathcal{B}, f)} \text{Gram}(e_i)$$

In particular, f and f^ have the same rank*

3. *If moreover \mathcal{B} is orthonormal, we have*

$$\text{Mat}(\mathcal{B}, f^*) = {}^t \overline{\text{Mat}(\mathcal{B}, f)}.$$

Proof. Let us denote $G = \text{Gram}(e_i)$ and $A = \text{Mat}(\mathcal{B}, f)$. We write the sought identity in terms of matrices taking into account $\langle x, y \rangle = {}^t \overline{X} G Y$ (17.2.5.3)

$${}^t (\overline{A X}) G Y = {}^t \overline{X} {}^t \overline{A} G Y = {}^t \overline{X} G G^{-1} {}^t \overline{A} G Y = {}^t \overline{X} G (\overline{G}^{-1} {}^t \overline{A} G) Y$$

which allow to defines f^* by the equality $\text{Mat}(\mathcal{B}, f^*) = \overline{G}^{-1} {}^t \overline{A} G$. All the items follow immediately. \square

For instance, isometries f of V are isomorphisms such that $f^{-1} = f^*$. Usual properties of transposition give the usual formulas (linearity of adjunction, $(f \circ g)^* = g^* \circ f^*$, $\text{Id}^* = \text{Id}$). Note that in this Euclidean case, f and f^* are similar (5.5.0.3).

17.4 Complex Normal endomorphisms

Definition 17.4.0.1. *We say that $f \in \text{End}(V)$ is Hermitian if $f = f^*$, skew-Hermitian if $f = -f^*$ and unitary if $f \circ f^* = \text{Id}$ or, equivalently, if $f^{-1} = f^*$, normal if $f \circ f^* = f^* \circ f$. The subgroup ${}^1U(V)$ (resp. $SU(V)$) of $GL(V)$ of unitary morphisms (resp. of $SL(V)$ of determinant 1 unitary matrices) is called the unitary group (resp. the special unitary group $SU(V)$). If V is \mathbf{C}^n with its standard hermitian form, we simply denote $U(V)$ (resp. $SU(V)$) by U_n (resp. by SU_n). Replacing f, f^* by the complex square matrices $M, {}^t \overline{M}$, we get the corresponding notions on matrices.*

17.4.1 Reduction of Complex Normal endomorphisms

Due to the fact that any complex endomorphism has at least one eigenvalue, it has always at least one stable line. Therefore, the reduction theorem of complex endomorphism is simpler than in the real case (12.6.1.3):

Theorem 17.4.1.1 (Reduction of Complex Normal endomorphisms). *Complex normal endomorphisms $f \in \text{End}(V)$ are endomorphisms such that there exists an orthonormal basis \mathcal{B} such that $\text{Mat}(c\mathcal{B}, f)$ is diagonal.*

Like in the Euclidean case, the theorem follows directly by induction from the following key lemma, complex variant of the key lemma (12.6.1.1)

Proposition 17.4.1.2. *The orthogonal of a stable subspace of a normal endomorphism f is stable by f and its restriction to each of these spaces is normal.*

Proof. We proceed like in (12.6.1.1): after choosing a suitable orthonormal basis, we are reduced to prove that if the complex matrix

$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ normal, then $C = 0$ because both A and B are obviously normal. We copy the proof of (17.4.1.2: the $(1, 1)$ block of $M^t \overline{M} - {}^t \overline{M} M$ is $C^t \overline{C} + A^t \overline{A} - {}^t A \overline{A}$ and is zero. Taking its trace, we have $\text{tr}(C^t \overline{C}) = \sum |c_{i,j}|^2 = 0$ and thus $C = 0$. \square

17.4.2 Reduction of Unitarian endomorphisms

We assume that E is a Hermitian space equipped with an orthonormal basis \mathcal{B} .

Lemma 17.4.2.1. *Let u be an endomorphism of E . The following properties are equivalent.*

- $u \in U(E)$;
- for any $x \in E$, we have $\|u(x)\| = \|x\|$;
- $u \circ u^* = \text{Id}$;
- $u^* \circ u = \text{Id}$;
- $\text{Mat}(u, \mathcal{B}) \cdot {}^t \overline{\text{Mat}(u, \mathcal{B})} = I$;
- ${}^t \overline{\text{Mat}(u, \mathcal{B})} \cdot \text{Mat}(u, \mathcal{B}) = I$;
- the columns of $\text{Mat}(u, \mathcal{B})$ form an orthonormal basis ;
- the rows of $\text{Mat}(u, \mathcal{B})$ form an orthonormal basis ;
- u transforms an orthonormal basis into an orthonormal basis.

Moreover, the modulus of any eigenvalue of u is 1

Proof. Only the last assertion is specific to the unitary case. Let x be a nonzero eigenvector of u with eigenvalue λ . One gets

$$0 \neq \langle x, x \rangle = \langle x, u^*u(x) \rangle = \langle u(x), u(x) \rangle = |\lambda|^2 \langle x, x \rangle$$

and therefore $|\lambda| = 1$. □

Corollary 17.4.2.2. *If $u \in \mathbf{U}(\mathbf{E})$, then $|\det u| = 1$ and thus $\mathbf{SU}(\mathbf{E}) = \det^{-1}\{1\} \cap \mathbf{U}(\mathbf{E})$ is a normal subgroup.*

From the lemma and the reduction theorem 17.4.1.1, we get

Corollary 17.4.2.3 (Reduction of Isometries). *A morphism u of \mathbf{V} is unitary if and only there exists an orthonormal basis \mathcal{B} of \mathbf{V} , real numbers θ_j (well defined (mod 2π) up to order) such that $\text{Mat}(u, \mathcal{B}) = \text{diag}(\exp(i\theta_j))$.*

Observing that $\text{diag}(i\theta_j)$ is skew-Hermitian, we get (compare with 12.6.2.5)

Corollary 17.4.2.4. *The continuous map $\exp : \mathcal{A}_n(\mathbf{C}) \rightarrow \mathbf{U}_n(\mathbf{C})$ (resp. $\exp : \mathcal{A}_{n,0}(\mathbf{C}) \rightarrow \mathbf{SU}_n(\mathbf{C})$) is surjective, where $\mathcal{A}_n(\mathbf{C})$ (resp. $\mathcal{A}_{n,0}(\mathbf{C})$) denotes the \mathbf{R} -vector space of skew-Hermitian matrices (resp. traceless skew-Hermitian matrices). In particular, $\mathbf{U}_n(\mathbf{C})$ (resp. $\mathbf{SU}_n(\mathbf{C})$) is path-connected.*

17.4.3 Reduction of Hermitian endomorphisms

Lemma 17.4.3.1. *The eigenvalues of an Hermitian matrix endomorphism u are real.*

Proof. Let x be a nonzero eigenvector of u with eigenvalue λ . One gets

$$\bar{\lambda} \langle x, x \rangle = \langle u(x), x \rangle = \langle x, u^*u(x) \rangle = \langle x, u(x) \rangle = \lambda \langle x, x \rangle$$

and therefore $\bar{\lambda} = \lambda$ because $\langle x, x \rangle \neq 0$. □

From the lemma and the general reduction theorem 17.4.1.1, we get

Theorem 17.4.3.2 (The Spectral theorem for Hermitian endomorphisms). *The Hermitian endomorphisms are the endomorphisms whose matrix in a suitable orthonormal basis is real diagonal. Matricially, Hermitian matrices H are the orthogonally real diagonalizable matrices: there exists a unitary matrix U and a real diagonal matrix Δ such that ${}^t\bar{U}U = \text{Id}$ and $U^{-1}MU = {}^t\bar{U}MU = \Delta$.*

We define (compare with chapter 19 and (12.6.3.4))

Definition 17.4.3.3. *Let H be an Hermitian matrix. We define the associated Hermitian bilinear form of S on $\mathbf{C}^n = M_{n,1}(\mathbf{C})$ by the formula $\langle X, Y \rangle_H = {}^t\bar{X}HY$. We say that H is positive definite if this form is an Hermitian scalar product.*

Corollary 17.4.3.4. *A Hermitian matrix H is positive definite if and only if its eigenvalues are > 0 .*

Proof. Let us write $H = U^{-1}\Delta U = {}^tO\Delta O$ with $\delta = \text{diag}(\lambda_i)$ real diagonal. Then, for any $X \in \mathbf{C}^n$

$$\langle X, X \rangle_H = {}^t\bar{X}HX = {}^t\bar{X}{}^t\bar{U}\Delta UX = \sum_i \lambda_i |\xi_i|^2$$

where ${}^t(\xi_i) = UX$

Assume $\lambda_i > 0$ for all i . Then, $\langle X, X \rangle_H \geq \inf(\lambda_i) \sum |\xi_i|^2 \geq 0$ and $\langle X, X \rangle_H = 0$ only if $\sum |\xi_i|^2 = 0$, that is to say of $UX = 0$ and therefore $X = 0$ because U is invertible being unitary.

Conversely, assume that H is positive definite and define $X_i = U^{-1}e_i$. Then, $\langle X_i, X_i \rangle_H = \lambda_i > 0$. \square

Notice that another way to understand the proof is that the Hermitian scalar product is simply given by the formula $\sum \lambda_i |\xi_i|^2$ in the coordinates ξ_i of an orthonormal basis of eigenvectors of H .

Exercise(s) 17.4.3.5. *Prove that the set of Hermitian positive definite matrices is convex and therefore connected.*

Exercise(s) 17.4.3.6 (Simultaneous Reduction). *Let $H, H' \in M_n(\mathbf{C})$ be two Hermitian matrices with H positive definite. Prove, there exists an invertible matrix $\Pi \in GL_n(\mathbf{C})$ and a real diagonal matrix Δ such that ${}^t\bar{\Pi}H\Pi = \text{Id}$ and ${}^t\bar{\Pi}H'\Pi = \Delta$ [Adapt the proof of (12.6.3.6)].*

Exercise(s) 17.4.3.7. *Let H_1, H_2 be Hermitian positive matrices. Prove $\det(H_1 + H_2) \leq \det(H_1) + \det(H_2)$.*

17.5 Topological Properties of the Unitary Group

Proposition 17.5.0.1. *The groups $U(E)$ and $SU(E)$ are compact and path connected.*

PROOF. *We already know the connectedness (17.4.2.4).*

We then consider the map $f : M \in M_n(\mathbf{C}) \mapsto {}^t \overline{M} M$. We have $U_n(\mathbf{C}) = f^{-1}(I_n)$, thus $U_n(\mathbf{C})$ is closed in $M_n(\mathbf{C})$. To show that $U_n(\mathbf{C})$ is bounded, we consider the norm induced by $\text{tr}^t \overline{M} M$. We have $U_n(\mathbf{C}) \subset B(0, \sqrt{n})$ proving the compactness of $U_n(\mathbf{C})$. The determinant map being continuous (it is polynomial), the compactness of $SU_n(\mathbf{C})$ follows.

17.5.1 Study of \mathcal{H}_n^{++}

We denote \mathcal{H}_n (resp. \mathcal{H}_n^{++}) as the set of Hermitian matrices (resp. positive definite). Like in the real symmetric case (13.3.0.4), we straightforwardly, we immediately get

Proposition 17.5.1.1. *With the notations above, we have*

1. \mathcal{H}_n and \mathcal{H}_n^{++} are convex.
2. The exponential map $\exp : \mathcal{H}_n \rightarrow \mathcal{H}_n^{++}$ is a homeomorphism compatible with the transposition.
3. The map $\text{Sq} : \mathcal{H}_n^{++}(\mathbf{R}) \rightarrow \mathcal{H}_n^{++}(\mathbf{R})$ defined by $\text{Sq}(H) = H^2 S$ is a homeomorphism whose inverse is denoted by $H \mapsto \sqrt{H}$. Moreover $\sqrt{{}^t H} = {}^t \sqrt{H}$.
4. For any $H \in \mathcal{H}_n^{++}$, the complex ellipsoid ${}^{\mathbf{C}}\mathcal{E}_H = \{Z \in \mathbf{C}^n \mid \langle Z, Z \rangle_H \leq 1\}$ is convex compact of volume $(\det S)^{-1/2} \text{vol } {}^{\mathbf{C}}\mathcal{E}_{\text{Id}} = \frac{\pi^n}{\Gamma(n+1)}$.
5. The volum map $\mathcal{H}_n^{++} \rightarrow]0, \infty[$ is strictly convex.
6. If K is a compact subset of \mathbf{C}^n whose interior contains the origin, then there exists a unique $H \in \mathcal{H}_n^{++}$ such that \mathcal{E}^H is the complex ellipsoid of minimal volume containing K .

17.6 Compact subgroups of $GL_n(\mathbf{C})$

From proposition 17.5.1.1, we get exactly like in the real case (13.5.0.2)

Theorem 17.6.0.1. *$SU_n(\mathbf{C})$ (resp. $U_n(\mathbf{C})$) is a maximal compact subgroup of $SL_n(\mathbf{C})$ (resp. of $GL_n(\mathbf{C})$). Moreover, any compact subgroup $GL_n(\mathbf{C})$ is conjugate to a subgroup of $U_n(\mathbf{C})$.*

17.6.1 Complex Polar Decomposition.

Finally, from proposition 17.5.1.1, we get like in (13.6.0.1) using that ${}^t\overline{M}M$ is Hermitian positive definite,

Theorem 17.6.1.1. *The map $\Phi : U(n) \times \mathcal{H}_n^{++} \rightarrow GL_n(\mathbf{C})$ defined by $\Phi(U, S) = US$ is a homeomorphism of inverse $\Psi : M \mapsto \Psi(M) = (M(\sqrt{{}^t\overline{M}M})^{-1}, \sqrt{{}^t\overline{M}M})$. In particular, $GL_n(\mathbf{C})$ is path connected.*

17.7 Additional Exercises

Exercise(s) 17.7.0.1. *Let H be an Hermitian finite dimensionnal space and $f \in \text{End}(H)$ such that for every $x \in H$, $\langle f(x), x \rangle = 0$.*

1. *Prove that for all $x, y \in H$, $\langle f(x), y \rangle = 0$.*
2. *Deduce $f = 0$.*
3. *Does the analogous result still remain true for Euclidean space?*

Part IV

Geometries

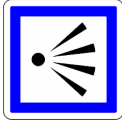
Chapter 18

A primer in Projective Geometry



Francesco di Giorgio Martini

18.1 Perspective



In this chapter, V will denote a \mathbf{k} -vector space of dimension $n + 1 > 0$ and $\mathbf{P}V$ the set of the vector lines of V : the n -dimensional projective space (of V). We will give basics on projective geometry considered as a powerful and natural compactification of affine geometry^a as we will see in the rest of the book. The main goal of this chapter is to understand homographies in terms of collinearity of points (18.5.1.3) and (??). It will be crucial to understand the automorphism group of the general group. Moreover, we will see that the main points are in dimension ≤ 2 but with a big difference between dimension 1 and dimension 2. This is not surprising because the notion of collinear points is empty in dimension 1!

^aFor basic notions of affine geometry, see 18.6

18.2 Introduction

Definition 18.2.0.1. A projective subspace of dimension d (or codimension $n - d$) of $\mathbf{P}V$ is the set of lines of W for some $(d + 1)$ -dimensional vector subspace W of V . If $p \in \mathbf{P}V$, the notation \vec{p} will refer to some nonzero vector in the line $p = \mathbf{k}\vec{p}$. The bijection $f : \mathbf{P}V \xrightarrow{\sim} \mathbf{P}V'$ induced by a linear isomorphism $\vec{f} : V \xrightarrow{\sim} V'$ (by $f(p) = \mathbf{k}\vec{f}(\vec{p})$) is called an homography which in turn determines \vec{f} up to multiplication by a nonzero scalar (*exercice*). If (p_i) is a family of points of $\mathbf{P}V$, we set $\langle p_i \rangle = \mathbf{P}\text{Span}(\vec{p}_i) \subset \mathbf{P}V$.

We therefore have the notion of projective line, plane, hyperplane, collinear points (points lying in the same line) and so on. The projection $\pi : \vec{p} \mapsto \mathbf{k}\vec{p}$ induces a canonical identification $(V - \{0\})/\mathbf{k}^* \xrightarrow{\sim} \mathbf{P}V$: we will use freely these two points of view.

For $V = \mathbf{k}^{n+1}$, we write $\mathbf{P}_\mathbf{k}^n$ or \mathbf{P}^n for $\mathbf{P}(\mathbf{k}^{n+1})$ and $(X_0 : \dots : X_n)$ for $p = \pi(X_0, \dots, X_n)$: these X_i 's, well defined up to multiplication by the same nonzero scalar are called the homogeneous coordinates of p . One has

$$\forall \lambda \neq 0, (X_0 : \dots : X_n) = (\lambda X_0 : \dots : \lambda X_n).$$

This formula shows that any family $(P_j)_j$ of homogeneous polynomial equations in the X_i 's¹ (for instance linear forms) defines a subset of $\mathbf{P}_\mathbf{k}^n$:

$$V((P_j)_j) := \{(X_0 : \dots : X_n) \mid \forall j, P_j(X_0, \dots, X_n) = 0\}.$$

This is the case for linear subspaces $\mathbf{P}W$ as above: take the linear forms in $W^\perp \subset V^*$.



Note that any nonzero value of an homogeneous function f at p does not make any sense because its value depends on the choice of \vec{p} over p (unless $\mathbf{k} = \mathbf{F}_2$!).

¹Or more generally homogeneous functions on \mathbf{k}^{n+1} .

The dimension formula of intersection in linear algebra immediately gives the inequality

$$\dim(\mathbf{P}V_1 \cap \mathbf{P}V_2) \geq \dim(\mathbf{P}V_1) - \operatorname{codim}(\mathbf{P}V_2)$$

proving that in projective geometry lines and hyperplanes always meet! In particular, two lines in the projective plane always intersect.

18.3 Topology of real or complex projective space

Lemma 18.3.0.1. *Let G be a group acting (on the left) isometrically on a metric space (X, d) . Assume that for all $x, x' \in X$, the real $(d/G)(x \bmod G, x' \bmod G) = \inf_{G \times G} d(gx, g'x')$ is reached. Then, (d/G) is a distance on the quotient $G \backslash X$. This is the case for instance if G is a compact group acting continuously on X .*

Proof. If $(d/G)(x \bmod G, x' \bmod G) = 0$, then for some $g, g' \in G$ one has

$$0 = (d/G)(x, x') = d(gx, g'x')$$

hence $gx = g'x'$ implying $x \bmod G = x' \bmod G$.

For the triangle inequality, let us consider $x_1, x_2, x_3 \in X$ such that $(d/G)(x_i \bmod G, x_{i+1} \bmod G) = d(x_i, x_{i+1})$, which is possible because G acts isometrically. For any $g_i \in G$, one has

$$(d/G)(x_1 \bmod G, x_3 \bmod G) \leq d(g_1x_1, g_3x_3) \leq d(g_1x_1, g_2x_2) + d(g_2x_2, g_3x_3)$$

Let $g_1, g_2 \in G$ such that

$$(d/G)(x_1 \bmod G, x_2 \bmod G) = d(g_1x_1, g_2x_2)$$

Then, choose $\gamma_2, \gamma_3 \in G$ such that

$$(d/G)(x_2 \bmod G, x_3 \bmod G) = (d/G)(g_2x_2 \bmod G, x_3 \bmod G)$$

Because,

$$(d/G)(g_2x_2 \bmod G, x_3 \bmod G) = d(\gamma_2(g_2x_2), \gamma_3x_3) = d(\gamma_2(g_2x_2), \gamma_2^{-1}\gamma_3x_3)$$

and set $g_3 = \gamma_2^{-1}\gamma_3$, we get the triangle inequality for (d/G) .

The last point is due to the fact that the map $(g, g') \rightarrow d(gx, g'x')$ is continuous on a compact and therefore that its infimum is a minimum. \square

Corollary 18.3.0.2. *Let S be the Euclidean sphere in \mathbf{k}^{n+1} with $\mathbf{k} = \mathbf{R}$ or $\mathbf{k} = \mathbf{C}$. We endow S with its geodesic distance $d(x, y) = \arccos \operatorname{Re} \langle x, y \rangle$ (12.3.0.2). Let G be the multiplicative group of module 1 elements of \mathbf{k}^* acting diagonally on X and denote $\mathbf{P}^n(\mathbf{k})$ the quotient $G \backslash S$.*

1. *The group G acts by isometry.*
2. *$(d/G)(x \bmod G, y \bmod G) = \arccos |\langle x, y \rangle|$.*
3. *$(x, y) \mapsto \arccos |\langle x, y \rangle|$ defines a distance on $\mathbf{P}^n(\mathbf{k})$.*

Proof. For item (1), observe that for any z in the sphere and any $X, Y \in \mathbf{k}^{n+1}$ one has

$$\operatorname{Re}({}^t(z\bar{X})(zY)) = \operatorname{Re}(|z|^{2t}(\bar{X})(Y)) = \operatorname{Re}({}^t(\bar{X})(Y))$$

For item (2) in the complex case, let us write $\langle x, y \rangle = r \exp(i\theta)$ with $r \geq 0$ and θ real. Then,

$$\operatorname{Re} \langle e^{i\alpha} x, e^{i\alpha'} y \rangle = r \operatorname{Re} \exp(i(\theta - \alpha + \alpha')) = r \cos(\theta - \alpha + \alpha')$$

whose supremum is $r = |\langle x, y \rangle|$ (obtained for $\theta - \alpha + \alpha' = 0$). Because \arccos is decreasing, item (2) follows for the complex case, the real case being analogous (and simpler).

The two first items together with the previous proposition gives (3). □

Exercise(s) 18.3.0.3. *TBD Définir la topo quotient, mettre que projection ouverte ssi saturé ouvert. Critère séparation. En déduire directement la compacité du projectif et que la topologie métrique est la topo quotient.*

18.4 Algebraic and Geometric descriptions

18.4.1 Homogeneous coordinates

Definition 18.4.1.1. *Let S be a set (or a family) of points of an $n \geq 0$ -dimensional projective space. We say that S is in linear general position if any subset (or subfamily) of $k \leq n$ points spans a $(k-1)$ -plane.*

Remark(s) 18.4.1.2.

- *Note that if S has at least $n+1$ points, then S is in linear general position if every subset of $n+1$ points spans \mathbf{P}^n .*

- $\varepsilon_0 = [1 : 0 : \cdots : 0]$, $\varepsilon_1 = [0 : 1 : \cdots : 0]$, ..., $\varepsilon_n = [0 : 0 : \cdots : 1]$, $\varepsilon_{n+1} = [1 : 1 : \cdots : 1]$ are in general position in $\mathbf{P}_{\mathbf{k}}^n$.
- If $n = 1$, any family of distinct points is in general position.

Lemma 18.4.1.3. Let $\varphi : S \rightarrow S'$ be a bijection between projective frames of n -dimensional spaces $\mathbf{P}V, \mathbf{P}V'$. There exists a unique homography $f : \mathbf{P}V \rightarrow \mathbf{P}V'$ induced by φ .

Proof. Let us choose \vec{s}_i, \vec{s}'_i , $i \leq n+1$ in V, V' projecting onto S, S' and a linear map inducing the sought f . The $n+1$ -first elements form a basis of V, V' and we have

$$\vec{s}_{n+1} = \sum_{i \leq n} \lambda_i \vec{s}_i \text{ and } \vec{s}'_{n+1} = \sum_{i \leq n} \lambda'_i \vec{s}'_i$$

with $\prod \lambda_i \lambda'_i \neq 0$ because the points of S, S' are in general position. The equality $\varphi(s_i) = \vec{s}'_i$ forces $\vec{f}(s_i) = \mu_i \vec{s}'_i$ and therefore for $i = n+1$

$$\sum_{i \leq n} \lambda_i \mu_i \vec{s}'_i = \mu_{n+1} \sum_{i \leq n} \lambda'_i \vec{s}'_i$$

giving $\mu_i = \mu_{n+1} \frac{\lambda'_i}{\lambda_i}$. □

Definition 18.4.1.4. A family $p_i \in \mathbf{P}V$ of $(n+2)$ points in general positions is called a projective frame, ε_i being called the standard projective frame of $\mathbf{P}_{\mathbf{k}}$ (18.4.1.2). The image $\pi(p)$ of any $p \in \mathbf{P}V$ by the unique homography $\pi : \mathbf{P}V \rightarrow \mathbf{P}_{\mathbf{k}}^n$ mapping p_i to ε_i is called the homogeneous coordinates of p .

Exercise(s) 18.4.1.5. If $\mathbf{k} = \mathbf{R}, \mathbf{C}$, prove that homogeneous coordinates define a continuous open map $V - \{0\} \rightarrow \mathbf{P}_{\mathbf{k}}^n$ with the above topology of $\mathbf{P}_{\mathbf{k}}^n$ (18.3). Deduce that $\mathbf{P}_{\mathbf{R}}^1$ is homeomorphic to a circle.

18.4.2 Affine charts

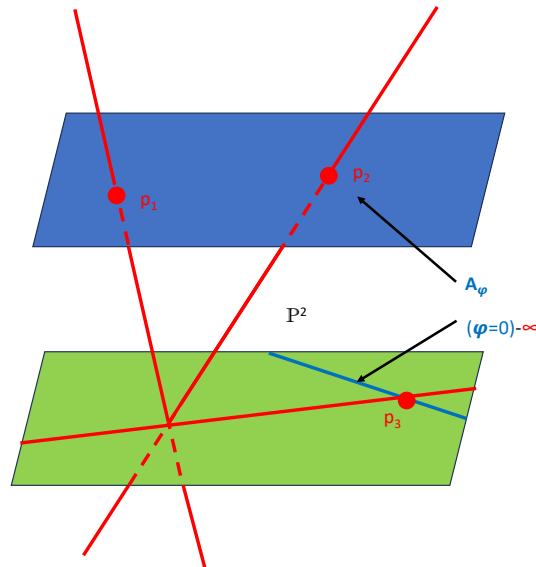
Let $p \in V$ and $\varphi \in V^*$ a linear form non vanishing at p . As before, the equation $\varphi(p) = 0$ makes sense because the vanishing of $\varphi(\vec{p})$ depends only on p . One certainly have the partition

$$\mathbf{P}V = ((\varphi \neq 0) = \{p \in \mathbf{P}V | \varphi(\vec{p}) \neq 0\}) \sqcup ((\varphi = 0) = \{p \in \mathbf{P}V | \varphi(\vec{p}) = 0\}).$$

One has the straightforward lemma

Lemma 18.4.2.1. Let \mathbf{A}_φ be the n -dimensional affine hyperplane of V of (affine) equation $\varphi = 1$. Then, the map $a \mapsto ka$ is a bijection $\mathbf{A}_\varphi \xrightarrow{\sim} (\varphi \neq 0) \subset \mathbf{P}V$ of inverse $p \mapsto \vec{p}/\varphi(\vec{p})$.

In particular, p is contained in the n -dimensional affine space \mathbf{A}_φ with complement the $n - 1$ dimensional projective hyperplane $\mathbf{P} \text{Ker}(\varphi)$, called the hyperplane at infinity.

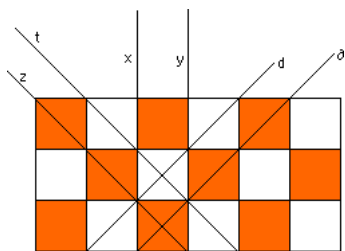


Example(s) 18.4.2.2. If $n = 1$, an hyperplane is just a point and $\mathbf{P}_k^1 = \mathbf{A}_k^1 \sqcup \{\infty\}$. If φ is $(X_0, X_1) \mapsto X_1$, one has $\infty = (0; 1)$ and $(X_0; X_1) = X_0/X_1 = k = \mathbf{A}_k^1$. Geometrically, it "looks like a circle". With this point of view, any family of three distinct points is a projective frame of \mathbf{P}_k^1 , the family is $0, 1, \infty$ being the standard one (18.4.1.2).

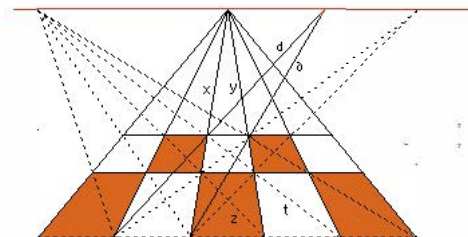
Proposition 18.4.2.3. With the above notations, one has

- The affine lines of \mathbf{A}_φ identify with projective lines on $\mathbf{P}V$ not contained in $\mathbf{P} \text{Ker}(\varphi)$.
- Two affine lines of \mathbf{A}_φ are parallel if and only if they meet in (one) point of $\mathbf{P} \text{Ker}(\varphi)$ (ones says that they mette at infinity).
- The notions of affine and projective collinearity of points of \mathbf{A}_φ coincide.

Proof. Left to the reader. □



Affine lines



Projective lines

The above plane picture illustrate the fact that parallel lines intersect in the horizon line, except *a priori* if these lines are parallel to the horizon (their meeting point is the infinite point of the (unique) horizon projective line).

Explicitely if $V = \mathbf{k}^{n+1}$ and $X_i, i \leq n$ the standard coordinates, $\mathbf{P}^n_{\mathbf{k}}$ is covered by the $n + 1$ affine spaces $(X_i \neq 0)$, each having (affine) coordinates $X_j/X_i, j \neq i$ in the sense where the map

$$\begin{cases} (X_i \neq 0) & \xrightarrow{\sim} & \mathbf{A}^n_{\mathbf{k}} \\ (X_0 : \dots : X_n) & \mapsto & (X_j/X_i)_{j \neq i} \end{cases}$$

is bijective.

Exercise(s) 18.4.2.4. If $\mathbf{k} = \mathbf{R}, \mathbf{C}$, prove that $(\varphi \neq 0)$ is open and dense in $\mathbf{P}^n_{\mathbf{k}}$ with the above topology (18.3)².

18.4.3 Lifting of affine isomorphisms, tbd

Compare with 18.6.1.2 and 18.4.2.3.

We will

18.5 The Fundamental theorem of Projective Geometry

18.5.1 Statement of the main theorem

Definition 18.5.1.1. Let \mathbf{A}, \mathbf{A}' be two affine spaces over two fields k, k' with underlying vector spaces V, V' . Let $\alpha : \mathbf{A} \rightarrow \mathbf{A}'$ and $\kappa : \mathbf{P}V \xrightarrow{\sim} \mathbf{P}V'$ be bijective maps.

1. α is a skew-affine isomorphism if there exists a field isomorphism $\sigma : k \xrightarrow{\sim} k'$ and a σ -linear isomorphism $\vec{\alpha} : V \xrightarrow{\sim} V'$ such that for any $a, b \in \mathbf{A}$ $\overline{\alpha(a)\alpha(b)} = \vec{\alpha}(\overline{ab})$.
2. κ is skew-homography if there exists a field isomorphism $\sigma : k \xrightarrow{\sim} k'$ and a σ -linear isomorphism $\vec{\kappa} : V \xrightarrow{\sim} V'$ such that for any $p \in \mathbf{P}V$, one has $\kappa(p) = \vec{\kappa}(\langle p \rangle)$.
3. κ is a collineation of dimension n if κ preserves collinearity.

²The savant reader will check that the above open covering defines a variety structure in the real or complex case.

Of course, if $\mathbf{k} = \mathbf{F}_p, \mathbf{Q}$, any skew-homography is an homography, simplifying the preceding definitions. But it is also the case for the real field.

Lemma 18.5.1.2. *The identity is the only field endomorphism of \mathbf{R} .*

Proof. Let f be any field endomorphism of \mathbf{R} . The kernel of f is an ideal of the field \mathbf{R} which cannot be \mathbf{R} because $f(1) = 1$: this is $\{0\}$ and f is injective. By additivity, f is the identity on \mathbf{Z} and therefore on \mathbf{Q} by multiplicativity and injectivity. Because any positive number is a square, f maps positive numbers on positive nummbers (multiplicativity) and therefore is increasing (additivity). By density of the rationnal, one concludes $f = \text{Id}$. \square

Of course, a skew-homography is a collineation. We would like to prove the reverse statement in dimension > 1 (in dimension ≤ 1 any bijective map is a collineation!).

Theorem 18.5.1.3. *Any collineation of dimension $n > 1$ is a skew-homography.*

As often in 2, the key argument lies in dimension 2. Even it is not necessary, we have therefore chosen to split the proof in two parts (dimension 2 and reduction to the dimension 2 case. We think that the elementary nature of the arguments are more transparent this way.

18.5.2 The dimension 2 case

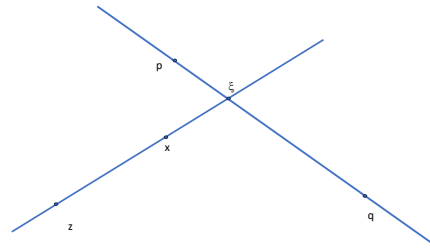
Let's assume $n = 2$ and let κ be a collineation of projective planes. We will prove a few lemmas to reduce to an affine plane statement which will be solved by elementary geometric arguments.

Let $p \neq q$ be two points of \mathbf{PV} ant let $p' = \kappa(p), q' = \kappa(q)$. By assumption, $\kappa(\langle p, q \rangle) \subset \langle p', q' \rangle$. Because $n > 1$, one has $\langle p, q \rangle \neq \mathbf{PV}$.

Lemma 18.5.2.1. *Let $z \notin \langle p, q \rangle$.*

1. $z' = \kappa(z) \notin \langle p', q' \rangle$.
2. *The images of points in general positions are in general positions.*
3. $\kappa(\langle p, q \rangle) = \langle p', q' \rangle$.

Proof. Assume $z' \in \langle p', q' \rangle$ and let's prove that this should imply that $\kappa(\mathbf{PV}) \subset \langle p', q' \rangle$ contradicting the surjectivity. Let $x \neq z$ and ξ the point of $\langle z, x \rangle \cap \langle p, q \rangle$. Then $\kappa(\xi) \in \langle p', q' \rangle$ and therefore $\kappa(x) \in \langle z', \kappa(\xi) \rangle = \langle p', q' \rangle$ because $z' \in \langle p', q' \rangle$, proving the first point.



The second point is a reformulation of the first one.

Assume that $z' = \kappa(z) \in \langle p', q' \rangle$ (κ is surjective). The first point implies $z \in \langle p, q \rangle$ as wanted. \square

Let D be a projective line of $\mathbf{P}V$ and $D' = \kappa(D)$. Let \mathbf{A}, \mathbf{A}' be the affine planes complement of the lines at infinity D, D' in $\mathbf{P}V, \mathbf{P}V'$ respectively. Because parallel affine lines identify with projective lines intersecting at infinity, the lemma 18.5.2.1 shows that κ induces a bijective map $\alpha : \mathbf{A} \xrightarrow{\sim} \mathbf{A}'$ mapping affine lines onto lines and preserving parallelism.

Lets us first show that it's enough to study α to determine κ .

Lemma 18.5.2.2. *If $\alpha : \mathbf{A} \rightarrow \mathbf{A}'$ is skew-affine, then κ is a skew-homography.*

Proof. Let us choose homogeneous coordinates $X_i, X'_i, i = 0, 1, 2$ on $\mathbf{P}V$ and $\mathbf{P}V'$ such that the equation of $D, D' \rangle$ are $X_2 = 0$ and $X'_2 = 0$ respectively. The affine coordinates on \mathbf{A} are $x_i = X_i/X_2, x'_i = X'_i/X'_2, i = 0, 1$. By construction, α can be matricially written as

$$\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \mapsto A \begin{pmatrix} \sigma(x_0) \\ \sigma(x_1) \end{pmatrix} + B$$

for $A \in GL_2(k'), B \in k'^2$. Let $h : \mathbf{P}V \rightarrow \mathbf{P}V'$ be the skew-homography defined by the σ -isomorphism matricially defined by

$$\begin{pmatrix} X_0 \\ X_1 \\ X_2 \end{pmatrix} \mapsto \begin{pmatrix} A \begin{pmatrix} \sigma(X_0) \\ \sigma(X_1) \end{pmatrix} + \sigma(X_2)B \\ \sigma(X_2) \end{pmatrix}$$

Then, $\delta = \kappa \circ h^{-1}$ is a collineation which is the identity on $\mathbf{A} = \mathbf{P}V - D$. Let us chose $O \in \mathbf{A}$ and $\xi \in D$. Then,

$$\delta(\langle O, \xi \rangle) = \langle O, \delta(\xi) \rangle \text{ and } \delta(D) = D$$

thanks to lemma 18.5.2.1. Taking the intersubsection with D we get $\delta(\xi) = \xi$ and $\kappa = h$. \square

The proof of the main theorem 18.5.1.3 in our dimension 2 situation will directly follow from

Lemma 18.5.2.3. *Any bijective map $\alpha : \mathbf{A} \xrightarrow{\sim} \mathbf{A}'$ of affine planes mapping affine lines onto affine lines and preserving parallelism is a skew-affine isomorphism.*

Proof. Choosing an origin $O \in \mathbf{A}$ and $f(O)$ in \mathbf{A}' , one can assume that $\mathbf{A} = \vec{\mathbf{A}}, \mathbf{A}' = \vec{\mathbf{A}'}$ are vector planes.

Let x, y be two non-zero collinear vectors. Then $\langle 0, x \rangle \parallel \langle y, x+y \rangle$ and $\langle 0, y \rangle \parallel \langle x, x+y \rangle$ implies $\langle 0, \alpha(x) \rangle \parallel \langle \alpha(y), \alpha(x+y) \rangle$ and $\langle 0, \alpha(y) \rangle \parallel \langle \alpha(x), \alpha(x+y) \rangle$ meaning α maps the parallelogram $(O, x, x+y, y)$ to the parallelogram $(O, \alpha(x), \alpha(x+y), \alpha(y))$ implying $\alpha(x+y) = \alpha(x) + \alpha(y)$ in this case. If they are collinear, choose z non collinear to x, y and therefore not collinear to $x+y$. By the previous observation, we get $\alpha((x+y)+z) = \alpha(x+y) + \alpha(z)$. But x is also not collinear to $y+z$ and y is not collinear to z . Therefore $\alpha(x+(y+z)) = \alpha(x) + \alpha(y+z) = \alpha(x) + \alpha(y) + \alpha(y+z)$ implying the additivity of α .

Let $\lambda \in k$ and $x \in \mathbf{A} - \{0\}$. We have $\alpha(x) \neq 0$. Because $0, x, \lambda x$ and therefore so are $0, \alpha(x), \alpha(\lambda x)$ are collinear, there exists $\sigma(\lambda, x) \in k'$ such that $\alpha(\lambda, x) = \sigma(\lambda, x)\alpha(x)$. One sets $\alpha(\lambda, 0) = 0$ (notice that the previous equality remains valid). Let's first check that $\sigma(x, \lambda)$ does not depend on $x \neq 0$. As before, let first consider x, y two nonzero points. One has $\alpha(\lambda(x+y)) = \sigma(\lambda, x+y)(x+y)$ and (additivity) $\alpha(\lambda(x+y)) = \sigma(\lambda, x)x + \sigma(\lambda, y)y$ implying $\sigma(\lambda, x) = \sigma(\lambda, y)$ in this case. If x, y are collinear (and non zero) choose $z \notin \langle x, y \rangle$ and observe $\sigma(\lambda, x) = \sigma(\lambda, z) = \sigma(\lambda, y)$ by the previous result. Let us therefore choose $z_0 \neq 0$ and define $\sigma(\lambda) = \sigma(\lambda, z_0)$. For any x , including zero, one has $\alpha(\lambda x) = \sigma(\lambda)x$. The formulas

$$\alpha((\lambda + \mu)z_0) = \alpha(\lambda z_0) + \alpha(\mu z_0), \alpha((\lambda\mu)z_0) = \alpha(\lambda(\mu z_0)), \sigma(k)\alpha(z_0) = \alpha(\langle z_0 \rangle) = \langle \alpha(z_0) \rangle = k'\alpha(z_0)$$

shows that σ is a field isomorphism as wanted. \square

Remark(s) 18.5.2.4. *The condition of preserving parallelism is superfluous for a bijection of affine planes that sends lines surjectively onto lines (*exercice*).*

18.5.3 The general case

Let us assume $n > 2$ and let κ be a collineation as above. We will show how to adapt or use the dimension 2 results and strategy.

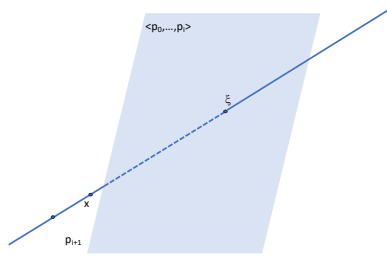
The first step to reduce to the affine case was lemma 18.5.2.1 which we generalize to

Lemma 18.5.3.1. *Let $d \leq n$ and p_0, \dots, p_d points in general positions.*

1. *The images $p'_i = \kappa(p_i)$ are in general positions.*

$$2. \kappa(\langle p_0, \dots, p_d \rangle) = \langle p'_0, \dots, p'_d \rangle.$$

Proof. Let us chose p_{d+1}, \dots, p_n such that $p_i, i \leq n$ are in general position and let us first prove by induction on i that for any $i \leq n$ $\kappa(\langle p_0, \dots, p_i \rangle) \subset \langle p'_0, \dots, p'_i \rangle$ with $p'_i = \kappa(p_i)$. The $i = 0$ statement is clear! Assume $\kappa(\langle p_0, \dots, p_i \rangle) \subset \langle p'_0, \dots, p'_i \rangle$ and let $x \in \langle p_0, \dots, p_{i+1} \rangle$. By induction, one can assume $z \notin \langle p_0, \dots, p_i \rangle$. The line $\langle p_{i+1}, x \rangle$ is contained in the projective space $\langle p_0, \dots, p_{i+1} \rangle$ and therefore meets its hyperplane $\langle p_0, \dots, p_i \rangle$ in ξ . By assumption, $\kappa(x) \in \langle p_{i+1}, \xi \rangle \subset \langle p_0, \dots, p_{i+1} \rangle$ proving the inclusion $\kappa(\langle p_0, \dots, p_{i+1} \rangle) \subset \langle p'_0, \dots, p'_{i+1} \rangle$. For $i = n$, the surjectivity of κ shows that p'_0, \dots, p'_n are in general position, and a fortiori so are p'_0, \dots, p'_d proving the first point.



Conversely, let $z' = \kappa(z) \in \langle p'_0, \dots, p'_d \rangle$ (κ is surjective). We argue like in the dimension 2 case : if $z \notin \langle p_0, \dots, p_d \rangle$, then z, p_0, \dots, p_d are in general position and by the first point so are z', p'_0, \dots, p'_d proving $z' \notin \langle p'_0, \dots, p'_d \rangle$, a contradiction, proving the subsecond point. \square

The reduction to the affine case reads as follows. Let H be a projective hyperplane of $\mathbf{P}V$. Thanks to the preceding lemma, $H' = \kappa(H)$ is an hyperplane of $\mathbf{P}V'$ and κ maps lines onto lines. Let \mathbf{A}, \mathbf{A}' be the affine spaces complement of the hyperplanes at infinity H, H' in $\mathbf{P}V, \mathbf{P}V'$ respectively. Because parallel affine lines identify with projective lines intersubsecting at infinity, the lemma 18.5.3.1 shows that κ induces a bijective map $\alpha : \mathbf{A} \xrightarrow{\sim} \mathbf{A}'$ mapping affine lines onto lines and preserving parallelism.

With these notations, we have nothing to change to genralize lemma 18.5.3.2 (except we are in higher dimension)

Proposition 18.5.3.2. *If $\alpha : \mathbf{A} \rightarrow \mathbf{A}'$ is skew-affine, then κ is a skew-homography.*

Proof. Same proof of the dimension 2 lemma replacing size 2 matrices by size n matrices. \square

We conclude the proof of the main theorem as above by the proof of

Lemma 18.5.3.3. *Any bijective map $\alpha : \mathbf{A} \xrightarrow{\sim} \mathbf{A}'$ of affine planes mapping affine lines onto affine lines and preserving parallelism is a skew-affine isomorphism.*

Proof. One choose origins $O, O' = \kappa(O)$ and let $x, y \in \mathbf{P}(V)$. Let P be a projective plane containing O, x, y . By 18.5.3.1, κ maps P to a projective plane P' . Moreover, $\dim(H \cap P) = 1$ because $P \not\subset H$ ($O \in P$) and $\dim(H \cap P) \geq \dim(P) - \text{codim}(H)$. Therefore $H \cap P$ is a line D like its image $\kappa(D) = D'$. We just have to apply 18.5.2.2 to $\kappa|_P : P \rightarrow P'$ and $\alpha|_{P-D} = \kappa|_{P-D} : P - D \rightarrow P' - D'$ to conclude that α is skew-linear. \square

18.5.4 The dimension 1 case

We assume in this section 18.5.4 $n = 1$.

Definition 18.5.4.1. Let $a, b, c, d \in \mathbf{P}V$ avec a, b, c distinct and $h : \mathbf{P}V \rightarrow \mathbf{P}_k^1$ be the unique homography mapping the projective frame (a, b, c) of $\mathbf{P}V$ to the projective frame $(0, 1, \infty)$ of $\mathbf{P}_k^1 = \mathbf{k} \sqcup \{\infty\}$. The value $h(d) = [a, b, c, d]$ is called the cross-ratio of a, b, c, d . In other words, the cross-ratio is characterized by

$$(h(a), h(b), h(c), h(d)) = (0, 1, \infty, [a, b, c, d])$$

In particular, one has

$$[0, 1, \infty, d] = d.$$

Exercise(s) 18.5.4.2. If moreover $a, b, c, d \in \mathbf{k} = \mathbf{P}_k^1 - \{\infty\}$ are four distinct points, prove the formula $[a, b, c, d] = \frac{c-a}{c-b} / \frac{d-a}{d-b}$. How many values does take the cross-ratio when you are permuting these entries?

The theorem 18.5.1.3 becomes in the dimension 1 case:

Proposition 18.5.4.3.

1. Any homography preserves the cross-ratio.
2. Any bijection of projective lines preserving the cross-ratio is an homography.

Proof. We keep the notation above. Let h' be an homography and denote by (a', b', c', d') the image of any (a, b, c, d) with a, b, c distinct. So are a', b', c' because h' is injective. Moreover, the homography $h \circ h'$ maps (a', b', c') to $(0, 1, \infty)$ showing

$$[a', b', c', d'] = (h \circ h')(d') = h(d) = [a, b, c, d] = [h'(a'), h'(b'), h'(c'), h'(d')].$$

Conversely, let $f : \mathbf{P}V \rightarrow \mathbf{P}V'$ be a bijection of projective lines such that for any (a, b, c, d) with a, b, c distinct

$$[a, b, c, d] = [h(a), h(b), h(c), h(d)].$$

Fix (a, b, c) three distinct points, defining therefore a projective frame of $\mathbf{P}V$. Because f is injective, so does their images. Let $h : \mathbf{P}V \rightarrow \mathbf{P}V'$ the unique homography mapping (a, b, c) to $(f(a), f(b), f(c))$ and $\iota : \mathbf{P}V \rightarrow \mathbf{P}_k^1$ mapping (a, b, c) to $(0, 1, \infty)$. Because both f and homographies preserves the cross ratio, so does the bijection $\varphi = \iota \circ f \circ h^{-1} \circ \iota^{-1}$ of \mathbf{P}_k^1 . But, φ fixes the three points $0, 1, \infty$ and for any $x \in \mathbf{P}_k^1$, one has

$$x = [0, 1, \infty, x] = [\varphi(0), \varphi(1), \varphi(\infty), \varphi(x)] = [0, 1, \infty, \varphi(x)] = \varphi(x)$$

proving $\varphi = \text{Id}$ and $f = h$. □

Exercise(s) 18.5.4.4. *More generally, characterize skew-homographies in terms of cross-ratio.*

18.6 Reminder on Affine Geometry

For convenience of the reader, we recall some elementary basic facts about affine geometry.

An affine space is the data of a simply transitive (right) action of a vector space $\vec{\mathcal{A}}$ (finite dimensional for us) on a (nonempty) set \mathcal{A} . For short, we say that \mathcal{A} is an affine space. If $a_1, a_2 \in \mathcal{A}$, we denote by $a_2 - a_1 = \vec{a_1 a_2}$ the unique vector of $\vec{\mathcal{A}}$ such that $a_2 = a_1 + \vec{a_1 a_2}$. By definition, the dimension of the affine space is the dimension of the underlying vector space.

Example(s) 18.6.0.1.

- *Any vector space E is an affine by its own action by translation. Conversely, the choice of a point (an origin) in $\vec{\mathcal{A}}$ identifies \mathcal{A} and $\vec{\mathcal{A}}$.*
- *The space of solutions of a differential equation $\sum a_i y^{(i)} = f$ is an affine space under the space of solutions of the homogeneous equation $\sum a_i y^{(i)} = 0$.*
- *If $\varphi \in E^*$ is any nonzero linear form, the set $\varphi^{-1}(1)$ is an affine hyperplane of E (under the action of the hyperplane $\varphi^{-1}(0)$) but has no "natural" origin. We will see in a while that this example is indeed universal.*
- *As we will see, the space of vector lines not contained in a given hyperplane of a vector space has a natural structure of affine space (??).*

An affine frame of \mathcal{A} is a set of $n+1$ vectors $a_i, i = 0, \dots, n$ such that $a_i - a_0, i > 0$ is a basis of $\vec{\mathcal{A}}$. This notion does not depend on the numbering of the a_i 's : the dimension of \mathcal{A} is by definition the dimension n of $\vec{\mathcal{A}}$.

An affine map (or affine morphism) $f : \mathcal{A} \rightarrow \mathcal{A}'$ is an application such that there exists $\vec{f} \in \text{Hom}(\vec{\mathcal{A}}, \vec{\mathcal{A}}')$ satisfying

$$\overrightarrow{f(a_1)f(a_2)} = \vec{f}(\vec{a_1 a_2}) \text{ for all } a_1, a_2 \in \mathcal{A}$$

Such an f is unique (hence the notation). Of course, linear affine maps exist and an affine morphism is uniquely defined by the images of an affine frame like in the linear situation. Fixing an origin O of \mathcal{A} , we get the usual formula $f(a) = f(O) + \vec{f}(\vec{Oa})$. For instance, for $\mathcal{A} = \mathbf{R}^n$ with its standard affine structure and origin, we get the usual formula for (endo) affine maps $f(X) = AX + B$ where $A \in M_n(\mathbf{R}), B \in \mathbf{R}^n$. An affine subspace B of \mathcal{A} is a (nonempty) subset of \mathcal{A} which is an affine space under some vector subspace \vec{B} of $\vec{\mathcal{A}}$ (through its induced action on B). In this case, we have simply, as a set, $\vec{B} = \{b - b', b, b' \in B\}$. The dimension theory of vector subspaces translate to the affine situation once we have verified that the spaces we are looking at are nonempty allowing to reduce to the linear case. For instance, the intersection of two affine spaces has codimension the sum of the two codimensions unless the intersection is empty: the reader knows from it childhood that parallel lines exist in the affine plane! This lack of intersection points is precisely why we will introduce projective spaces (18.2).

18.6.1 Universal vector envelop of an affine space

We have seen in example 18.6.0.1 that a non zero linear form φ the natural hyperplane space $\varphi^{-1}(1)$. Let us give two (equivalent) reverse constructions. Let \mathcal{A} be an affine space. The geometrical intuition is clear.

Extrinsic construction. Let us choose an origin O of \mathcal{A} , identifying it with the vector space $\vec{\mathcal{A}}$, and consider the vector space $\widehat{\mathcal{A}} = \vec{\mathcal{A}} \times \mathbf{k}$. Then, the map $(\vec{a}, 1) \mapsto O + \vec{a}$ has inverse $a \mapsto 0 + \xrightarrow{0a}$ and is an affine isomorphism from \mathcal{A} to the affine hyperplane with equation $x_{n+1} = 1$ (where x_{n+1} is the coordinate on the \mathbf{k} factor).

Intrinsic construction. The reader can skip this intrinsic version though it is not only completely formal. Keeping the notations of example 18.6.0.1, we have

$$E = \sqcup_{\lambda \in k} \lambda \varphi^{-1}(\lambda) = \varphi^{-1}(0) \sqcup \sqcup_{\lambda \neq 0} \varphi^{-1}(1) \times \{\lambda\}$$

where identify $e \in \varphi^{-1}(\lambda)$ and $(e/\lambda, \lambda)$ for any nonzero λ . In particular, for any nonzero λ , we have the quite "strange" formula for the scalar multiplication $\mu(e, \lambda) = (e, \mu\lambda)$ if $\mu \neq 0$ and $\mu(e, \lambda) = 0$ if $\mu = 0$.

As a set, we define

$$\widehat{\mathcal{A}} = \vec{\mathcal{A}} \sqcup \sqcup_{\lambda \neq 0} \mathcal{A} \times \{\lambda\}$$

Mapping $\vec{\mathcal{A}}$ to zero and its complement thanks to the second projection defines a map of sets $\varphi : \widehat{\mathcal{A}} \rightarrow k$. Of course, we make the identifications thanks to the first projection $\varphi^{-1}(1) = \mathcal{A}$.

We define a scalar multiplication on each $\varphi^{-1}(\lambda)$ by analogy to our previous example: if $\lambda = 0$, it is the scalar multiplication of $\vec{\mathcal{A}} = \varphi^{-1}(0)$ and if $\lambda \neq 0$, we define $\mu \odot (a, \lambda) = (a, \mu\lambda)$ if $\mu \neq 0$ and $\mu \odot (a, \lambda) = 0$ if $\mu = 0$.

Let us define the sum \oplus of two elements $\widehat{a}, \widehat{a}' \in \widehat{\mathcal{A}}$.

If $\varphi(\widehat{a}) = \varphi(\widehat{a}') = 0$, the sum is given by the vector space structure on $\varphi^{-1}(0)$.

If $\varphi(\widehat{a}) = \lambda \neq 0, \varphi(\widehat{a}') = 0$ we have necessarily

$$\widehat{a} \oplus \widehat{a}' = (a, \lambda) \oplus a' = \lambda \odot ((a, 1) \oplus \frac{1}{\lambda} a')$$

and we define

$$(a, 1) \oplus \frac{1}{\lambda} a' = a + \frac{1}{\lambda} a' \in \mathcal{A}$$

as the right action of $\frac{1}{\lambda} a' \xrightarrow{\widehat{\mathcal{A}}}$ on $a \in \mathcal{A}$.

If $\varphi(\widehat{a}) = 0$ $\varphi(\widehat{a}') \neq 0$, we define

$$a \oplus a' = a' \oplus a$$

If $\varphi(\widehat{a})\varphi(\widehat{a}') \neq 0$ because we have the necessary formula $\widehat{a} = \frac{1}{\varphi(a)} \odot a$, one can assume further $\varphi(a) = \varphi(a') = 1$.

If characteristic of k is not 2, we define

$$a \oplus a' = 2 \odot \left(\frac{1}{2} a \oplus \frac{1}{2} a' \right)$$

If the characteristic is 2, we must have $a \oplus a' = (-a) \oplus a'$ and we define $a \oplus a' = \overrightarrow{aa'}$. By a both straightforward and tedious computation³, we get

Proposition 18.6.1.1. *With these two laws, $\widehat{\mathcal{A}}$ is a $n+1$ dimensional vector space, the map φ is linear and \mathcal{A} is canonically isomorphic to $\varphi^{-1}(1)$. In particular, \mathcal{A} is canonically embedded in its so called projective completion $\mathbf{P}\widehat{\mathcal{A}}$ as an affine chart.*

Exercise(s) 18.6.1.2. *Let $f : \mathcal{A} \rightarrow \mathcal{A}'$ be an affine map. Prove that there exists a unique morphism $\widehat{f} : \widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{A}'}$ lifting the natural inclusions of these affine spaces to their envelops. Deduce from the uniqueness property the compatibility between composition and envelop. Prove that if f is moreover bijective then it can be uniquely lifted to an homography $\mathbf{P}\mathcal{A} \xrightarrow{\sim} \mathbf{P}\mathcal{A}'$.*

³In fact, choosing an affine frame of \mathcal{A} , the reader will observe that $\widehat{\mathcal{A}}$ is identified with \mathbf{R}^{n+1} and that our two laws \odot, \oplus become the usual external product and sums of \mathbf{R}^{n+1} proving the assertion without too much calculations.

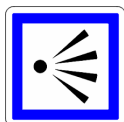
Chapter 19

Sesquilinear Forms and Projective Geometry



Besma sesquilinearia

19.1 Perspective



We generalize the Euclidean or Hermitian scalar products to general sesquilinear forms over some field \mathbf{k} without assuming any symmetry *a priori*. We will see that a lot of notions/methods are valid in this general case, notions that will be crucial for our study of general quadratic forms (see chapter 21). This generality is not "just for fun". General sesquilinear forms are strongly related to projective geometry and even arise in quantum mechanics, at least historically ([5], § 13 and 14)!

19.2 Introduction

In the linear case, two equivalent perspectives -morphisms and matrices- complement each other fruitfully giving rise naturally to an equivalence relation on square matrices, the similarity relation (see chapter 5). Like in the linear case, two equivalent perspectives -sesquilinear forms and matrices- complement each other fruitfully giving rise naturally to a new equivalence relation on square matrices, the congruence relation. Unlike the linear case, we will see in this bilinear context¹ that there is no hope of describing the classes of congruences in a unified way over any field as in the case of similarity relations. Observe it is already impossible in dimension 1: a bilinear form up to congruence is a scalar up to multiplication by a non-zero square of \mathbf{k} and that the quotient set heavily depend on the field \mathbf{k} !

19.2.1 Notations and Reminders

- σ is an automorphism of \mathbf{k} of a field \mathbf{k}^2 .
- A σ -linear map $f : E \rightarrow F$ between \mathbf{k} vector space is an additive map such that $f(\lambda x) = \sigma(\lambda)f(x)$ for any $x \in E, \lambda \in \mathbf{k}$.
- In this chapter, E, \tilde{E} are \mathbf{k} -vector spaces³ over \mathbf{k} with the same finite dimension n (although most formal definitions generalize without this assumption as the reader will easily convince themselves).
- We denote by $\mathcal{B} = (e_i), \tilde{\mathcal{B}} = (\tilde{e}_i)$ basis of E, \tilde{E} .
- We recall that if $\mathcal{B} = (e_i)$ and $\mathcal{B}' = (e'_i)$ are bases of E (of finite dimension), the columns of the base change matrix $P = \text{Mat}_{\mathcal{B}, \mathcal{B}'}(\text{Id}_E)$ are the coordinates of the vectors of \mathcal{C}' in the base \mathcal{B} . If $X = [x]_{\mathcal{B}}$ (resp. $X' = [x]_{\mathcal{B}'}$) are the coordinates of $x \in E$ in \mathcal{B} (resp. in \mathcal{B}'), then $X = PX'$.

19.3 Sesquilinear Forms

Definition 19.3.0.1. We say that $b_\sigma : E \times E \rightarrow \mathbf{k}$ is sesquilinear if σ is

- additive in each of the variables
- σ -linear in the first variable: $b_\sigma(\lambda x, \tilde{x}) = \sigma(\lambda)b_\sigma(x, \tilde{x})$ for all $x \in E, \tilde{x} \in \tilde{E}, \lambda \in \mathbf{k}$
- linear in the second variable: $b_\sigma(x, \lambda \tilde{x}) = \lambda b_\sigma(x, \tilde{x})$ for all $x \in E, \tilde{x} \in \tilde{E}, \lambda \in \mathbf{k}$.

¹Apart from the alternating case (20.5).

²The most important case for us being the case where σ is an involution, in particular when $\mathbf{k} = \mathbf{R}$ and $\sigma = \text{Id}$ or $\mathbf{k} = \mathbf{C}$ with σ being the complex conjugation.

³We have chosen E as in Euclid rather than V as in vector to distinguish the quadratic/Hermitian context from the general vector context.

If $\sigma = \text{Id}$, a sesquilinear form is simply called a bilinear form. We define $\text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}}) = (b_\sigma(e_i, \tilde{e}_j))_{i,j} \in M_n(\mathbf{k})$.

This introduces the notion of sesquilinear (resp. bilinear) space (a vector space equipped with a sesquilinear form b_σ , a morphism between such spaces being a linear map that preserves the forms), a vector subspace defining a (sub-)sesquilinear space by restriction of b_σ .

Remark(s) 19.3.0.2. A fundamental example is $\mathbf{k} = \mathbf{R}, \sigma = \text{Id}, E = \tilde{E}, \mathcal{B} = \tilde{\mathcal{B}}$ (resp. $\mathbf{k} = \mathbf{C}, \sigma = \text{conjugation}, E = \tilde{E}, \mathcal{B} = \tilde{\mathcal{B}}$) and b_σ is a real (resp. complex hermitian) scalar product. In these cases, the above matrix is nothing but its Gram matrix relative to \mathcal{B} .

Like in the Euclidean or Complex Hermitian case (12.2.6) and (17.2.5), we get for $X = [x]_{\mathcal{B}}$ and $\tilde{X} = [\tilde{x}]_{\tilde{\mathcal{B}}}$

$$(i) \quad \begin{array}{l} b_\sigma(x, \tilde{x}) = {}^t\sigma(X) \cdot \text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}}) \cdot \tilde{X} \\ \text{Mat}(b_\sigma, \mathcal{B}', \tilde{\mathcal{B}}') = {}^t\sigma(P) \text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}}) \tilde{P} \end{array}$$

where P, P' are the base change matrices from $\mathcal{B}, \tilde{\mathcal{B}}$ to basis $\mathcal{B}', \tilde{\mathcal{B}}'$ of E, \tilde{E} . **Thus, thanks to the previous formulas (i), the choice of basis $\mathcal{B}, \mathcal{B}'$ allows us to identify b_σ with its matrix $\text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}})$, which we will do freely. All properties of b_σ will have its matrix counterpart.**

19.3.1 Non-degenerate Forms

We denote E_σ the vector space whose underlying group is E whose external multiplication \cdot_σ is twisted by σ , i.e., $\lambda \cdot_\sigma x = \sigma(\lambda)x$. A basis \mathcal{B} of E is still a basis noted \mathcal{B}_σ of E_σ implying that E and E_σ have the same dimension.

A σ -linear application $f : E \rightarrow F$ is therefore a linear map $f \in \text{Hom}(E, F_\sigma)$.

In particular, the twisted dual $E_\sigma^* = (E_\sigma)^* = \text{Hom}_{\mathbf{k}}(E_\sigma, \mathbf{k})$ of E is the space of σ -linear forms with a σ -dual basis \mathcal{B}_σ^* which is the dual basis of \mathcal{B}_σ .

We can associate with b_σ the linear applications

$$(ii) \quad \check{b}_\sigma : \begin{cases} \tilde{E} & \rightarrow & E_\sigma^* \\ \tilde{x} & \mapsto & (x \mapsto b_\sigma(x, \tilde{x})) \end{cases} \quad \text{and} \quad \hat{b}_\sigma : \begin{cases} E_\sigma & \rightarrow & \tilde{E}^* \\ x & \mapsto & (\tilde{x} \mapsto \sigma^{-1}(b_\sigma(x, \tilde{x}))) \end{cases}$$

We observe that $\text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}}) = \text{Mat}(\check{b}_\sigma, \mathcal{B}, \tilde{\mathcal{B}}_\sigma^*)$ and ${}^t\text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}}) = \sigma^{-1}(\text{Mat}(\hat{b}_\sigma, \mathcal{B}_\sigma, \mathcal{B}^*))$.

Definition 19.3.1.1. We define the left kernel $\text{Ker } b_\sigma$ of b_σ as

$$\text{Ker } b_\sigma = \text{Ker } \hat{b}_\sigma = \{\tilde{x} \in E, \forall x \in E, b_\sigma(x, \tilde{x}) = 0\}.$$

We say a bilinear form b_σ is non-degenerate if its kernel is null, i.e., if its matrix in a basis \mathcal{B} is invertible.

In particular, we have $b_\sigma(\mathbb{E}, \text{Ker}(b_\sigma)) = \{0\}$.

A non-degenerate form thus identifies $\tilde{\mathbb{E}}$ with the twisted dual \mathbb{E}_σ^* via ii).

Remark(s) 19.3.1.2. • If we identify $\mathbb{E}, \tilde{\mathbb{E}}$ with \mathbf{k}^n thanks to $\mathcal{B}, \tilde{\mathcal{B}}$, by (i) we have $\text{Ker}(b_\sigma) = \text{Ker}(\text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}}))$.

• If we had swapped the roles of x and \tilde{x} , i.e. used \widehat{b}_σ to define the right kernel, it would have been the kernel of the transpose ${}^t\text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}})$. The notion of degeneracy would not have changed.

• In the non-degenerate case, ${}^t\widehat{b}_\sigma \in \text{Hom}(\tilde{\mathbb{E}}, \mathbb{E}_\sigma^*)$ is an isomorphism: we then traditionally define the asymmetry

$$\beta = {}^t\widehat{b}_\sigma^{-1} \circ \check{b}_\sigma \in \text{End}_{\mathbf{k}}(\mathbb{E}_\sigma^*) = \text{End}_{\mathbf{k}}(\tilde{\mathbb{E}})$$

of the non-degenerate form b_σ . It is the unique isomorphism verifying

$$b_\sigma(y, x) = b_\sigma(\beta(x), y)$$

for all $x, y \in \mathbb{E}$. Its matrix⁴ is simply $\sigma(\text{Mat}(\widehat{b}_\sigma), \mathcal{B}_\sigma, \mathcal{B}^*)^{-1} \text{Mat}(\check{b}_\sigma, \mathcal{B}, \tilde{\mathcal{B}}_\sigma^*)$.

If b_σ is degenerate, then the associated matrix in a basis adapted to the direct sum of the kernel and an arbitrary complement has $\dim \text{Ker } b_\sigma$ zero columns.

19.4 (Left) Orthogonality

Definition 19.4.0.1. If $F \subset \mathbb{E}$ is a subset, the left orthogonal of F is the vector subspace of \mathbb{E} .

$$F^\perp = \{\tilde{x} \in \mathbb{E}, \forall x \in F, b_\sigma(x, \tilde{x}) = 0\}.$$

Thus we have $b(F, F^\perp) = \{0\}$. But in this generality, we do not have in general $b(F^\perp, F) = \{0\}$ so that usually we have $F \not\subset (F^\perp)^\perp$ (**Exercise** find an example!).

We then have part of the usual properties of orthogonality (compare with 20.4.0.2)

Proposition 19.4.0.2. Suppose b_σ is non-degenerate and let F, G be subspaces of \mathbb{E} .

1. $\dim F + \dim F^\perp = \dim \mathbb{E}$,
2. $(F + G)^\perp = F^\perp \cap G^\perp$,

⁴In the bilinear case ($\sigma = \text{Id}$), this matrix is called the *co-square* of $\text{Mat}(b_\sigma, \mathcal{B})$.

3. Orthogonality is decreasing.

PROOF. By definition, $\widehat{b}_\sigma(F)$ is the orthogonal in duality of F in E_σ^* . The first formula follow from the dimension calculation of the orthogonal in duality (7.5.0.1). The last two are formal (but useful).

Remark(s) 19.4.0.3. The notion of orthogonality for general sesquilinear forms without hermitian symmetry is delicate. We reserve the notion of orthogonal direct sum of two subspaces E_1, E_2 to the case where the subspaces are both direct sums and orthogonal on the right and left, i.e., $b_\sigma(E_1, E_2) = b_\sigma(E_2, E_1) = \{0\}$. We then write $E_1 \perp \oplus E_2$. Matrix-wise, in a basis adapted to the direct sum, this means that the matrix of b_σ is block diagonal and the form is non-degenerate if and only if the blocks are.

19.4.1 Adjoint

The following proposition is just a formal generalization of the existence of adjoint in the complex hermitian context (17.3.0.1).

Proposition 19.4.1.1. Let $b_\sigma : E \times \tilde{E} \rightarrow \mathbf{k}$ be a **non-degenerate** sesquilinear form and f an endomorphism of E . There exists a unique endomorphism f^* of \tilde{E} , called the adjoint of f (relative to b_σ), such that for all $x \in E, \tilde{x} \in \tilde{E}$,

$$b_\sigma(f(x), \tilde{x}) = b_\sigma(x, f^*(\tilde{x})).$$

We have

$$(iii) \quad \text{Mat}(f^*, \tilde{\mathcal{B}}) = \text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}})^{-1} \sigma({}^t \text{Mat}(f, \mathcal{B})) \text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}})$$

In particular, f and f^* have the same rank and if \mathcal{B} is orthonormal, we have

$$\text{Mat}(\mathcal{B}, f^*) = {}^t \sigma(\text{Mat}(\mathcal{B}, f))$$

Proof. Let $M = \text{Mat}(b_\sigma, \mathcal{B}, \tilde{\mathcal{B}})$ be the matrix of b_σ $\mathcal{A} = \text{Mat}(f, \mathcal{B})$ that of f . We write the sought identity matrix-wise taking into account the computation (i)

$$b_\sigma(f(x), \tilde{x}) = {}^t \sigma(\text{AX}) \cdot M \cdot \tilde{X}.$$

We get

$${}^t \sigma(\text{AX})M\tilde{X} = {}^t \sigma(X){}^t \sigma(\mathcal{A})M\tilde{X} = {}^t \sigma(X)MM^{-1}{}^t \sigma(\mathcal{A})M\tilde{X} = {}^t \sigma(X)M(M^{-1}{}^t \sigma(\mathcal{A})M)\tilde{X}.$$

□

Usual propositions of transposition give the usual formulas (linearity of adjunction, $(f \circ g)^* = g^* \circ f^*$). Note that in the bilinear case ($\sigma = \text{Id}$), f and f^* are similar (5.5.0.3).

Exercise(s) 19.4.1.2. Show that the isomorphism $\check{b} : \tilde{E} \rightarrow E_\sigma^*$ defined by b_σ (cf. 19.3.1) identifies the adjoint f^* of $f \in \text{End}_{\mathbf{k}}(E)$ with its (σ -twisted transpose ${}^t f \in \text{End}_{\mathbf{k}}(E_\sigma^*) = \text{End}_{\mathbf{k}}(\tilde{E})$) (cf. 7.7.0.1).

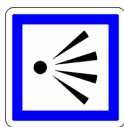
Chapter 20

ε -Hermitian Forms



Taj Mahal

20.1 Perspective



TBD

Generalizing the usual Euclidean or Hermitian classical situation, we focus our attention to σ -sesquilinear forms where σ is not only an automorphism but an involution of \mathbf{k} with a suitable symmetry property. As we will see, the first fundamental tool in their classification, which allows many problems to be reduced to forms on lines or planes like in the Euclidean case.

20.2 Introduction

The theory of sesquilinear forms serves as the foundation for geometry: real symmetric bilinear forms lead to Riemannian geometry, are ubiquitous in number theory, and complex Hermitian forms to complex and

holomorphic geometry. Moreover, the imaginary part of a complex Hermitian form defines an alternating form, which lies at the heart of symplectic geometry. These geometries are fundamental in classical physics (Euclidean geometry), relativistic physics (Lorentzian geometry for special relativity, general Riemannian geometry for general relativity), mechanics (symplectic geometry), and quantum mechanics (Hermitian geometry).

20.3 Definitions

In this chapter, σ is an involution of \mathbf{k} and we fix $\varepsilon \in \{\pm 1\}$. By analogy with the complex conjugation, for any $\lambda \in \mathbf{k}$, we'll often denote $\sigma(\lambda)$ by

$$\sigma(\lambda) = \bar{\lambda}.$$

We keep the notations of 19.2.1.

Definition 20.3.0.1. *We say that the sesquilinear form is*

- *Hermitian if for all $x, y \in E$ one has $b_\sigma(y, x) = \sigma(b_\sigma(x, y)) = \overline{b_\sigma(x, y)}$,*
- *skew-Hermitian if for all $x, y \in E$ one has $b_\sigma(y, x) = -\sigma(b_\sigma(x, y)) = -\overline{b_\sigma(x, y)}$,*
- *ε -Hermitian if it is either Hermitian or skew-Hermitian.*

If $\sigma = \text{Id}$, an ε -Hermitian form is called an ε -symmetric form.

If E is of finite dimension and if $\mathcal{B} = \{e_i\}_{1 \leq i \leq n}$ is a basis, we define $\text{Mat}(b_\sigma, \mathcal{B}) = (b_\sigma(e_i, e_j))_{i,j}$. The basis \mathcal{B} is said to be orthogonal (resp. orthonormal) if $\text{Mat}(b_\sigma, \mathcal{B})$ is diagonal (resp. the identity).

Remark(s) 20.3.0.2. *Let $H = \text{Mat}(b_\sigma, \mathcal{B})$ and X, Y the columns of coordinates of $x, y \in V$ with respect to \mathcal{B} .*

- *The ε -symmetry above forces σ to be an involution and the equality $\varepsilon = \pm 1$, at least if b_σ is non degenerate.*
- *Like in the classical complex Hermitian case (17.2.5.3) or more generally in the sesquilinear case 19.3.i, we get the formula*

$$b_\sigma(x, y) = {}^t\bar{X} \cdot H \cdot Y.$$

- *Notice that ${}^tH = \varepsilon\bar{H}$. Conversely, any such matrix $H \in M_n(\mathbf{k})$, ε -Hermitian matrix say, defines an ε -Hermitian form on \mathbf{k}^n by the formula $b_\sigma(x, y) = {}^t\bar{X}HY$.*
- *If H' is the matrix of b_σ in another basis and P is the corresponding base change matrix, we also get*

$$H' = {}^t\bar{P}HP.$$

In this chapter, we consider b_σ an ε -Hermitian form. Thus we have

$$b_\sigma(x, y) = 0 \Rightarrow b_\sigma(y, x) = 0.$$

We recall (19.3.1) that a sesquilinear form b_σ which is ε -Hermitian defines a non-degenerate form on $E/\text{Ker}(b_\sigma)$, or equivalently, on any complement of the kernel thus reducing their study to the non-degenerate case. Sometimes this kernel is called the "radical" of b_σ or of (E, b_σ) or even the radical $\text{rad}(E)$ of E when no confusion is to be feared.

From the obvious property

$$b_\sigma(\text{Ker}(b_\sigma), E) = b_\sigma(E, \text{Ker}(b_\sigma)) = \{0\},$$

one can straightforwardly reduce the study of Hermitian form to the non degenerate case.

Proposition 20.3.0.3. *Let b_σ be an ε -Hermitian form on E and F be a supplementary subspace of $\text{Ker}(b_\sigma)$.*

1. *There exists a unique form \tilde{b}_σ on $E/\text{Ker}(b_\sigma) = \tilde{E}$ inducing b_σ , i.e. such that the diagram*

$$\begin{array}{ccc} E \times E & \xrightarrow{b_\sigma} & \mathbf{k} \\ \text{quotient} \downarrow & \nearrow \tilde{b}_\sigma & \\ \tilde{E} \times \tilde{E} & & \end{array}$$

commutes and \tilde{b}_σ is non degenerate (\tilde{E} is the associated non degenerate quotient space).

2. *The direct sum $E = \text{Ker}(b_\sigma) \oplus F$ is orthogonal and induces an isomorphism on (non degenerate) ε -Hermitian space $F \simeq \tilde{E}$.*
3. *Two ε -Hermitian spaces are isomorphic if and only if they have the same rank and their non degenerate quotient spaces are isomorphic.*

20.4 Orthogonality

Definition 20.4.0.1. *If $F \subset E$ is a subset, the left orthogonal of F is the vector subspace of E .*

$$F^\perp = \{x \in E, \forall y \in F, b_\sigma(x, y) = 0\}.$$

Thus we have $b(F^\perp, F) = \{0\}$ and therefore $b(F, F^\perp) = \{0\}$ so that $F \subset (F^\perp)^\perp$.

Proposition 20.4.0.2. *Suppose b_σ is non-degenerate and let F, G be subspaces of E .*

1. $\dim F + \dim F^\perp = \dim E$,
2. $F = (F^\perp)^\perp$,
3. $(F + G)^\perp = F^\perp \cap G^\perp$ and $(F \cap G)^\perp = F^\perp + G^\perp$.
4. *Orthogonality is decreasing.*
5. *The following conditions are equivalent*
 - (a) $F \cap F^\perp = \{0\}$
 - (b) b_σ non-degenerate on F
 - (c) b_σ non-degenerate on F^\perp

PROOF. By definition, $\widehat{b}_\sigma(F)$ is the orthogonal in duality of F in E_σ^* . The first two formulas follow from the dimension calculation of the orthogonal in duality (19.4.1.1). The last two are formal (but useful). For 5), the equivalence of a) and b) is tautological. But by invoking this equivalence for F^\perp , we conclude the equivalence of a) and b) thanks to 2).

The proof justifies *a posteriori* the abuse of notation \perp for the orthogonal which generally does not lead to confusion.

Definition 20.4.0.3. *We say that a vector x is isotropic if $b_\sigma(x, x) = 0$. We denote by $\mathcal{C}(b_\sigma)$ the cone of isotropic vectors; it contains $\text{Ker}(b_\sigma)$.*

A vector space F is isotropic if $F \cap F^\perp \neq \{0\}$, totally isotropic if $F \subset F^\perp$. It is anisotropic if $F \cap F^\perp = \{0\}$. Finally, we say that F is totally isotropic if $F \subset F^\perp$, in other words, if $b_\sigma|_{F \times F} \equiv 0$.

Example(s) 20.4.0.4. *If $E = \mathbf{R}^n$ and $b(X, Y) = x_1y_2 + y_1x_2$, then all vectors of the canonical basis are isotropic.*

Remark(s) 20.4.0.5. *If F is isotropic, then $F \cap F^\perp$ is totally isotropic.*

If b_σ is non-degenerate and F is totally isotropic, then $\dim F \leq \dim E/2$ (20.4.0.2); however, if F is anisotropic, then we have $E = F \oplus F^\perp$.

Note that in general, we obviously have $\text{rad}(E_1 \oplus E_2) = \text{rad}(E_1) \oplus \text{rad}(E_2) =$ which will allow us to systematically reduce to the non-degenerate case.

Proposition 20.4.0.6. *Suppose b_σ is non-degenerate and let $a \in \text{End}(E)$ and a^* its adjoint. Then we have*

1. $\text{Ker}(a^*) = \text{Im}(a)^\perp$.
2. $\text{Im}(a^*) = \text{Ker}(a)^\perp$.

Proof. We can deduce it formally from the analogous property for the transpose. Here is a direct proof (which is a copy of the proof by duality). First, it suffices to prove one of the two formulas (replace a by a^* and use the involutivity of adjunction and orthogonality). Then, $\text{Im}(a^*) = \text{Ker}(a)^\perp$ having the same dimension according to 7.7.0.2 and 20.4.0.2, it suffices to prove $\text{Im}(a^*) \subset \text{Ker}(a)^\perp$. Now, if $a(x) = 0$, we have $b_\sigma(x, a^*(y)) = b_\sigma(a(x), y) = 0$ and therefore $a^*(y) \in \text{Ker}(a)^\perp$. \square

20.5 Alternating Forms

We will give in the alternating case a first illustration of the principle of the introduction: much of the geometry is contained in dimensions ≤ 2 , with larger dimensions obtained by orthogonality.

20.5.1 Classification

So let b be an alternating bilinear form on V of finite dimension d .

Theorem 20.5.1.1. *Alternating forms are classified by their rank. A non-degenerate form has even rank $2n$ and its matrix in a suitable basis is*

$$W_n = \text{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)$$

Proof. By restricting to a complement of the kernel, we can assume b is non-degenerate and $\dim(V) > 0$. Let $x_1 \in V$ be non-zero. Since the kernel of b is null, we can choose y_1 such that $b(x_1, y_1)$ is non-zero, so that $\Pi = \langle x_1, y_1 \rangle$ is a plane (if y_1 were collinear with x_1 we would have $b(x_1, y_1) = 0$). The dimension of the orthogonal of Π is $n - 2$. If $z = \alpha x_1 + \beta y_1 \in \Pi^\perp$, we have $0 = b(z, x_1) = -\beta$ and $0 = b(z, y_1) = -\alpha$, so $z = 0$, and we have an orthogonal decomposition $V = \Pi \oplus \Pi^\perp$ where the restriction of b to Π^\perp is non-degenerate (19.3.1.2). We conclude by a straightforward recurrence. \square

In particular, non-degenerate alternating forms only exist in even dimensions.

Exercise(s) 20.5.1.2. *Using the principle of extension of algebraic identities, show that the determinant of any odd-sized alternating matrix with coefficients in a ring is zero.*

Exercise(s) 20.5.1.3. *State and prove a result analogous to 20.5.1.1 in the anti-Hermitian case.*

20.5.2 Pfaffian



Johann Friedrich Pfaff

Theorem 20.5.1.1 proves that an invertible alternating matrix is of size $2n$ and is congruent to W_n . Since its determinant is 1, the determinant of any alternating matrix is a square in \mathbf{k} . This is universally true. Let $R = \mathbf{Z}[T_{i,j}], 1 \leq i < j \leq 2n$ be the ring of polynomials with integer coefficients in $(n(2n - 1))$ indeterminates. It is a unique factorization domain, and we denote K its field of fractions (11.3.3.1). Let $M \in M_n(K)$ be the matrix with coefficients $\text{sign}(i - j)T_{i,j}$. It is a polynomial alternating matrix and therefore defines a matrix function on the alternating matrices of $M_n(\mathbf{k})$.

Proposition 20.5.2.1. *There exists a unique polynomial $\text{Pf}(M) \in R$ with square $\det(M)$ and which equals 1 when $M = W_n$.*

Proof. If we had a second polynomial Q satisfying the proposition, we would have $Q^2 = \text{Pf}^2$ and thus $Q = \pm \text{Pf}$ by the integrality of R . But by looking at the value on W_n , we conclude $Q = \text{Pf}$. For existence, observe that $\det(M) \in K$ is non-zero (because it is true when $M = W_n$). So let $P \in \text{GL}_{2n}(K)$ such that ${}^tPW_nP = M$ (20.5.1.1). Then we have $\det(M) = (\det(P))^2$ with $\det(M) \in K$ and $\det(P) \in K^*$. Write the decomposition $\det(M) = \prod p_i^{v_i}$ into irreducible factors in the unique factorization domain R , and similarly, by writing those of the numerators and denominators of $\det(P)$, write $\det(P) = u \prod p_i^{w_i}$ with u invertible in R . Then we have $v_i \geq 0$, $w_i \in \mathbf{Z}$, and by uniqueness of the decomposition, $2w_i = v_i \geq 0$, $u^2 = 1$. Then we set $\text{Pf} = \pm u \prod p_i^{w_i} \in R$ choosing the sign to have $\text{Pf}(W_n) = 1$. \square

Symplectic geometry is the study of properties that are invariant under the symplectic group $\text{Sp}_{2n}(\mathbf{k})$ of matrices P preserving W_n , i.e., such that ${}^tPW_nP = W_n$. It is very rich, full of open questions but goes beyond the chosen framework.

20.6 Supplementary exercises

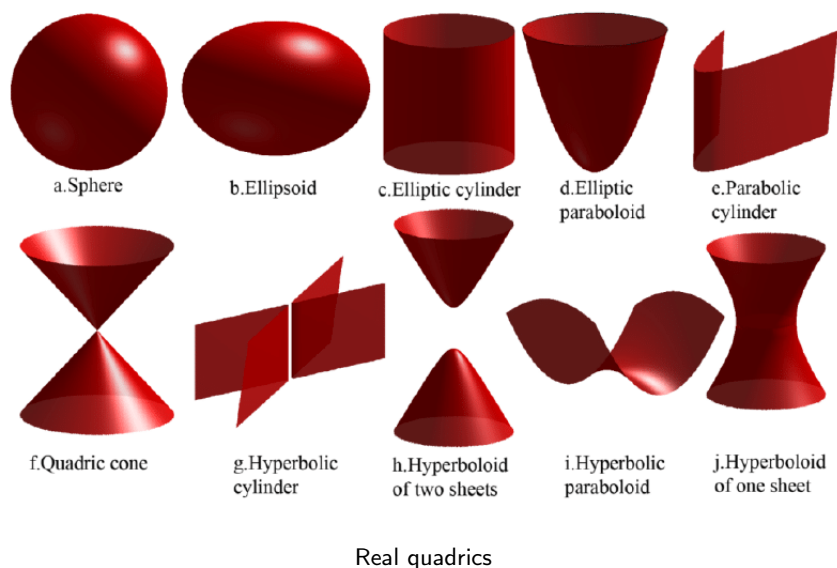
Exercise(s) 20.6.0.1. Let A be an alternating matrix of size $2n$. Prove the formula

$$\text{Pf}(A) = \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} \varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(2i-1), \sigma(2i)}$$

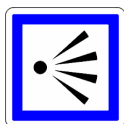
where $\varepsilon(\sigma)$ is the sign of the permutation $\sigma \in S_{2n}$.

Chapter 21

Quadratic Forms



21.1 Perspective



Quadratic forms in finite dimension can be seen as homogeneous polynomials of degree 2 in n variables or¹ as symmetric bilinear forms over \mathbf{K}^n -or their associated matrix-. The linear group acts on the former by variable change and on the latter by congruence. We will use both perspectives for their study, a study that heavily depends on the field \mathbf{K} unlike ordinary linear algebra.

We provide a general definition. If the theory of quadratic forms in characteristic 2 is useful and interesting, it differs significantly from the case of characteristic $\neq 2$. Therefore,

¹In characteristic different from 2 at least.

Unless otherwise stated, \mathbf{k} denotes from 21.2 a field of characteristic different from 2.

Definition 21.1.0.1. An map $q : E \rightarrow \mathbf{k}$ is a quadratic form if

1. q is homogeneous of weight 2, i.e., for all $x \in E, \lambda \in \mathbf{k}$, $q(\lambda x) = \lambda^2 q(x)$;

2. the map ${}^2b : \begin{cases} E \times E & \rightarrow & \mathbf{k} \\ (x, y) & \mapsto & q(x+y) - q(x) - q(y) \end{cases}$ is bilinear (symmetric).

Example(s) 21.1.0.2. Let $E = \mathbf{k}^n$ and $P \in \mathbf{k}[X_1, \dots, X_n]$ a homogeneous polynomial of degree 2 in n variables. Then, P defines a quadratic form $(x_i) \mapsto P(x_i)$ and conversely once a basis \mathcal{B} of E is given.

Note that we have ${}^2b(x, x) = 2q(x)$. Thus,

1. If $\text{char}(\mathbf{k}) = 2$, then b_2 is both symmetric and alternating. Note that $(x_i) \mapsto x_1^2$ on \mathbf{k}^n is a quadratic form, as is any homogeneous polynomial of degree 2 in n variables. However, ${}^2b(x, y) = (x_1 + y_1)^2 - x_1^2 - y_1^2 = 0$, so q and 0 have the same associated bilinear form!
2. If $\text{char}(\mathbf{k}) \neq 2$, then $q(x) = \frac{1}{2}{}^2b(x, x)$ being given q is equivalent to being given a symmetric bilinear form or a symmetric matrix.

21.2 Polar Form

Definition 21.2.0.1. The polar form of q is the symmetric bilinear form on $E \times E$ defined by $b(x, y) = \frac{1}{2}{}^2b(x, y) = \frac{1}{2}(q(x+y) - q(x) - q(y))$.

The notions defined for a bilinear form extend to quadratic forms. Thus, we will say quadratic space, quadratic space morphism, matrix of a quadratic form, etc., for bilinear space, matrix of a bilinear form, etc. An isomorphism of quadratic spaces is often called an isometry² by analogy with the usual Euclidean case. If q is the quadratic form on \mathbf{k}^n defined by $q(x_i) = \sum_{i \leq j} a_{i,j} x_i x_j$, $a_{i,j} \in \mathbf{k}$, its matrix $M(\mathcal{B}, q) = S$ in the canonical basis \mathcal{B} is defined by $S_{i,j} = a_{i,j}/2$ if $S_{i,i} = a_{i,i}$ with the formula

$$b(X, Y) = {}^t XSY$$

²Sometimes in the literature the term isometry is used for a morphism of quadratic spaces. We will not use it in this sense because in the degenerate case it can be confusing, such a morphism not necessarily being an isomorphism contrary to the usual Euclidean case from which the terminology is derived.

once E is identified with \mathbf{k}^n through \mathcal{B} . As before, a quadratic form q with matrix $M(\mathcal{B}, q) = S$ can therefore be seen as

1. a symmetric bilinear form b ;
2. a symmetric matrix S ;
3. a homogeneous polynomial of weight 2,

equivalent viewpoints that we will freely use. Recall (20.3.0.3), that q is invariant under a if and only if ${}^tASA = S$ with $A = \text{Mat}(\mathcal{B}, a)$.

Example(s) 21.2.0.2. • If (X, μ) is a measured space, then the formula

$$q(f) = \int f^2 d\mu$$

defines a quadratic form (in infinite dimension in general) on $L^2(X, \mu; \mathbf{R}) \rightarrow \mathbf{R}$ with polar form

$$b(f, g) = \int fg d\mu.$$

- If $\varphi_1, \dots, \varphi_r \in E^*$ and if $(\lambda_i) \in \mathbf{k}^r$, then the formula

$$q(x) = \sum_{1 \leq j \leq r} \lambda_j (\varphi_j(x))^2$$

defines a quadratic form on E with polar form

$$b(x, y) = \sum_{1 \leq j \leq r} \lambda_j \varphi_j(x) \varphi_j(y).$$

The reader will verify (*exercise*) that its rank is r as soon as the λ_i are non-zero and the φ_i are independent.

- If $M \in M_n(\mathbf{k})$, then $q(M) = \text{tr}({}^tMM)$ defines a quadratic form with polar form $b(M, N) = \text{tr}({}^tMN)$
- We define on $M_2(\mathbf{k})$ the form $q(M) = \det M$. We notice that q is a homogeneous quadratic polynomial in the coefficients of M . Moreover, (direct verification or Cayley-Hamilton theorem), we have

$$M^2 - (\text{tr } M) \cdot M + (\det M) \cdot I_2 = 0.$$

Taking the trace, we find

$$q(M) = \frac{(\text{tr } M)^2 - \text{tr } M^2}{2}.$$

Therefore, the associated polar form is

$$b(M, N) = \frac{(\text{tr } M)(\text{tr } N) - \text{tr}(MN)}{2}.$$

- If \mathbf{k} is a finite extension of \mathbf{Q} , the multiplication by $x \in \mathbf{k}$ defines a linear \mathbf{Q} endomorphism and therefore has a trace denoted $\text{tr}_{\mathbf{k}/\mathbf{Q}}(x)$. The map $x \mapsto \text{tr}_{\mathbf{k}/\mathbf{Q}}(x^2)$ is a bilinear form on the \mathbf{Q} -vector space \mathbf{k} , which can be shown without too much difficulty to be non-degenerate.

21.3 Orthogonal Bases

Recall (20.3.0.3) that a basis \mathcal{B} is orthogonal for q if and only if its matrix is diagonal or if in the associated coordinates we have $q(x) = \sum \lambda_i x_i^2$. This diagonal form is traditionally denoted by $\langle \lambda_1, \dots, \lambda_n \rangle$. Considering a diagonal congruence $\text{diag}(t_i)$ shows

$$\langle \lambda_1, \dots, \lambda_n \rangle \cong \langle t_1^2 \lambda_1, \dots, t_n^2 \lambda_n \rangle$$

and thus $\langle \lambda_1, \dots, \lambda_n \rangle$ depends only on the class of the λ_i in $\mathbf{k}^*/(\mathbf{k}^*)^2$.

Theorem 21.3.0.1 (Gauss). *Any quadratic space³ (E, q) admits an orthogonal basis (e_i) . The rank of q is then the number of indices i such that $q(e_i) \neq 0$. Moreover, such a basis can be obtained by the algorithm infra, called the Gauss algorithm.*

Proof. We proceed by induction on n . We can assume $n > 0$ and q non-zero. If q is degenerate, we take a non-zero vector e_n from the kernel and an orthogonal basis (by recursion) of the restriction of q to a complementary (necessarily orthogonal) which form an orthogonal basis. If q is non-degenerate, we then choose $e_{n+1} \in E$ such that $q(e_{n+1}) \neq 0$ so that the orthogonal $H = e_n^\perp$ is a hyperplane not containing e_n since it is non-isotropic (20.4.0.5). It suffices to complete e_n with an orthogonal basis of (H, q) (recursively). \square

Gauss Algorithm. We start with $q(X) = \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$ which we will recursively transform by changes of variables into $\sum \lambda_i \varphi_i^2$ where the φ form a basis of E^* .

At each step, the algorithm brings out a square and makes a coordinate disappear in the remaining form. Initialization. We can assume $n \geq 2$ (in dimension ≤ 1 we don't really have a choice...) and q not identically zero (in this case we take for φ_i the coordinates associated to the canonical basis for example and $\lambda_i = 0$).

Recursion.

- If there exists i such that $a_{i,i} \neq 0$, we can assume, by permuting the variables, that $a_{11} \neq 0$, we factor out x_1 in all possible monomials:

$$\begin{aligned} q(X) &= a_{11} \left(x_1^2 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_1 x_j \right) + \sum_{2 \leq i < j \leq n} a_{ij} x_i x_j \\ &= a_{11} \left(x_1 + \frac{1}{2} \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j \right)^2 - \frac{a_{11}}{4} \left(\sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j \right)^2 + \sum_{2 \leq i < j \leq n} a_{ij} x_i x_j \\ &= a_{11} \varphi_1(x)^2 + \sum_{2 \leq i < j \leq n} \alpha_{ij} x_i x_j \end{aligned}$$

³The proof is identical in the Hermitian case

with $\varphi_1(x) = x_1 + \frac{1}{2} \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j$ independent of the $n - 1$ forms $x_j, j \geq 2$ and we apply (recursively) the algorithm to $\sum_{2 \leq i \leq j \leq n} \alpha_{ij} x_i x_j$ which has only $n - 2$ variables left.

- If all square terms are zero ($a_{jj} = 0$), then, by permuting the variables, we can assume $a_{1,2} \neq 0$. We write

$$x_1 x_2 = \frac{(x_1 + x_2)^2 - (x_1 - x_2)^2}{4}.$$

We then set

$$\begin{cases} \varphi_1(x) = \frac{x_1 + x_2}{2} \\ \varphi_2(x) = \frac{x_1 - x_2}{2} \end{cases}$$

so that $q(x) = a_{1,2} \varphi_1(x)^2 - a_{1,2} \varphi_2(x)^2 + \sum_{2 \leq i \leq j \leq n} a_{ij} x_i x_j$ with $\varphi_1, \varphi_2, x_i, i \geq 3$ independent and we apply (recursively) the algorithm to \tilde{q} which has only $n - 2$ variables left.

Once we have the φ_i , their ante-dual basis (19.4.1.1) is the desired orthogonal basis.

Exercise(s) 21.3.0.2. *Implement (in SAGE for example) the Gauss algorithm. Discuss its numerical stability. Give a matrix version of the algorithm using only congruences by permutation matrices and transvections.*

Example(s) 21.3.0.3. *We consider on \mathbf{R}^3 the form*

$$q(x, y, z) = xy + yz + xz.$$

In the canonical basis \mathcal{B} , we have

$$\text{Mat}(q, \mathcal{B}) = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

We then set

$$\begin{cases} u = \frac{x+y}{2} \\ v = \frac{x-y}{2} \end{cases}$$

It follows that

$$q(x, y, z) = u^2 - v^2 + (u - v)z + (u + v)z = u^2 + 2uz - v^2;$$

thus,

$$q(x, y, z) = (u + z)^2 - v^2 - z^2.$$

Let \mathcal{B} denote the canonical basis, and

$$P^* = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

the change-of-basis matrix from \mathcal{B}^* to a basis \mathcal{C}^* the dual base of an orthogonal basis \mathcal{C} that we want to determine. We compute $P = ({}^tP^*)^{-1}$ and obtain

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

In this basis, $q(x, y, z) = x^2 - y^2 - z^2$.

21.4 Quadratic Spaces

In accordance with our strategy, let us study quadratic spaces.

Proposition 21.4.0.1. *Let (E, q) be a quadratic space of dimension 2.*

One of four possibilities:

1. q is anisotropic ($-\text{disc}(q) \notin (k^*)^2$);
2. there is exactly one isotropic line ($\text{disc}(q) = 0$), q is degenerate and in an appropriate basis, $q(x, y) = ax^2$, for $a \neq 0$;
3. there are exactly two isotropic lines ($-\text{disc}(q) \in (k^*)^2$), in an appropriate basis $q(x, y) = xy$ and in this case, it is said that E is hyperbolic.
4. there are at least three isotropic lines ($\text{disc}(q) = 0$) and q is identically zero.

PROOF. We may assume q is not identically zero.

By multiplying q by a non-zero scalar, q is then written in a suitable orthogonal basis⁴ $q(x, y) = x^2 - \lambda y^2$.

If λ is not a square ($\text{disc} = (q)$ non-square) and q is anisotropic.

Otherwise, write $\lambda = \mu^2$ and $q(x, y) = x^2 - \mu^2 y^2 = (x - \mu y)(x + \mu y)$.

If λ is zero, we are in case 2) - the vertical axis $x = 0$ is the only isotropic line - if λ is non-zero we are in case 3) - the lines with equations $x - \mu y = 0$ and $x + \mu y = 0$ being the only isotropic lines.

⁴Gauss' algorithm then reduces to the canonical factorization of a polynomial of degree 2! Indeed, the statement amounts to saying that a non-null second degree trinomial admits at most 2 roots...

With the vocabulary of 20.4.0.3, we deduce the following characterization of hyperbolic planes.

Proposition 21.4.0.2. *A quadratic plane is hyperbolic if and only if it is non-degenerate and isotropic or if its discriminant is a non-zero square.*

21.5 Anisotropic spaces

We can provide models of anisotropic spaces as follows. Let $\alpha \in \mathbf{k}^*$ such that $-\alpha$ is not a square and let $K = k[\sqrt{-\alpha}] = \mathbf{k}[T]/(T^2 + \alpha)$. It is a space of dimension 2 over \mathbf{k} with basis $\mathcal{B} = (1, \sqrt{-\alpha})$ and the endomorphism σ with matrix $\text{diag}(1, -1)$ is a field morphism (as is complex conjugation...). Naturally, the elements fixed by σ are exactly the elements of \mathbf{k} . We then define $N(z) = z\sigma(z) = x^2 - \alpha y^2$ for any $z = x + \sqrt{-\alpha}y \in K$: it is a quadratic form on K with values in \mathbf{k} whose matrix in \mathcal{B} is $\text{diag}(1, -\alpha)$. We will simply denote this quadratic plane as Π_α . According to 21.4.0.1, any anisotropic space is equivalent to $\Pi_{-\text{disc}(q)}$.

21.6 Invariants of Quadratic Forms

Definition 21.6.0.1. *Two spaces (E, q) and (E', q') are equivalent (\approx) if there exists an isometry $u : (E, q) \rightarrow (E', q')$, i.e. an isomorphism $u : E \rightarrow E'$ such that, for every $x \in E$, we have $q'(u(x)) = q(x)$, in other words, if the matrices of q and q' are congruent. An invariant is an application on the corresponding quotient $\{(E, q)\}/\approx$.*

We currently have two invariants by congruences of a quadratic form q (or a symmetric matrix): the rank $\text{rk}(M)$ and the discriminant $\text{disc}(q) = \det(M) \in k/k^{*2}$ (19.3) of a matrix of q . We will see that if \mathbf{k} is algebraically closed, the rank classifies the quadratic forms (21.8.0.1). In this case, $k/k^{*2} = \{0, 1\}$ and $\text{disc}(q) = \text{sign}(r)$.

We will see that in the case of finite fields, rank and discriminant classify quadratic forms (21.11.0.2).

In general, these two invariants are not sufficient. For example, the real forms in four variables $x^2 - y^2 - z^2 - t^2$ and $x^2 + y^2 - z^2 - t^2$ have the same rank but are not equivalent because their discriminants are -1 and 1 respectively, which are different in $\mathbf{R}/\mathbf{R}^{*2} = \{-1, 0, 1\}$. We will see (21.9.0.1) that a third invariant is nonetheless necessary, the index (21.7.0.1), these three invariants being summarized in the signature of the real quadratic form.

In all these cases, there are only a finite number of equivalence classes. This is not true in general.

Let us give an example. We define, for every prime number p the quadratic form

$$q_p(x) = \sum_{1 \leq j < n} x_j^2 + px_n^2$$

on \mathbf{Q}^n . They are pairwise non-equivalent as soon as $n > 0$ and differ by their discriminant which is $p \pmod{(\mathbf{Q}^{*2})}$ (cf. the exercise 21.14.0.3). It is possible to classify generally over \mathbf{Q} but new invariants related to the classification over finite fields, the Hilbert symbols (cf. the magnificent work [29]) are necessary.

In general, classification is an extremely difficult problem. This can be seen in reverse: quadratic forms allow defining subtle field invariants (cf. the exercise 21.14.0.2).

21.7 Isotropy and Index

We will define a third invariant: the index.

Definition 21.7.0.1. *The index ν of a quadratic form q is the maximum dimension of totally isotropic spaces. If $\nu = 0$ i.e., if $q(x) = 0 \Rightarrow x = 0$, it is said that q is anisotropic or defined.*

For example, if $\mathbf{k} = \mathbf{R}$, a continuity argument assures if q is defined, then, either for every $x \neq 0$ we have $q(x) > 0$, or, for every $x \neq 0$ we have $q(x) < 0$.

We will deduce from this the general decomposition of a space (E, q) . We start with a lemma.

Lemma 21.7.0.2. *Let (E, q) be a non-degenerate quadratic plane.*

1. *If x is isotropic, there is a hyperbolic plane containing x .*
2. *The index ν of $\bigoplus_{1 \leq j \leq r} P_j$ is r .*
3. *There are r hyperbolic planes P_j and (F, q) is anisotropic such that*

$$E = \left(\bigoplus_{1 \leq j \leq r} P_j \right) \perp F.$$

4. *If such a decomposition exists, then $r = \nu$.*

If q is arbitrary, we have a decomposition

$$E = \text{rad}(E) \perp \left(\bigoplus_{1 \leq j \leq r} P_j \right) \perp F$$

with $r + \dim(\text{rad}(E)) = \nu$.

PROOF. There exists y such that $b(x, y) \neq 0$. Consequently, x and y indeed generate a quadratic plane whose discriminant $-b(x, y)^2 \neq 0$. Since it has an isotropic line, it is a hyperbolic plane (21.4.0.1) hence 1).

For 2), we may assume q is isotropic and $n \geq 3$ (according to 21.4.0.1). Let then P be a hyperbolic plane contained within E . We show $E = P \oplus P^\perp$. Since q is non-degenerate, the dimensions are correct. If $v \in P \cap P^\perp$ then $b(v, P) = \{0\}$ contradicting the non-degeneracy of isotropic spaces. This contradicts that q is non-degenerate. Therefore $E = P \oplus P^\perp$ and we apply the recursion hypothesis to P^\perp .

For 3), denote e_i, e'_i a basis for the two isotropic lines of P_i . Obviously, $\nu \geq r$ since $\text{Span}(e_i)$ is totally isotropic of dimension r . As $\text{Span}(e_i + e'_i)$ is anisotropic of codimension r , we also have $\nu \geq r - r'$.

For 4), let G be isotropic of dimension ν and denote p the orthogonal projection onto F (parallel to $\bigoplus_{1 \leq j \leq r} P_j$). The space $p(G)$ is isotropically constructed therefore null since F is anisotropic. It follows that G is included in $\bigoplus_{1 \leq j \leq r} P_j$ and thus $\nu \leq r$ according to 3). Conversely, since $\text{Span}(e_i)$ is isotropic, we have $r \leq \nu$.

The last point follows from the previous ones by the additivity of radicals by orthogonal direct sum.

Theorem 21.7.0.3 (Witt's Simplification). Let (E, q) be a quadratic plane and

$$E = \text{rad}(E) \oplus \left(\bigoplus_{1 \leq j \leq \nu} P_j \right) \oplus F$$

as in lemma 21.7.0.2.

1. The quadratic isomorphism class of F is well determined and $r = \nu$.
2. If there is an isomorphism of quadratic planes

$$E \oplus F \cong E \oplus F'$$

then

$$F \cong F'.$$

3. If two families of non-zero scalars satisfy

$$\langle a, a_1, \dots, a_n \rangle \cong \langle a, b_1, \dots, b_n \rangle$$

then

$$\langle a_1, \dots, a_n \rangle \cong \langle b_1, \dots, b_n \rangle$$

PROOF. We will show 3) then the two implications 3) \Rightarrow 2) and 2) \Rightarrow 1).

Proof of 3). We denote x_i , $0 \leq i \leq n$ the coordinate linear forms on \mathbf{k}^{n+1} , in other words, the dual basis of the canonical basis (e_i) . Saying

$$\langle a, a_1, \dots, a_n \rangle \cong \langle a, b_1, \dots, b_n \rangle$$

means the existence of independent linear forms $\varphi_i \in (\mathbf{k}^{n+1})^*$ (the rows of the matrix defining the congruence) such that

$$(*) \quad ax_0^2 + \sum_{i \geq 1} a_i x_i^2 = a\varphi_0(x)^2 + \sum_{i \geq 1} b_i \varphi_i^2$$

At least one of the two linear forms $x_0 \pm \varphi(x)$ does not cancel e_0 : let $\varepsilon = \pm 1$ such that $\langle x_0 + \varepsilon\varphi_0(x), e_0 \rangle = \lambda \neq 0$ so that $x_0 + \varepsilon\varphi_0(x)$ is written as

$$x_0 + \varepsilon\varphi_0(x) = \lambda x_0 - \lambda\psi(x_1, \dots, x_n)$$

with ψ a linear form on \mathbf{k}^n . If we substitute $x_0 = \psi(x_1, \dots, x_n)$ in (*), we therefore have

$$\sum_{i \geq 1} a_i x_i^2 = \sum_{i \geq 1} b_i \varphi_i(\psi(x_1, \dots, x_n), x_1, \dots, x_n)^2 = \sum_{i \geq 1} b_i \tilde{\varphi}_i(x_1, \dots, x_n)^2$$

where $\tilde{\varphi}$ are linear forms on \mathbf{k}^n . If $\Psi \in M_n(\mathbf{k})$ is the matrix they define, then we have $\text{diag}(a_i) = {}^t\Psi \text{diag}(b_i)\Psi$ ensuring the invertibility of Ψ by taking the determinants and therefore

$$\langle a_1, \dots, a_n \rangle \approx \langle b_1, \dots, b_n \rangle$$

thus 3).

3) \Rightarrow 2) We note q, q' the associated quadratic forms, r the rank of the restriction of q to E , and ρ, ρ' those of the restrictions of q, q' to F, F' . As the sums are orthogonal, we have $\text{rank}(q) = r + \rho = r + \rho'$ so that $\rho = \rho'$ and of course $\dim(F) = \dim(F')$. The radicals of q and q' being generated by the vectors of the corresponding orthogonal bases with indices such as $a_i = b_j = 0$ and $a_i = b'_j = 0$, we can assume by passing to the quotient by the radicals that the forms are non-degenerate. We thus have

$$\langle a_1, \dots, a_r, b_1, \dots, b_\rho \rangle \approx \langle a_1, \dots, a_r, b'_1, \dots, b'_\rho \rangle$$

so that

$$\langle b_1, \dots, b_\rho \rangle \approx \langle b'_1, \dots, b'_\rho \rangle$$

according to 3) which proves 2).

2) \Rightarrow 1) As above, we reduce to the non-degenerate case by passage to the quotient. Assume that

$$(\bigoplus_{1 \leq j \leq r} P_j) \oplus F \approx (\bigoplus_{1 \leq j \leq r'} P'_j) \oplus F'$$

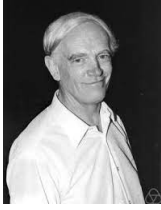
with F, F' anisotropic, the P_i, P'_j hyperbolic and for example $r' \leq r$. We have already seen in lemma 21.7.0.2 $r = \nu = r'$. Thanks to 2), we then deduce $F \approx F'$.

The key to this important result is thus point 3), the rest being quite formal. But ultimately it's just a very simple calculation but which to my knowledge only appeared recently ([11]). The classic proof is proposed as an exercise below (21.14.0.1). Naturally, according to the previous discussion, every quadratic plane (degenerate or not) decomposes into

$$E = \text{Ker}(q) \oplus (\bigoplus_{1 \leq j \leq \nu} P_j) \oplus F$$

with F well defined up to isometry (and anisotropic).

Thus, we have defined a fourth invariant: the anisotropic part $(E, q)^{\odot} = (F, q) \bmod \approx$ of (E, q) .



Ernst Witt

The character of Witt is controversial due in part, at a minimum, to his active collaboration with the Nazi regime (a member of the Nazi party from 1933 then SA). The complex character of Witt however seems to somewhat mitigate his actions. See <https://mathshistory.st-andrews.ac.uk/Biographies/Witt/> for a biography.

21.8 Classification over an algebraically closed field

Given a basis \mathcal{B} of E of dimension n , as always $(x_j) = [x]_{\mathcal{B}}$ denote the coordinates of x in \mathcal{B} .

Theorem 21.8.0.1. *If \mathbf{k} is algebraically closed, then, for any q , there exists a basis \mathcal{B} such that*

$$q(x) = \sum_{1 \leq j \leq \text{rk}(q)} x_j^2.$$

There are exactly $n + 1$ classes of equivalences, which are differentiated by the rank of q .

PROOF. *In an orthogonal basis $\mathcal{B} = (e_1, \dots, e_n)$, we have*

$$q(x) = \sum_{1 \leq j \leq \text{rk}(q)} \lambda_j x_j^2,$$

with $\lambda_j \in \mathbf{k}^$. Since \mathbf{k} is algebraically closed, there exists $\mu_j \in \mathbf{k}$ such that $\mu_j^2 = \lambda_j$. By setting $f_j = e_j/\mu_j$ if $i \leq \text{rk}(q)$ and $f_j = e_j$ otherwise, we obtain*

$$q(x) = \sum \lambda_j x_j^2 = \sum (\mu_j x_j)^2 = \sum \xi_j^2$$

with $(\xi_j = \mu_j x_j)$ the coordinates of x in the base (f_j)

Exercise(s) 21.8.0.2. *Verify in this case the formulas $\nu = \dim(\text{rad}(q)) + [\text{rk}(q)/2]$ and $E^{\oplus} = \{0\}$ or $E^{\oplus} \cong \langle 1 \rangle$ depending on the parity of the rank.*

21.9 Classification over \mathbf{R}

Theorem 21.9.0.1 (Sylvester's Inertia Theorem). *If $\mathbf{k} = \mathbf{R}$, then for every q , there exists a unique pair of natural numbers (s, t) called the signature of q such that there exists a basis in which*

$$q(x) = \sum_{1 \leq j \leq s} x_j^2 - \sum_{s+1 \leq j \leq \text{rk}(q)} x_j^2.$$

We then have:

1. $s + t = \text{rk}(q)$.
2. $s + \dim(\text{Ker}(q)) = \max\{\dim(F) \mid q|_F \geq 0\}$.
3. $t + \dim(\text{Ker}(q)) = \max\{\dim(F) \mid q|_F \leq 0\}$.

There are exactly $(\dim E + 1)(\dim E + 2)/2$ equivalence classes, which are distinguished by the signature of q .

Proof. By separating the vectors into a suitable orthogonal basis $\mathcal{B} = (e_1, \dots, e_n)$, we have in an orthogonal basis $\mathcal{B} = (e_1, \dots, e_n)$,

$$q(x) = \sum_{1 \leq j \leq \text{rk}(q)} \lambda_j x_j^2,$$

with $\lambda_j \in \mathbf{R}^*$. We can assume that the first s scalars are positive and the last t are negative. We then set $\mu_j = \sqrt{|\lambda_j|}$ and $f_j = e_j$, otherwise we get

$$q(x) = \sum_{1 \leq j \leq s} \xi_j^2 - \sum_{s+1 \leq j \leq \text{rk}(q)} \xi_j^2$$

with $(\xi_j = \mu_j x_j)$ being the coordinates of x in the basis (f_j) . It remains to show items 1), 2), and 3), the rest follows immediately. Item 1) is clear and item 3) follows from 2) by changing q to $-q$.

Now, let $F' = \text{Span}\{e_1, \dots, e_s\} \oplus \text{Ker}(q)$ and $G = \text{Span}\{e_{s+1}, \dots, e_n\}$. Since $n = \dim(\text{Ker}(q)) + \text{rk}(q)$, we have $\dim(G) = t$. Since q is ≥ 0 on F' , we have $s + \dim(\text{Ker}(q)) \leq \max\{\dim(F) \mid q|_F \geq 0\}$. Conversely, if there exists F of dimension $> s + \dim(\text{Ker}(q)) = n - t$, we would have $\dim(F \cap G) > 0$ and thus a vector x such that $q(x) \geq 0$ because $x \in F$ and $q(x) < 0$ because $x \in G - \{0\}$. \square

Exercise(s) 21.9.0.2. Verify in this case the formulas $\nu = \dim(\text{rad}(q)) + \inf(p, q)$ and $\dim(E^\ominus) = |s - t|$ with $E^\ominus = \langle \text{signe}(s - t), \dots, \text{signe}(s - t) \rangle$ if $s \neq t$.

21.10 Conics and Quadrics in \mathbf{R}^2 and \mathbf{R}^3 , Ellipsoid

A conic is given by an equation of the form $q(x, y) = 1$, where q is a homogeneous polynomial of degree 2. In other words, q is a quadratic form. According to theorem 12.6.3.1, there exists an orthonormal basis of \mathbf{R}^2 such that P has a canonical form, which gives us the notion of ellipse, hyperbola...

We define a quadric in \mathbf{R}^3 as the set

$$\mathcal{Q} = \{(x, y, z) \in \mathbf{R}^3, q(x, y, z) = 1\}$$

where q is a non-degenerate quadratic form. We discuss according to the signature of q the form of the quadric.

$\text{sig}(q)=(0,3)$ In an adapted basis of \mathbf{R}^3 , we have $q(x, y, z) = -x^2 - y^2 - z^2$, thus $\mathcal{Q} = \emptyset$.

$\text{sig}(q)=(1,2)$ In an adapted basis of \mathbf{R}^3 , we have $q(x, y, z) = x^2 - y^2 - z^2$. Thus \mathcal{Q} has two connected components depending on whether $x \geq 1$ or $x \leq -1$. The quadric intersects the plane $\{x = \text{const}\}$, for $|x| \geq 1$, in a circle of radius $x^2 - 1$. We say that \mathcal{Q} is a two-sheeted hyperboloid.

$\text{sig}(q)=(2,1)$ In an adapted basis of \mathbf{R}^3 , we have $q(x, y, z) = x^2 + y^2 - z^2$. Thus \mathcal{Q} is connected. The quadric intersects the plane $\{z = 0\}$ in a circle. We say that \mathcal{Q} is a one-sheeted hyperboloid.

An important property of this quadric is that it is exactly doubly ruled.

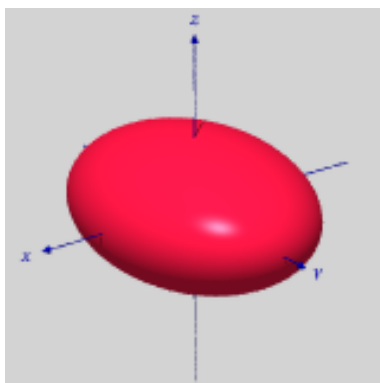
A point belongs to \mathcal{Q} if $(y - z)(y + z) = (1 - x)(1 + x)$. We can guess the equation of two families of lines included in \mathcal{Q} .

$$\Delta_a \begin{cases} y - z = a(1 - x) \\ (y + z)a = 1 + x \end{cases} \quad a \in \mathbf{R} \quad \text{and} \quad \Delta_\infty \begin{cases} y = -z \\ x = 1 \end{cases}$$

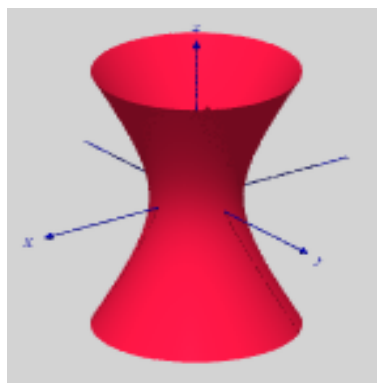
and also

$$D_b \begin{cases} y + z = b(1 - x) \\ (y - z)b = 1 + x \end{cases} \quad b \in \mathbf{R} \quad \text{and} \quad D_\infty \begin{cases} y = z \\ x = 1 \end{cases}$$

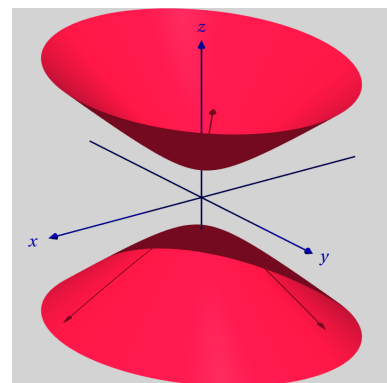
A simple calculation shows that these families are transverse, and that only one line per family passes through a given point of \mathcal{Q} .



ellipsoid



one-sheeted hyperboloid



two-sheeted hyperboloid

21.11 Classification over finite fields.

If \mathbf{k} is a finite field (of characteristic different from 2), we know that its cardinality is of the form p^d where $p \neq 2$ is its characteristic -and thus is prime- (simply because it is a \mathbf{F}_p -vector space). According to the general strategy, let us see what happens in the crucial case of dimension ≤ 2 starting with dimension 1, namely understanding the discriminant.

Lemma 21.11.0.1. *The multiplicative group $\mathbf{F}_{p^d}^*/(\mathbf{k}^*)^2$ has two elements $1 \neq \alpha$ so that the discriminant of a quadratic form over \mathbf{k} is valued in the set $\{0, 1, \alpha\}$. There are exactly two non-degenerate quadratic planes equivalent to $\langle 1, 1 \rangle$ and $\langle 1, \alpha \rangle$ distinguished by the discriminant.*

PROOF. We have an exact sequence $\{1\} \rightarrow \text{Ker}(sq) \rightarrow \mathbf{k}^* \xrightarrow{sq} (\mathbf{k}^*)^2 \rightarrow \{1\}$ of multiplicative groups where sq is the squaring morphism. As \mathbf{k} is integral, the equation $x^2 = 1$ has solutions ± 1 . Since $p \neq 2$, we have $1 \neq -1$. We deduce $\text{Card}((\mathbf{k}^*)^2) = \frac{q-1}{2}$ and $\text{Card}(\mathbf{k}^*/(\mathbf{k}^*)^2) = 2$ which proves the first point. Then, consider a non-degenerate quadratic plane of form q . We may assume $q(x_1, x_2) = ax_1^2 + bx_2^2$ with $ab \neq 0$. Thus, there are $1 + \frac{q-1}{2} = \frac{q+1}{2}$ squares in \mathbf{k} so that the cardinalities of $\{at^2, x \in \mathbf{k}\}$ and $\{1 - bt^2, y \in \mathbf{k}\}$ are $\frac{q+1}{2}$ and therefore have at least one intersection point that defines $e_1 = (x_1, x_2)$ such that $q(x_1, x_2) = 1$. Then, let e_2 be a basis of the orthogonal of e_1 . Writing q in this basis, we have $q \approx \langle 1, \text{disc}(q) \rangle$ (cf. 21.3 for invariance by congruence of the notation).

Here is then.

Theorem 21.11.0.2. *If \mathbf{k} is finite with characteristic $p \neq 2$, then every non-degenerate form is (uniquely) equivalent to either $\langle 1, \dots, 1 \rangle$ or $\langle 1, \dots, 1, \alpha \rangle$. These classes are distinguished by their discriminants.*

PROOF. We proceed by induction on n . We may assume $n \geq 3$ and (e_i) an orthogonal basis. As the quadratic plane $\text{Span}(e_1, e_2)$ is non-degenerate, we can choose ε_1 in this plane $q(\varepsilon_1) = 1$. Then $E = \mathbf{k}\varepsilon_1 \oplus \varepsilon_1^\perp$ such that ε_1^\perp is a hyperbolic plane to which we can apply the induction hypothesis. The second point is clear.

Exercise(s) 21.11.0.3. Let \mathbf{k} be finite with characteristic $p \neq 2$ and q non-degenerate on E .

1. Using $\text{Card}(\mathbf{k}^*) = \frac{p^d-1}{2}$, show that -1 is a square if and only if $p \equiv 1 \pmod{4}$.
2. Calculate $\nu(E)$ and E^{\oplus} depending on $n, \text{disc}(q)$ and $p \pmod{4}$.

21.12 Witt's Extension Theorem

Theorem 21.12.0.1 (de Witt). *Let $u : F \rightarrow E$ be an injective morphism of quadratic planes with E non-degenerate. Then, there exists $\tilde{u} \in O(q)$ such that $\tilde{u}|_F = u$.*

PROOF. The injectivity of the morphism u ensures $F \approx u(F)$. We proceed by induction on $\text{codim}(F)$. We may assume $\text{codim}(F) > 0$.

- If $q_{\mathbb{F}}$ is non-degenerate, i.e. $\mathbf{K} = \mathbb{F} \cap \mathbb{F}^\perp = \{0\}$, then $\mathbb{E} = \mathbb{F} \oplus \mathbb{F}^\perp$. If $u(f) \in u(\mathbb{F}) \cap u(\mathbb{F}^\perp)$, we have $0 = b(u(f), u(\mathbb{F}^\perp)) = b(f, \mathbb{F}^\perp) = \{0\}$ thus $f = 0$ and so $u(\mathbb{F}) \cap u(\mathbb{F}^\perp) = \{0\}$ such that

$$\mathbb{F} \oplus \mathbb{F}^\perp = u(\mathbb{F}) \oplus u(\mathbb{F}^\perp) \approx \mathbb{F} \oplus u(\mathbb{F}^\perp)$$

so that $\mathbb{F}^\perp \approx u(\mathbb{F}^\perp)$ according to Witt's Simplification Theorem 21.7.0.3. We can therefore choose isomorphisms of quadratic planes $u^\perp : \mathbb{F}^\perp \rightarrow u(\mathbb{F}^\perp)$ and $\tilde{u} \oplus u^\perp \in \mathcal{O}(q)$ fits.

- If $q_{\mathbb{F}}$ is degenerate, choose x non-null in $\mathbb{F} \cap \mathbb{F}^\perp$ and an isotropic space \mathbb{P} that contains it (21.7.0.2). If y directs the second isotropic line of \mathbb{P} , we have $b(x, y) = 1$ so that $y \notin \mathbb{F}$ and

$$\text{codim}(\mathbb{G} = \mathbb{F} \oplus \mathbf{k}y) = \text{codim}(\mathbb{F}) - 1$$

It suffices to extend u to \mathbb{G} (preserving b). We then seek $z \in \mathbb{E}$ such that

1. $b(u(f), z) = b(f, y)$ for all $f \in \mathbb{F}$
2. $b(u(x), z) = 1$
3. $b(z, z) = 0$

which ensures that the extension of u defined by $\tilde{u}(y) = z$ is an isometry (on its image). Note that 1) \Rightarrow 2) and changing z by $z + \lambda u(x)$ does not change 1) because $x \in \mathbb{F}^\perp$. Yet,

$$b(z + \lambda u(x), z + \lambda u(x)) = b(z, z) + 2\lambda b(u(x), z) = b(z, z) + 2\lambda b(x, y) = b(z, z) + 2\lambda$$

so that by changing z by $z - \frac{b(z, z)}{2}x$, we will have the sought extension. It suffices therefore to find z fulfilling 1).

Let then \mathbb{S} be an arbitrary complement of $u(\mathbb{F})$ in \mathbb{E} and $\varphi \in \mathbb{E}^*$ the linear form null on \mathbb{S} and equaling $u^{-1}(t), y$ for any $t \in u(\mathbb{F})$ - recalling that u is assumed injective therefore bijective from \mathbb{F} to $u(\mathbb{F})$ -. Since b is non-degenerate, there exists $z \in \mathbb{E}$ such that $\varphi = b(\cdot, z)$ and z fulfills 1).

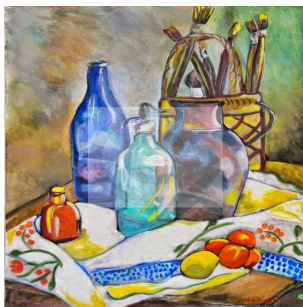
21.13 Appendix: Quadratic Pencils

The classification of quadratic forms allows for the study of quadrics as seen in the real case (or in the complex case -cf. tutorial-). It is the first step towards understanding their arithmetic. Here, we focus on the intersections of two quadrics, one of which is non-degenerate (a relatively mild assumption -why?-). For another proof, see V..

Let q_1, q_2 be two quadratic forms on a \mathbf{k} -vector space V of finite dimension n with q_1 non-degenerate. Denote b_1, b_2 as the associated bilinear forms and $S_\ell = (b_\ell(e_i, e_j))_{i, j}, \ell = 1, 2$ the matrices in a basis $\mathcal{B} = (e_i)$ chosen arbitrarily. Thus $S_1 \in \text{GL}_n(\mathbf{k})$.

If $k = \mathbf{R}$ and q_1 is definite, the reduction theorem 12.6.3.1 ensures that we can find a basis such that the matrix of q_1 is the identity and that of q_2 is diagonal, thus reducing to a «diagonal pencil»):

$$q_1(x) = \sum x_i^2 \text{ and } q_2(x) = \sum \lambda_i x_i^2.$$



Pencils

What about in the general case? Generally speaking, we will talk about a diagonal pencil for

$$q_1(x) = \sum \mu_i x_i^2 \text{ and } q_2(x) = \sum \lambda_i x_i^2.$$

Thus (q_1, q_2) is a diagonal pencil in a suitable basis if and only if there is a basis of co-orthogonalization.

Theorem 21.13.0.1. *With the notations above, q_1, q_2 is a diagonal pencil in a suitable basis if and only if $S_1^{-1}S_2$ is diagonalizable over \mathbf{k} . This is particularly the case if*

$$\text{Card}\{\lambda \in \mathbf{k} | q_2 + t q_1 \text{ is degenerate}\} = n.$$

As we have seen, generally, $S_1^{-1}S_2$ is not diagonalizable, even with $S_1 = \text{Id}$ (cf. 12.6.3.2). Note also that if \mathbf{k} is algebraically closed, we can take $\mu_i = 1$ (changing e_i to $e_i/\sqrt{\mu_i}$), the eigenvalues of $S_1^{-1}S_2$ are the quotients λ_i/μ_i .

Proof. Let u be the endomorphism defined by $S_1^{-1}S_2$: it is self-adjoint for q_1 (cf. 19.4.1.2).

Suppose u is diagonalizable. We proceed by induction on n , the case $n = 1$ being trivial. Assume $n > 1$ and the proposition true in dimension $< n$.

If u is a homothety μId , then $q_2 = \mu q_1$ and (q_1, q_2) is a diagonal pencil in any orthogonal basis of q_1

Suppose then that u is not a homothety. The eigenspaces of u are orthogonal to each other for q_1 : if $x \in \text{Ker}(u - \lambda \text{Id})$, $x' \in \text{Ker}(u - \lambda' \text{Id})$ with $\lambda \neq \lambda'$, we have

$$\lambda'(x, x')_1 = (x, u(x'))_1 = (u^*(x), x')_1 = (u(x), x')_1 = \lambda(x, x')_1$$

and thus $(x, x')_1 = 0$. By taking for each eigenspace an orthogonal basis q_1 , we thus obtain an orthogonal basis $q_1 \bar{\mathcal{B}} = (\bar{e}_i)$ of eigenvectors of u . Let P be the passage matrix from \mathcal{B} to $\bar{\mathcal{B}}$ (the columns of P are the coordinates of the vectors of $\bar{\mathcal{B}}$ relative to \mathcal{B}). We have

$$\text{Mat}_{\bar{\mathcal{B}}}(u) = P^{-1}S_1^{-1}S_2P = (P^{-1}S_1^{-1}P^{-1})({}^tPS_2P) = ({}^tPS_1P)^{-1}({}^tPS_2P) = \bar{S}_1^{-1}\bar{S}_2.$$

with \bar{S}_ℓ matrices of q_ℓ in \mathcal{B}' . By construction, both \bar{S}_1 and $\text{Mat}_{\bar{\mathcal{B}}}(u) = \bar{S}_1^{-1}\bar{S}_2$ are diagonal and thus so is \bar{S}_2 which is their product.

The converse is clear because if the pencil is diagonal considering $\overline{\mathcal{B}}$ a basis \mathcal{B} of diagonalization, both \overline{S}_1 and \overline{S}_2 are diagonalized and thus the same for $\overline{S}_1^{-1}\overline{S}_2$, which is similar to $S_1^{-1}S_2$ as we just saw.

The last point means that u has n distinct eigenvalues: u is indeed diagonalizable. \square

21.14 Supplementary Exercises

Exercise(s) 21.14.0.1. *TBD*

Exercise(s) 21.14.0.2. *TBD*

Exercise(s) 21.14.0.3. *Let \mathcal{P} be the set of prime numbers, $v_p(x)$ the exponent of p in the factorization of $x \in \mathbf{Q}^*$, and $v_\infty(x) \in \mathbf{Z}/2\mathbf{Z}$ defined by $\text{sign}(x) = (-1)^{v_\infty(x)}$. Show that the application $x \mapsto (v_i(x))$ defines a group isomorphism $\mathbf{Q}^* / (\mathbf{Q}^*)^2 \simeq (\mathbf{Z}/2\mathbf{Z})^{(\mathcal{P} \cup \{\infty\})}$.*

Chapter 22

The orthogonal group of a non-degenerate quadratic form

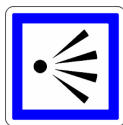
FORMIS AEQVATIONIBVSQVE INDETERMINATIS

SECVNDI GRADVS,

153. **I**n hac sectione imprimis de functionibus duarum indeterminatarum x, y , huius formae, $axx + 2bxy + cyy$, vbi a, b, c sunt integri dati tractabimus, quas *formas secundi gradus*, siue simpliciter *formas* dicemus. Huic disquisitioni superstruetur

Gauss' Disquisitiones¹

22.1 Perspective



In this chapter, (E, q) denotes a non-degenerate quadratic space of dimension $n > 0$ and S is the (invertible) matrix of q in a given basis \mathcal{B} .

¹On Forms and Indeterminate Equations of the Second Degree.//153. In this section, we will mainly discuss functions of two indeterminates of the form $ax^2 + 2bxy + cy^2$, where a, b , and c are given integers, functions that we will call second degree forms, or simply forms...

22.2 Definition

Note that an endomorphism u with matrix $M = \text{Mat}(u, \mathcal{B})$ preserves q if and only if ${}^t\text{MSM} = S$ so that $\det(M) \neq 0$. Thus, u is automatically an isometry (or an orthogonal endomorphism as preferred). The set of isometries of (E, q) forms a subgroup $O(q)$ of $\text{GL}(E)$. It is noted that the orthogonal groups of two equivalent quadratic forms are isomorphic (**exercise**).

Remark(s) 22.2.0.1. *If b is Hermitian or alternating, we can still talk about isometries. In the Hermitian case, we refer to the unitary group $U(b)$ and in the alternating case to the symplectic group $\text{Sp}(b)$*

Since $u \in O(q)$ if and only $u^* = u^{-1}$, that is if

$${}^t\text{Mat}(u, \mathcal{B}) \cdot \text{Mat}(q, \mathcal{B}) \cdot \text{Mat}(u, \mathcal{B}) = \text{Mat}(q, \mathcal{B}).$$

we have $(\det u)^2 = 1$. We then define the special orthogonal group $\text{SO}(q)$ as the normal subgroup of isometries with determinant 1, resulting in an exact sequence

$$\{1\} \rightarrow \text{SO}(q) \rightarrow O(q) \rightarrow \{\pm 1\} \rightarrow \{1\}$$

22.3 The case of dimension 2

A non-degenerate quadratic plane is either hyperbolic or anisotropic (21.4.0.1) depending on whether $-\text{disc}(q)$ is a square or not. Let's study $\text{SO}(q)$ in each of these cases.

Proposition 22.3.0.1. *The special orthogonal group of a hyperbolic plane is commutative and isomorphic to \mathbf{k}^* . Explicitly $q(x, y) = xy$, we have $\text{SO}(q) = \{\text{diag}(a, a^{-1}), a \in \mathbf{k}^*\} \simeq \mathbf{k}^*$.*

PROOF. *We choose coordinates such that $q(x, y) = xy$ so that $\text{SO}(q)$ identifies with matrices $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of determinant 1 such that ${}^t\text{MSM} = S$ with $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The corresponding equations are then*

$$2ac = 0$$

$$bc + ad = 1$$

$$2bd = 0$$

$$ad - bc = 1$$

From the second and fourth we deduce $ad = 1$ thus a and d are invertible. From the first and third we then deduce $b = c = 0$.

Let $-\alpha$ be a non-square in \mathbf{k}^* and q_α with discriminant α . Recall (21.5) that an anisotropic plane is isometric to the field $K = \mathbf{k}[\sqrt{-\alpha}]$ (isomorphic to \mathbf{k}^2 as a \mathbf{k} -vector field) equipped with the form $N(z) = z\sigma(z)$ where $\sigma(x + \sqrt{-\alpha}y) = x - \sqrt{-\alpha}y$.

Proposition 22.3.0.2. *The special orthogonal group of a hyperbolic plane is commutative and isomorphic to the subgroup $\{z \in K^* \mid N(z) = 1\}$ of K^* , which acts by multiplication on K equipped with the quadratic form N .*

PROOF. *The matrix for multiplication by $z = a + \alpha c$ in K in its natural basis $(1, \sqrt{-\alpha})$ is $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and the determinant of M is $N(z)$. In this basis, $q(x, y) = x^2 + \alpha y^2$ so that we have well defined an injective morphism of $\{z \in K^* \mid N(z) = 1\}$ into $SO(q_\alpha)$. For surjectivity, it's a calculation analogous to the previous one with $q(x, y) = x^2 + \alpha y^2$ so that $SO(q)$ identifies with matrices $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of*

determinant 1 such that ${}^tMSM = S$ with $S = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$.

As $\det(M) = 1$, we have

$$\begin{pmatrix} a & \alpha c \\ \alpha^{-1}b & d \end{pmatrix} = S^{-1}{}^tMS = M^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

and thus $a = d$, $c = -b/\alpha$ with $\det(M) = a^2 + \alpha c^2 = 1$ proving surjectivity.

The reader will recognize for $\mathbf{k} = \mathbf{R}$ and $\alpha = 1$ the isomorphism of the plane rotation group with the unit circle of the complex plane.

22.4 Orthogonal Symmetries

A symmetry is an endomorphism u such that $u \circ u = \text{Id}$. In particular, they are invertible. Moreover, the eigenvalues are ± 1 , and there exists a decomposition of E into the direct sum $E = E_+ \oplus E_-$, where E_+ is the eigenspace associated with the eigenvalue 1, and E_- with the eigenvalue -1 .

Proposition 22.4.0.1. *A symmetry is orthogonal if and only if the E_+ and E_- are orthogonal. In this case, these spaces are non-isotropic.*

Conversely, if F is a non-isotropic subspace, then there exists a unique orthogonal symmetry such that F is exactly the eigenspace associated with the eigenvalue 1.

PROOF. If u is orthogonal, then, for $x \in E_+$ and $y \in E_-$, we have

$$b(x, y) = b(u(x), u(y)) = -b(x, y)$$

therefore $b(x, y) = 0$ because $\text{char } \mathbf{k} \neq 2$.

Conversely, if these spaces are orthogonal, then, let $x, y \in E$. We write

$$\begin{cases} x = x_+ + x_-, & (x_+, x_-) \in E_+ \times E_- \\ y = y_+ + y_-, & (y_+, y_-) \in E_+ \times E_- \end{cases}$$

It follows

$$b(u(x), u(y)) = b(x_+, y_+) + b(x_-, y_-) = b(x, y).$$

Let F be non-isotropic. We denote $H = F^\perp$. Thus, $E = F \oplus H$, and we can define $u \in O(q)$ by $u|_F = \text{Id}$ and $u|_H = -\text{Id}$.

Definition 22.4.0.2. When $\dim E_- = 1$, u is called a reflection, and when $\dim E_- = 2$, it is referred to as a inversion (half-turn around axis E_+).

Example(s) 22.4.0.3. If y is anisotropic, the endomorphism s_y defined by $s_y(x) = x - 2\frac{b(x,y)}{b(y,y)}y$ is the reflection with $E_+ = \{y\}^\perp$ and $E_- = \mathbf{k}y$. In particular, if $q(x) = q(y)$ and $x - y$ is anisotropic, then $s_{x-y}(x) = y$. It should be noted that the previous formula proves that the restriction of a reflection s_x to a stable space containing x is still a reflection and that the opposite of a reflection in dimension 3 is a reversal.

22.5 Orthogonal Similitude

As in the Euclidean case, orthogonal similitudes, or similitudes for short, are endomorphisms u such that there exists a scalar $\lambda \in \mathbf{k}^*$ such that $b(u(x), u(y)) = \lambda \cdot b(x, y)$. They form a group $\text{GO}(q)$ and we have the exact sequence

$$1 \rightarrow O(q) \rightarrow \text{GO}(q) \rightarrow \mathbf{k}^*$$

where the last arrow is given by the scalar λ . For convenience of the reader, we adapt by copying the Euclidean characterization below.

Matrix-wise, we obtain the following identity:

$${}^t \text{Mat}(u, \mathcal{B}) \cdot \text{Mat}(q, \mathcal{B}) \cdot \text{Mat}(u, \mathcal{B}) = \lambda \cdot \text{Mat}(q, \mathcal{B}).$$

Thus, $\det^2 u = \lambda^n$.

When \mathbf{k} is algebraically closed, or when $(\mathbf{k}^*)^2 = \mathbf{k}^*$, we have the short exact sequence:

$$1 \rightarrow O(q) \rightarrow GO(q) \rightarrow \mathbf{k}^* \rightarrow 1.$$

Indeed, if $\mu \in \mathbf{k}^*$, consider $\lambda \in \mathbf{k}$ such that $\lambda^2 = \mu$, and $u = \lambda I$. It follows that $b(u(x), u(y)) = \lambda^2 \cdot b(x, y) = \mu \cdot b(x, y)$. Generally, it is not obvious to find a section that gives the square root of a scalar.

We have the following characterization of similarities.

Proposition 22.5.0.1. *Let E be a finite-dimensional \mathbf{k} -vector space equipped with a non-degenerate quadratic form q . Let $u \in GL(E)$. Then, u is a similitude if and only if u preserves orthogonality, that is*

$$\forall x, y \in E, x \perp y \iff u(x) \perp u(y).$$

PROOF. *It is straightforward to verify that a similitude preserves orthogonality. Conversely, consider an orthogonal basis $\mathcal{B} = (e_1, \dots, e_n)$ of E . Let $\varepsilon_i = u(e_i)$, $i = 1, \dots, n$, which also form an orthogonal basis by assumption.*

Since q is non-degenerate, $q(e_i)$, $q(\varepsilon_i)$ are non-zero, hence there exists $\lambda_i \in \mathbf{k}^$ such that $q(\varepsilon_i) = \lambda_i q(e_i)$. It suffices to show that λ_i is independent of i to conclude that u is a similitude. We take two indices $i \neq j$ and set $\lambda = -q(e_i)/q(e_j)$. It follows*

$$b(e_i + e_j, e_i + \lambda e_j) = q(e_i) + \lambda q(e_j) = 0$$

therefore these vectors are orthogonal. Consequently, $u(e_i + e_j) = \varepsilon_i + \varepsilon_j$ and $u(e_i + \lambda e_j) = \varepsilon_i + \lambda \varepsilon_j$ are also orthogonal and we deduce

$$\lambda = -\frac{q(\varepsilon_i)}{q(\varepsilon_j)} = -\frac{\lambda_i}{\lambda_j} \cdot \frac{q(e_i)}{q(e_j)} = \lambda \cdot \frac{\lambda_i}{\lambda_j}.$$

This clearly shows that λ_i is a constant function of i .

22.6 Generators of the orthogonal group

We will demonstrate that reflections (resp. reversals) generate $O(q)$ (resp $SO(q)$.) Let's start with a simple case.

Lemma 22.6.0.1. *If q is anisotropic, any isometry u is the product of at most $\text{rk}(u - \text{Id})$ reflections.*

PROOF. *We use induction on*

$$d = \text{rk}(u - \text{Id}) = n - \dim \text{Ker}(u - \text{Id}) \leq n$$

If $d = 0$, the identity is indeed a product of 0 symmetries. Assume $0 < d \leq n$ and the theorem proven for any isometry v such that $\text{rk}(v - \text{Id}) \leq d - 1$.

As $d > 0$, we can choose x not belonging to $\text{Ker}(u - \text{Id})$, that is, not fixed by u . Let $y = u(x) \neq x$. We have $q(x) = q(y)$ and $y - x$ is anisotropic since it is non-null. We have $s_{x-y}(x) = y$ according to 22.4.0.3 and thus $v(x) = x$ with $v = s_{x-y} \circ u$. But if $z \in \text{Ker}(u - \text{Id})$, then

$$b(z, x - y) = b(z, x) - b(z, u(x)) = b(z, x) - b(u(z), u(x)) = b(z, x) - b(z, x) = 0$$

thus $\text{Ker}(u - \text{Id}) \subset \{x - y\}^\perp = \text{Ker}(s_{x-y} - \text{Id})$. Hence, $\mathbf{k}x \oplus \text{Ker}(u - \text{Id}) \subset \text{Ker}(v - \text{Id})$ and thus the codimension of $\text{rk}(v - \text{Id}) \leq d - 1$. We conclude by applying the induction hypothesis to v .

Proposition 22.6.0.2. *If q is non-degenerate, any isometry is a product of at most $2n$ hyperplane symmetries.*

PROOF. We proceed by induction on n and let $u \in \text{O}(q)$. Suppose x is non-isotropic such that $u = u(x)$ is also. One of the vectors $x - y$ or $x + y$ is thus non-isotropic according to the polarization formula.

If $x - y$ is non-isotropic, $v = s_{x-y} \circ u$ fixes x . As x is non-isotropic, the hyperplane $\mathbb{H} = x^\perp$ is a complement of $\mathbf{k}x$ which is stable by v . We then apply the induction hypothesis to the restriction of v to \mathbb{H} noting that (\mathbb{H}, q) is non-degenerate.

If $x + y$ is non-isotropic, we define $v = s_y \circ s_{x+y} \circ u$ and conclude as above.

Remark(s) 22.6.0.3. *It can be proven that n symmetries suffice (Cartan-Dieudonné).*

Proposition 22.6.0.4. *If $n \geq 3$, every element of $\text{SO}(q)$ is the product of at most n reversals.*

PROOF. Every element of $\text{SO}(q)$ is the product of an even number of reflections such that it is a matter of demonstrating that the product $u = s_x \circ s_y$ of two reflections is a composite of reversals.

We may assume that x, y are non-collinear so that they generate a plane \mathbb{P} . As $q|_{\mathbb{P}}$ is non-null (x and y anisotropic by definition), its kernel $\mathbb{P} \cap \mathbb{P}^\perp$ is at most of dimension 1. Note that the restriction of v to $\mathbb{P}^\perp = x^\perp \cap y^\perp$ is the identity and that \mathbb{P} is stable by v .

If $\mathbb{P} \cap \mathbb{P}^\perp = \{0\}$, then $\mathbb{E} = \mathbb{P} \oplus \mathbb{P}^\perp$ and \mathbb{P}^\perp is non-degenerate as \mathbb{E} so we can take $z \in \mathbb{P}^\perp$ non-isotropic. Then, $\mathbb{W} = \text{Span}(x, y, z)$ is non-degenerate such that we have a decomposition $\mathbb{E} = \mathbb{W} \oplus \mathbb{W}^\perp$ which is stable by v and we conclude by induction.

If $\mathbb{P} \cap \mathbb{P}^\perp$ is a line \mathbb{D} , let $z \notin \mathbb{D}^\perp$ and $\mathbb{W} = \text{Span}(x, y, z)$. As $\mathbb{D} \subset \mathbb{P}^\perp$, we have $\mathbb{P} \subset \mathbb{D}^\perp$ such that $z \notin \mathbb{P}$ and \mathbb{W} is dimension 3. Let's show that \mathbb{W} is non-degenerate. Suppose thus $w \in \mathbb{W} \cap \mathbb{W}^\perp$ non-null. Since

$w \in z^\perp$, we have $w \in D \subset P$. But $W^\perp \subset P^\perp$ thus $w \in P \cap P^\perp = D$ and $D = \mathbf{k}w$. But $b(w, z) = 0$, thus $z \in D^\perp$, a contradiction. Then, $E = W \oplus W^\perp$. As u acts on the identity on $P^\perp \subset W^\perp$, it also acts on the identity on W^\perp and thus leaves stable W . But the restrictions of $-s_x$ and $-s_y$ to W are reversals (22.4.0.3) as well as their extensions r_x, r_y by the identity on W^\perp . And we have $y = r_x \circ r_y$.

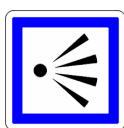
Exercise(s) 22.6.0.5. Show that the conjugate of a reflection by an isometry is an isometry, specifying its characteristic elements. Deduce from this the centers of the orthogonal and special orthogonal groups.

Chapter 23

Automorphisms of classical groups



23.0.1 Perspective



We first explain how the consideration of involutions in the orthogonal groups allows to intrinsically recover the projective space of \mathbf{R}^n in $O_n(\mathbf{R})/\{\pm \text{Id}\}$ with its combinatorial collinearity structure and how this allows to compute its automorphism group. We adapt the method to the general linear group (see (23.1.2.3) and [Dieu-1951] for a vast but difficult to read generalization).

23.1 Automorphisms of $O_n(\mathbf{R})$

Our goal in this section is to prove that the inner automorphism

$$\begin{cases} O_n(\mathbf{R}) & \rightarrow & O_n(\mathbf{R}) \\ g & \mapsto & \text{Ad}(g_0)(g) = g_0 g g_0^{-1} \end{cases}$$

are the only automorphisms of the orthogonal group in dimension ≥ 3 ! The key point is the study of involutions.

23.1.1 Involutions

Lemma 23.1.1.1. *Let S be a commuting family of involutions of $O_d(\mathbf{R})$, $d \geq 1$. For $s \in S$, let us define $p(s) = \min\{\dim \text{Fix}(s), \dim \text{Fix}(s)^\perp\} \in [0, \dots, n/2]$.*

1. *s is conjugate under $SO_d(\mathbf{R})$ to $\pm \text{diag}(\text{Id}_{p(s)}, -\text{Id}_{n-p(s)})$.*
2. *$\text{Fix}(s') = \text{Fix}(s)$ if and only if $s' = \pm s$.*
3. *s, s' are conjugate if and only if $p(s) = p(s')$.*
4. *$p(s) = 0$ if and only if $s \in Z(G) = \{\pm \text{Id}\}$.*
5. *$\pm s$ is an orthogonal reflection if and only if $p(s) = 1$: we will call them extremal involutions.*
6. *The restrictions to $\text{Fix}(s)$ and $\text{Fix}(s)^\perp$ defines an isomorphism*

$$C(s) \xrightarrow{\sim} O(\text{Fix}(s)) \times O(\text{Fix}(s)^\perp).$$

In particular, two different reflections commutes if and only if their skew lines are orthogonal.

7. *There exists $g \in SO_d(\mathbf{R})$ such that gSg^{-1} is a family of diagonal matrices with ± 1 entries¹.*

2

Proof.

1. Recall that $V = \text{Fix}(s) \oplus \text{Fix}(s)^\perp$ and $\text{Fix}(s)^\perp = \text{Ker}(s + \text{Id})$ for an orthogonal symmetry. Then, changing the sign of one of the eigenvectors of a corresponding orthonormal basis of eigenvectors, we get the first three points.
5. $p(s) = 0$ means either $\dim \text{Fix}(s) = n$ and $s = \text{Id}$ or $\dim \text{Fix}(s)^\perp = n$ and $s = -\text{Id}$ (which is the only non trivial central element of $O_n(\mathbf{R})$).

¹Compare with 5.7.0.4.

6. An matrix is in the centralizer of a diagonalizable matrix if and only if it leaves globally invariant its eigenspaces. It is moreover orthogonal if and only the corresponding restrictions are orthogonal, proving the item.
7. Induction on d , the case $d = 1$ being trivial. One can assume than one involution s of the family is $\neq \text{Id}$. Both $F_+ = \text{Fix}(s)$ and its orthogonal F_- are dimension $< d$ and are stable by all members s' of the family because they are its ± 1 eigenspaces of s which commutes with all s' (5.7.0.3). Let \mathcal{B}_\pm be orthonormal basis of F_\pm . Changing one vector to this opposite if necessary, one can assume that $\mathcal{B}_+ \sqcup \mathcal{B}_-$ is a direct orthogonal basis of \mathbf{R}^d which identifies V_\pm to standard Euclidean of dimension $< d$. With this identification, by induction there exists g_\pm in the special orthogonal group such that $g_\pm S_\pm g_\pm^{-1}$ is a family of diagonal matrices with ± 1 entries and we set $g = \text{diag}(g_+, g_-)$.

□

Corollary 23.1.1.2. *Let $d, d' \geq 1$. Then $O_d(\mathbf{R}) \xrightarrow{\sim} O_{d'}(\mathbf{R})$ or $SO_d(\mathbf{R}) \xrightarrow{\sim} SO_{d'}(\mathbf{R})$ if and only if $d = d'$.*

Proof. By the above lemma, the maximal numbers of commuting involution of $O_d(\mathbf{R})$ is 2^d (2^{d-1} for $SO_d(\mathbf{R})$), number which is invariant through group isomorphism. □

Exercise(s) 23.1.1.3. *Let $d, d' \geq 1$. Prove that $SO_d(\mathbf{R})$ and $O_{d'}(\mathbf{R})$ are never isomorphic.*

Let $G = O_n(\mathbf{R}), n \geq 3$ acting naturally on the left on $V = \mathbf{R}^n$. We denote by $C(S)$ the centralizer of a subset $S \subset G$ and by $D(H)$ the derived subgroup $H \subset G$.

Proposition 23.1.1.4.

1. *The maximal cardinality of a family S of pairwise permuting involutions s' conjugate to a given involution s of $O_n(\mathbf{R})$ is $\binom{n}{p(s)}$.*
2. *The involution s of $O_n(\mathbf{R})$ is extremal if and only if the maximal number of commuting family of involutions in the conjugacy class of s is n . In this case, there a unique line $D = D(s)$ of V such that $s = \pm s_D$.*
3. *Let D_i be lines of V . One has*

$$D(C(s_{D_1}, \dots, s_{D_i})) \xrightarrow{\sim} SO(D_1 + \dots + D_i)^\perp$$

where the second member acts trivially on $D_1 + \dots + D_i$.

Proof. 1. By lemma 23.1.1.1, one can assume that all matrices are diagonal with ± 1 entries with $\pm \text{diag}(\text{Id}_{p(s)}, -\text{Id}_{n-p(s)}) \in S$. A matrix in $\pm S$ is diagonal with d coefficients equal to 1, the other one being equal to -1 because it is conjugate to s . There is at least $\binom{n}{p(s)}$ such matrices and the maximal set is obtained by choosing an arbitrary subset of cardinality $p(s)$ in the indices $\{1, \dots, n\}$ where the coefficients of matrices of $\pm S$ will be equal to 1.

2. The formula $\binom{n}{1} = n$ gives the first point. Because, $\pm s$ is conjugate to $\text{diag}(1, -\text{Id}_{n-1})$, the existence of D follows. If we had two such different lines D, D' , we would have $s_D = -s_{D'}$ which is only possible for $n = 2$.

3. As in the above lemma, because $V = D \oplus D^\perp$, one has $g \in C(s_D)$ if and only if $g|_D = \pm 1$ and $g|_{D^\perp} \in O(D^\perp)$. More generally, $g \in C(s_{D_1}, \dots, s_{D_i})$ if and only if $g(D_i) \subset D_i$ meaning that its restriction to $(D_1 + \dots + D_i)$ acts by ± 1 on each generating lines and that its orthogonal is stable (with an orthogonal restriction of g). If G is the (finite abelian) group of all orthogonal transformations acting by ± 1 on each D_i 's, this means that, up to the obvious diagonal embedding, one has

$$\{\text{Id}\} \times O(D_1 + \dots + D_i)^\perp \subset C(s_{D_1}, \dots, s_{D_i}) \subset G \times O(D_1 + \dots + D_i)^\perp$$

which gives the item because $Z(G)$ is trivial and $D(O_d(\mathbf{R})) = D(\text{SO}_d(\mathbf{R}))$ (13.7).

□

23.1.2 The main theorem

Theorem 23.1.2.1. *Any automorphism Φ of $O_n(\mathbf{R})$, $n \geq 3$ is inner.*

Proof. Φ maps the center $\{\pm \text{Id}\}$ onto itself and therefore non central involutions to non central involutions. Because extremal involutions (23.1.1.1) are purely characterized in group theory terms by 23.1.1.4 and 23.1.1.1, the image of the extremal involution S_D is an extremal involution

$$\Phi(s_D) = \varepsilon(D)s_\varphi(D)$$

for some line $\Phi(D)$ and sign $\varepsilon(D) \in \{\pm 1\}$. Let Φ' be the inverse of Φ and φ' the corresponding map at the line levels. One has $s_D = \pm s_{\varphi' \circ \varphi(D)}$. Because the defining of an extremal involution is line is uniquely defined, one has $\varphi' \circ \varphi = \text{Id}$ and analogously $\varphi \circ \varphi' = \text{Id}$ showing that Φ is a bijection.

If D_1, D_2, D_3 are distinct collinear points of $\mathbf{P}_{\mathbf{R}}^n$, then, as subvector spaces of V , the dimension of $D_1 + D_2 + D_3$ is 2 and thanks to (23.1.1.4)

$$Z(C(s_{D_1}, s_{D_2}, s_{D_3})) \xrightarrow{\sim} \text{SO}_{n-2}(\mathbf{R}).$$

Because Φ is a morphism, it induces an isomorphism

$$Z(C(s_{D_1}, s_{D_2}, s_{D_3})) \xrightarrow{\sim} Z(C(s_{\varphi(D_1)}, s_{\varphi(D_2)}, s_{\varphi(D_3)})) \xrightarrow{\sim} SO_{n-3}(\mathbf{R})$$

and therefore $SO_{n-2}(\mathbf{R}) \xrightarrow{\sim} SO_{n-3}(\mathbf{R})$. This is impossible if $n \geq 4$ by (23.1.1.2). If $n = 3$, one observes that $g \in C(s_{D_1}, s_{D_2}, s_{D_3})$ if and only if g leaves the plane $D_1 + D_2 + D_3$ invariant, giving an infinite centralizer and that $\gamma \in C(s_{\varphi(D_1)}, s_{\varphi(D_2)}, s_{\varphi(D_3)})$ if and only if γ leaves the independent lines $\varphi(D_i)$ invariant, showing that this group is finite (at most of order 2^3). In all situation, the bijection φ preserves collinearity.

By (18.5.1.2) and the fundamental theorem of projective geometry (18.5.1.3), there exists an homography defined by some $A \in GL_n(\mathbf{R})$ such that $\forall D \in \mathbf{P}_{\mathbf{R}}^n$, $\varphi(D) = A(D)$. If D, D' are orthogonal, s_D and $s_{D'}$ commute, so does their image $\pm s_{\varphi(D)}$ and $\pm s_{\varphi(D')}$ implying $\varphi(D)$ and $\varphi(D')$ orthogonal by 23.1.1.1. We deduce that A is an Euclidean similitude (13.8). We can therefore assume $A \in O_n(\mathbf{R})$. The formula $\text{Ad}(A)(s_D) = s_{A(D)}$ shows that the automorphism $\Psi = \text{Ad}(A^{-1}) \circ \Phi$ satisfies $\Psi(s_D) = \pm s_D$ for any line D . Because reflections generate the orthogonal group (13.7.0.2), we have

$$\Psi(g) = \varepsilon(g)g, g \in O_n(\mathbf{R})$$

where ε has values in $\{\pm 1\}$ and is a morphism of groups like Ψ . But ε is trivial on the derived group which is $SO_n(\mathbf{R})$ (13.7.0.5) and is accordingly induced by a endomorphism of the group with 2 elements proving $\varepsilon(g) = 1$ for all g . This proves $\Psi = \text{Id}$ and $\Phi = \text{Ad}(A)$ as wanted. \square

Exercise(s) 23.1.2.2. Prove that the surjective morphism $\text{Ad} : O_n(\mathbf{R}) \rightarrow \text{Aut}(O_n(\mathbf{R}))$ has kernel $\{\pm \text{Id}\}$.

Remark(s) 23.1.2.3. The key geometric observation is that there if V is Euclidean, the projective space

1. $\mathbf{P}V$ can be intrinsically defined in purely group theoretic terms in $O(V)/Z(O(V)) = O(V)/\{\pm \text{Id}\}$ as $\{\text{extremal involutions}\}/Z(O(V))$
2. the incidence relation defined by collinearity is also canonically defined in purely group theoretic terms.

Therefore, an automorphism preserves this combinatorial datum and therefore is induced by an homography: this is exactly the strategy we have used. In the next section 23.2, we will keep the same general strategy: defining a combinatorial object in purely group theoretic terms related with projective geometry which will be therefore invariant by any automorphism and then apply the fundamental theorem of projective geometry.

The reader will easily check that this argument works without modification for $SO(2n+1)(\mathbf{R})$ but fails for $SO(2n)(\mathbf{R})$. The reason is that reflections never belong to SO and that non central extremal involutions are up to sign Euclidean inversions. Therefore we have to study the so-called Grassmannian of planes

rather than the projective space of planes. This can be done (in an analogous but more complicated way see (23.2.6.2)) but only for $n > 4$. Indeed, we have already seen (see chapter 14) that in this case (14.3.4.1)

$$\mathrm{SO}_4(\mathbf{R}) \xrightarrow{\sim} \mathrm{SO}_3(\mathbf{R}) \times \mathrm{SO}_3(\mathbf{R}) / \{\pm \mathrm{Id}\}$$

and that the morphism flipping both factors is not inner (14.5.0.4).!

23.2 Automorphisms of $\mathrm{GL}(V)$

Let \mathbf{k} be a (commutative) field of characteristic $\neq 2$ and V a \mathbf{k} -vector space of dimension $n > 2$. We identify V with its bidual: a hyperplane in the dual defines a line $D \in \mathbf{P}V$, specifically H^\perp . Symmetrically, a line D in V defines a hyperplane $H \in \mathbf{P}V^*$, specifically H^\perp . The hypothesis ensures that a hyperplane H of V is never a line D of V , and vice versa.

23.2.1 Involutions

We adapt the results 23.1.1. As in the orthogonal case, we define for any symmetry s of a vector space, $E_\pm(s) = \mathrm{Ker}(s \pm \mathrm{Id})$ and $p(s) = \min_\pm \{\dim E_\pm\} \in [0, \dots, n/2]$. Involutions with $p(s) = 1$ are called *extremal* involutions.

Lemma 23.2.1.1. *Let s, s' involutions of $\mathrm{GL}_d(\mathbf{k})$*

1. s is conjugate under $\mathrm{SL}_d(\mathbf{k})$ to $\pm \mathrm{diag}(\mathrm{Id}_{p(s)}, -\mathrm{Id}_{n-p(s)})$.
2. s, s' are conjugate if and only if $p(s) = p(s')$.
3. $p(s) = 0$ if and only if s is central, i.e. $s = \pm \mathrm{Id}$.
4. The restrictions to E_+ and E_- defines an isomorphism $C(s) \xrightarrow{\sim} \mathrm{GL}(E_+) \times \mathrm{GL}(E_-)$.
5. For any commuting family S of involutions, there exists $g \in \mathrm{SL}(V)$ such that gSg^{-1} is a family of diagonal matrices with ± 1 entries.
6. Let $d, d' \geq 1$. Prove that $\mathrm{GL}_d(\mathbf{k})$ and $\mathrm{GL}_{d'}(\mathbf{k})$ are never isomorphic.
7. The maximal number of commuting symmetries³ with s is $\binom{n}{p(s)}$. In particular, s is extremal if and only if s is non central and this maximal number is n .

³Another group theoretic characterization of extremal symmetries follows from the fact that $p(s) > 1$ if and only if $D(C(s)) \xrightarrow{\sim} \mathrm{SL}_{p(s)}(\mathbf{k}) \times \mathrm{SL}_{n-p(s)}(\mathbf{k})$ has a non central proper normal subgroup.

Proof. Straightforward adaptation of the proof of 23.1.1.1 and 23.1.1.4. □

We will denote by $\mathcal{E} \subset GL(V)$ the set of extremal involutions and for any $s \in \mathcal{E}$ by $D(s) \in \mathbf{P}V$ its dimension 1 eigenspace and by $H(s) \in \mathbf{P}V^*$ its dimension $n - 1$ eigenspace (they are well defined because $1 \neq n - 1$). Finally, for any subset $S \subset \mathcal{E}$, we define this ad-hoc version of the commutant of S by

$$S^\diamond = \{s' \in \mathcal{E} \mid \forall s \in S, s \circ s' = s' \circ s\}.$$

Let us observe that $S \mapsto S^\diamond$ is decreasing and $S \subset S^{\diamond\diamond}$.

23.2.2 Pairs of extremal involutions

Definition 23.2.2.1. *A pair of involutions $\{\sigma_1, \sigma_2\}$ is minimal if the following conditions hold*

1. $\sigma_1, \sigma_2 \in \mathcal{E}$.
2. $\sigma_1 \neq \pm\sigma_2$ (i.e. $\sigma_1\sigma_2$ non central)
3. σ_1 and σ_2 have a common eigenspace.

Lemma 23.2.2.2. *Let $\sigma_1, \sigma_2 \in \mathcal{E}$. Then,*

1. σ_1 and σ_2 commute if and only
 - either $D_1 = D_2$ and $H_1 = H_2$ (or equivalently $\sigma_1\sigma_2$ is central)
 - or $D_1 \subset H_2$ and $D_2 \subset H_1$.

Assume σ_1 and σ_2 do not commute.

2. $\{\sigma_1, \sigma_2\}^\diamond = \{s \in \mathcal{E} \mid D(s) \subset H_1 \cap H_2 \text{ and } D_1 + D_2 \subset H(s)\}$.
3. If $H_1 \cap H_2 \subset D_1 + D_2$, then $\{\sigma_1, \sigma_2\}^\diamond = \emptyset$, $\{\sigma_1, \sigma_2\}^{\diamond\diamond} = \mathcal{E}$ and $n = 3$.
4. If $H_1 \cap H_2 \not\subset D_1 + D_2$ then $\{\sigma_1, \sigma_2\}^{\diamond\diamond} = \{s' \in \mathcal{E} \mid H_1 \cap H_2 \subset H(s') \text{ and } D(s') \subset D_1 + D_2\}$.
5. $\{\sigma_1, \sigma_2\}$ is minimal if and only if for any $\sigma'_1, \sigma'_2 \in \{\sigma_1, \sigma_2\}^{\diamond\diamond}$ such that $\sigma'_1 \neq \pm\sigma'_2$, one has $\{\sigma_1, \sigma_2\}^{\diamond\diamond} = \{\sigma'_1, \sigma'_2\}^{\diamond\diamond}$.

Proof.

Proof of (1).

Assume first σ_1 and σ_2 commute. Changing sign if necessary, in a suitable basis σ_1 has matrix $-\text{Id} + 2E_{1,1}$ and $\sigma_2 = \text{Id} + 2E_{i,i}$ for some $i = 1, \dots, n$. Then $D_1 = \langle e_1 \rangle, H_1 = \langle e_j, j \neq 1 \rangle$ and $D_2 = \langle e_j \rangle, H_2 = \langle e_j, j \neq i \rangle$ proving the direct implication of the first item.

Conversely, assume $D_1 \subset H_2$ and $D_2 \subset H_1$. Changing signs if necessary, we can assume that H_i are the space of fixed points of σ_i . Then $H_2 \neq H_1$ because $D_1 \cap H_1 = \{0\}$. The intersection $H_1 \cap H_2$ is therefore an hyperplane of H_1 and $H_1 = D_2 \oplus H_1 \cap H_2$ and by symmetry $H_2 = D_1 \oplus H_1 \cap H_2$. Therefore $V = D_1 \oplus D_2 \oplus H_1 \cap H_2$. It follows that $\sigma_1 \circ \sigma_2$ is the identity on $H_1 \cap H_2$ and $-\text{Id}$ on D_1, D_2 and so is $\sigma_2 \circ \sigma_1$ by symmetry proving that σ_1 and σ_2 commute.

Proof of (2).

Let $s \in \{\sigma_1, \sigma_2\}^\diamond$. Because σ_1 and σ_2 does not commute, $s = \text{not} = \pm\sigma_i, i = 1, 2$. By (1), we get therefore $D(s) \subset H_i$ and $D_i \subset H(s)$ hence the inclusion

$$\{\sigma_1, \sigma_2\}^\diamond \subset \{s \in \mathcal{E} \mid D(s) \subset H_1 \cap H_2 \text{ and } D_1 + D_2 \subset H(s)\}.$$

The reverse inclusion is obvious from (1).

Proof of (3).

If $s \in \{\sigma_1, \sigma_2\}^\diamond$, we have in our case $D(s) \subset H(s)$ which is impossible. Therefore, $\{\sigma_1, \sigma_2\}^\diamond = \emptyset, \{\sigma_1, \sigma_2\}^{\diamond\diamond} = \mathcal{E}$. Looking at dimensions in the inclusion $H_1 \cap H_2 \subset D_1 + D_2$, we get $n - 2 \leq 2$. If $n = 4$, we get $H_1 \cap H_2 = D_1 + D_2$ hence $D_1 \subset H_1$ which is impossible.

Proof of (4). We have $H_1 \cap H_2 \not\subset D_1 + D_2$.

Let

$$s' \in \{s' \in \mathcal{E} \mid H_1 \cap H_2 \subset H(s') \text{ and } D(s') \subset D_1 + D_2\}$$

and let

$$s \in \{\sigma_1, \sigma_2\}^\diamond \stackrel{(2)}{=} \{s \in \mathcal{E} \mid D(s) \subset H_1 \cap H_2 \text{ and } D_1 + D_2 \subset H(s)\}$$

One has therefore $D(s) \subset H_1 \cap H_2 \subset H(s')$ and $D(s') \subset D_1 + D_2 \subset H(s)$ and therefore s, s' commute thanks to (1) proving $s' \in \{\sigma_1, \sigma_2\}^{\diamond\diamond}$.

Conversely, assume that $s' \notin \{s' \in \mathcal{E} \mid H_1 \cap H_2 \subset H(s') \text{ and } D(s') \subset D_1 + D_2\}$ with for instance⁴ $H_1 \cap H_2 \not\subset H(s')$. But $H_1 + H_2 \not\subset D_1 \cap D_2$ by assumption and therefore, both the intersections of $H(s')$ and $D_1 + D_2$ with $H_1 \cap H_2$ are proper subspaces of $H_1 \cap H_2$ whose union cannot be the whole $H_1 \cap H_2$ (??). Let us chose a line $D = \langle d \rangle$ with d in the complement :

$$D \subset H_1 \cap H_2 \text{ and } D \not\subset H(s'), D \not\subset D_1 + D_2.$$

Let S be any supplement of $D \oplus (D_1 + D_2)$ in V and the hyperplane $H = (D_1 + D_2) \oplus S$ and let $s \in \mathcal{E}$ the involution $(d, h) \in D \oplus H \mapsto (-d, h)$. By construction,

$$D(s) = D \subset H_1 \cap H_2 \text{ and } D_1 + D_2 \subset H = H(s)$$

proving $s \in \mathcal{E}$ using (a) and $D(s) \not\subset H(s')$ proving that s and s' do not commute thanks to the first point.

⁴The case $D(s') \not\subset D_1 + D_2$ being reduced to this one by duality.

Proof of (5).

For the direct implication, assume that $\{\sigma_1, \sigma_2\}$ is minimal with for instance $D_1 = D_2 = D$ and therefore $H_1 \neq H_2$ because $\sigma_1 \neq \pm\sigma_2$. If we had $H_1 \cap H_2 \subset D_1 + D_2 = D$, we would have $H_1 \cap H_2 = D$ and $D_1 = H_1 \cap H_2 \subset H_1$, a contradiction. Therefore, we have $H_1 \cap H_2 \not\subset D_1 + D_2 = D$. Let $\sigma'_1, \sigma'_2 \in \{\sigma_1, \sigma_2\}^{\diamond\diamond}$ such that $\sigma'_1\sigma'_2 \neq \pm Id$. We will denote the eigenspaces of σ'_i by D'_i, H'_i with $\dim(D'_i) = 1$. One has

$$\sigma'_i \in \{\sigma_1, \sigma_2\}^{\diamond\diamond} \stackrel{(4)}{=} \{s \in \mathcal{E} | H_1 \cap H_2 \subset H(s) \text{ and } D(s) = D\}.$$

In particular, we $D'_i = D$ and $H_1 \cap H_2 \subset H'_1 \cap H'_2$. But $H'_1 \neq H'_2$ because $\sigma'_1 \neq \pm\sigma'_2$ and therefore $H_1 \cap H_2 = H'_1 \cap H'_2 \not\subset D'_1 + D'_2 = D$. We conclude the wanted equality $\{\sigma_1, \sigma_2\}^{\diamond\diamond} = \{\sigma'_1, \sigma'_2\}^{\diamond\diamond}$ by (4) applied to σ'_1, σ'_2 .

For the converse implication, assume now $\{\sigma_1, \sigma_2\}$ not minimal, that is $H_1 \neq H_2$ and $D_1 \neq D_2$. If $H_1 + H_2 \subset D_1 + D_2$, we have $\{\sigma_1, \sigma_2\}^{\diamond\diamond} \stackrel{(3)}{=} \mathcal{E}$ and $n = 3$. We just have to give one example of non commuting involutions of \mathbf{k}^3 with produce two $\{\sigma_1, \sigma_2\}^{\diamond\diamond} \subsetneq \mathcal{E}$: take for instance the permutation matrices of two transpositions of S_3 . Assume now $H_1 + H_2 \not\subset D_1 + D_2$. Let $v \in V$ such $v \pmod{H_1 \cap H_2}$ does not belong to the two lines images of D_1, H_1 in the plane $V/H_1 \cap H_2$ (7.1.0.1) and let $H = \langle v \rangle \oplus H_1 \cap H_2$. By construction, $D_1 \not\subset H$ and $H \neq H_1$. Let s be the reflection associated to D_1, H . The involutions σ_1 and s do not commute (because $H_1 \neq H$) and $H_1 \cap H = H_1 \cap H_2$ for dimension reasons. If $H_1 \cap H \subset D_1 + D$, we have $\{\sigma_1, s\}^{\diamond\diamond} \stackrel{(3)}{=} \mathcal{E}$ and $\mathcal{E} \neq \{\sigma_1, \sigma_2\}^{\diamond\diamond}$ because any symmetry s' with $D(s') \not\subset D_1 + D_2$ is not in $\{\sigma_1, \sigma_2\}^{\diamond\diamond}$ by (4). Assume finally that $H_1 \cap H \not\subset D_1 + D$. We have

$$\sigma_2 \notin \{s' \in \mathcal{E} | H_1 \cap H \subset H(s') \text{ and } D(s') = D_1\} \stackrel{(4)}{=} \{\sigma_1, s\}^{\diamond\diamond}$$

and certainly $\sigma_2 \in \{\sigma_1, \sigma_2\}^{\diamond\diamond}$. □

Lemma 23.2.2.3. *Up to sign, any transvection is a product of a minimal pair of involutions and conversely any such product is a transvection.*

Proof. The formula

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

shows that a transvection is the product of two extreme involutions with a common eigenspace (here $\mathbf{k}e_1$).

Conversely, by transposing if necessary, we can consider two extreme involutions with a common fixed line. In a suitable basis, the matrices of these involutions are necessarily

$$\pm \text{diag}(1, -1, \dots, -1) \text{ and } \pm \begin{pmatrix} 1 & L \\ 0 & -Id \end{pmatrix},$$

where $L = (L_i, i \geq 2) \neq 0$. Their product, up to sign, is $\begin{pmatrix} 1 & L \\ 0 & Id \end{pmatrix}$, which maps $x = (x_1, \dots, x_n) \in \mathbf{k}^n$ to $x + \sum_{i=2}^n L_i x_i e_1$. This is a transvection of \mathbf{k}^n with hyperplane H defined by $\sum_{i=2}^n L_i x_i = 0$ and line $\mathbf{k}e_1 \subset H$. □

Lemma 23.2.2.4. *For two transvections τ_1, τ_2 of respective types (D_1, H_1) and (D_2, H_2) , their product is again a transvection or the identity if and only if $D_1 = D_2$ or $H_1 = H_2$ and they commute if and only if $D_1 = D_2$ and $H_1 = H_2$*

Proof. Let $\tau_i(x) = x + \varphi_i(x)v_i$, where φ_i, v_i are nonzero and $\varphi_i(v_i) = 0$. Then:

$$\tau_1\tau_2(x) = x + \varphi_2(x)v_2 + \varphi_1(x + \varphi_2(x)v_2)v_1 = x + \varphi_2(x)v_2 + (\varphi_1(x) + \varphi_2(x)\varphi_1(v_2))v_1.$$

If $H_1 = H_2$, then $\varphi_2 = \lambda\varphi_1$ with $\lambda \neq 0$, so $\varphi_1(v_2) = 0$. Thus:

$$\tau_1\tau_2(x) = x + \varphi_1(x)(v_1 + \lambda v_2),$$

which is a transvection or the identity, depending on whether $v_1 + \lambda v_2 \neq 0$.

If $D_1 = D_2$, then $v_2 = \lambda v_1$ with $\lambda \neq 0$, so:

$$\tau_1\tau_2(x) = x + (\lambda\varphi_2 + \varphi_1)(x)v_1,$$

which is a transvection or the identity, depending on whether $(\lambda\varphi_2 + \varphi_1) \neq 0$.

Conversely, suppose $\tau_1\tau_2$ is a transvection, so $D = \text{Im}(\tau_1\tau_2 - \text{Id})$.

If $D_1 \neq D_2$, the pair (v_1, v_2) is linearly independent. For every x , the vector:

$$\varphi_2(x)v_2 + (\varphi_1(x) + \varphi_2(x)\varphi_1(v_2))v_1$$

is collinear to a direction vector of D , necessarily of the form $av_1 + bv_2$. Therefore:

$$\det \begin{pmatrix} \varphi_2(x) & \varphi_1(x) + \varphi_2(x)\varphi_1(v_2) \\ b & a \end{pmatrix} = (a - b\varphi_1(v_2))\varphi_2(x) - b\varphi_1(x) = 0,$$

which establishes a non-trivial linear relation between φ_1 and φ_2 , implying $H_1 = H_2$.

Thus, $D_1 = D_2$ or $H_1 = H_2$ is a necessary and sufficient condition. The commutation statement follows from the previous computation (sufficient condition) and the fact that commuting endomorphisms preserve eigenspaces (necessary condition). \square

We call a subgroup of $\text{SL}(V)$ a t -group if, up to sign, all its non-identity elements are transvections.

Corollary 23.2.2.5. *Let Σ be a t -group in $\text{SL}(V)$. Then:*

- *Either there exists a line D such that for every transvection $\pm\tau \in \Sigma$, $D(\tau) = D$,*
- *Or there exists a hyperplane H such that for every transvection $\pm\tau \in \Sigma$, $H(\tau) = H$.*

In particular, the maximal t -subgroups of $\Sigma \subset \text{SL}(V)$ are the non commutative subgroups:

1. $\Sigma_D \xrightarrow{\sim} \text{Hom}(V/D, D)$, whose non-central elements are, up to sign, the transvections τ with $D(\tau) = D$ for a given line D depending only on Σ .
2. $\Sigma_H \xrightarrow{\sim} \text{Hom}(V/H, H)$, whose non-central elements, up to sign are the transvections τ with $H(\tau) = H$ for a given hyperplane H depending only on Σ .

Remark(s) 23.2.2.6. Non central elements of $\Sigma_D \cap \Sigma_H$ are up to sign transvections of type (D, H) if $D \subset H$ and $\pm \text{Id}$ else, proving that $\Sigma_D \cap \Sigma_H$ is commutative. In particular, we never have $\Sigma_D = \Sigma_H$.

Exercise(s) 23.2.2.7. Prove that $\Sigma_D \cap \Sigma_H$ is equal to $\{\pm \text{Id}\}$ if $D \not\subset H$ and isomorphic to the additive group $\mathbf{Z}/2\mathbf{Z} \times \mathbf{k}$, if $D \subset H$.

23.2.3 Proof of the theorem

Proposition 23.2.3.1. Let Ψ be an automorphism of $GL(V)$. Then:

1. $\Psi(SL(V)) = SL(V)$.
2. Up to sign, Ψ transforms transvections into transvections.
3. Either there exists a semi-linear bijection f of V such that for every transvection τ , $(D(\Psi(\tau)), H(\Psi(\tau))) = (f(D(\tau)), f(H(\tau)))$, and we set $\varepsilon(\Psi) = 1$,
4. Or there exists a semi-linear bijection f of V to V^* such that for every transvection τ , $(D(\Psi(\tau)), H(\Psi(\tau))) = (f(H(\tau)), f(D(\tau)))$, and we set $\varepsilon(\Psi) = -1$.

Proof. Ψ maps commutators to 1 and item 1 follows from lemma 7.7.1.1.

Lemmas 23.2.1.1, 23.2.2.2 and 23.2.2.3 provide a theoretic characterization of transvections giving item 2.

For any line D and hyperplane H , by (23.2.2.5) there exists a unique line (resp. hyperplane) $f(D)$ and a unique hyperplane (resp. line) $f(H)$

$$\Psi(\Sigma_D) = \Sigma_{f(D)} \text{ and } \Psi(\Sigma_H) = \Sigma_{f(H)}$$

and f is bijective because so is Ψ .

Let D_1, D_2 be two lines and $g \in GL(V) | g(D_1) = D_2$. Then, $g\Sigma_{D_1}g^{-1} = \Sigma_{D_2}$ implying

$$\Sigma_{\gamma(f(D_1))} = \gamma\Sigma_{f(D_1)}\gamma^{-1} = \Sigma_{f(D_2)}$$

with $\gamma = \Psi(g)$. By (23.2.2.6), we have $\gamma(f(D_1)) = f(g(D_1))$ proving that f is always a line -we set $\varepsilon(\Psi) = 1$ in this case- or always a hyperplane -we set $\varepsilon(\Psi) = -1$ in this case-.

Moreover, by lemma 23.2.2.5, this formula shows

$$(i) \quad f \circ \Psi(g) = g \circ f \text{ for any } g \in \text{GL}(V)$$

with

- $f \in \text{Map}(\mathbf{P}V, \mathbf{P}V)$ if $\varepsilon(\Psi) = 1$ satisfying $g(\langle d \rangle) = \langle g(d) \rangle$ for any $\langle d \rangle \in \mathbf{P}V$ and $D(\Psi(\tau)) = f(D(\tau))$ for any transvection τ .
- $f \in \text{Map}(\mathbf{P}V, \mathbf{P}V^*)$ if $\varepsilon(\Psi) = -1$ with $g(\text{Ker}(\varphi)) = \text{Ker}({}^t g^{-1}(\varphi))$ for any $\text{Ker}(\varphi) \in \mathbf{P}V^*$ and $H(\Psi(\tau)) = f(D(\tau))$ if $\varepsilon(\Psi) = -1$ for any transvection τ .

Lemma 23.2.3.2. *The restriction of f to $\mathbf{P}V$ preserves collinearity.*

Proof. Let first observe that f preserves the incidence relation $D \subset H$. Precisely, let $D \subset H$ and $\tau \in \Sigma_D \cap \Sigma_H$ a transvection of type (D, H) . Then, $\Psi(\tau)$ is a non central and belongs to $\sigma_{f(D)} \cap \sigma_{f(H)}$: it is a transvection of type $(f(D), f(H))$ if $\varepsilon(\tau) = 1$ and $(f(H), f(D))$ if $\varepsilon(\tau) = -1$. Therefore, $f(D) \subset f(H)$ or $f(H) \subset f(D)$ depending on the value of $\varepsilon(\Psi)$: f preserves incidence of pairs (D, H) .

Let D_1, D_2 two distinct points of $\mathbf{P}V$. Then, by duality $D \in \langle D_1, D_2 \rangle$ if and only if all hyperplanes H containing D_1 and D_2 contain also D . If $\varepsilon(\Psi) = 1$, let $H' = f(H)$ an hyperplane of V containing $f(D_1)$ and $f(D_2)$. Then, H' contains D_1 and D_2 and therefore D by assumption. Because f preserves incidence, H' contains $f(D)$ and f preserves collinearity. Dually, if $\varepsilon(\Psi) = -1$, let $D' = f(D)$ a line contained in both $f(D_1)$ and $f(D_2)$. Then, D' is collinear with D_1 and D_2 and, because f preserves incidence, D' is contained in $f(D)$. By the fundamental theorem of projective geometry, f is induced by some sesquilinear bijection of $V \xrightarrow{\sim} V$ if $\varepsilon(\Psi) = 1$ and some sesquilinear bijection of $V \xrightarrow{\sim} V^*$, abusively denoted by f . \square

\square

Theorem 23.2.3.3. *Let Ψ be an automorphism of $\text{GL}(V)$.*

1. *Then, there exists a group isomorphism $\chi : \mathbf{k}^* \rightarrow \mathbf{k}^*$ such that Ψ is either of the form $g \mapsto \chi(\det(g))fgf^{-1}$, where f is a semi-linear isomorphism $V \rightarrow V$, or $g \mapsto \chi(\det(g))f^t g^{-1} f^{-1}$, where f is a semi-linear isomorphism $V \rightarrow V^*$.*
2. *Conversely, such morphism is an automorphism if and only if $\chi^n \cdot \text{Id}$ is a semi-linear automorphism of \mathbf{k}^* compatible to Ψ .*

Proof. With the notation of the above proposition, let us define $\Phi \in \text{Aut}(GL(V))$ by $\Phi : g \mapsto f \circ g \circ f^{-1}$ if $\varepsilon(\Psi) = 1$ and $\Phi : g \mapsto f \circ {}^t g^{-1} \circ f^{-1}$ if $\varepsilon(\Psi) = -1$. Then, $\delta = \Psi \circ \Phi^{-1} \in \text{Aut}(GL(V))$ and thanks (i) shows for all $g \in GL(V)$ and for all line D , one has $\delta(g)(D) = D$ proving that $\delta(g)$ is a scalar matrix $\chi(g)\text{Id}$ for some group morphism $\chi : GL(V) \rightarrow \mathbf{k}^*$. Then item 1 follows from the equality $\text{Ker}(\det) = \text{SL}(V) = \text{D}(GL(V))$.

Conversely, such a morphism induces a commutative diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \text{SL}(V) & \longrightarrow & \text{GL}(V) & \xrightarrow{\det} & \mathbf{k}^* \longrightarrow 0 \\ & & \downarrow \Psi & & \downarrow \Psi & & \downarrow \chi^n \cdot \text{Id} \\ 0 & \longrightarrow & \text{SL}(V) & \longrightarrow & \text{GL}(V) & \xrightarrow{\det} & \mathbf{k}^* \longrightarrow 0 \end{array}$$

with vertical isomorphisms if and only if Ψ is an isomorphism. □

23.2.4 Automorphisms of \mathbf{k}^*

We will illustrate that the automorphism group $\text{Aut}(\mathbf{k}^*)$ can be very diverse and in general is huge. This dramatically changes if one makes some continuity assumption.

Proposition 23.2.4.1.

1. There is a canonical group isomorphism⁵ $\mathbf{Q}^* \xrightarrow{\sim} \{\pm 1\} \times \mathbf{Z}^{(\mathcal{P})}$ identifying $\text{Aut}(\mathbf{Q}^*)$ with $GL_{\mathbf{Z}}\mathbf{Z}^{(\mathcal{P})}$. In particular, it contains the uncountable groups $\text{Bij}(\mathcal{P})$ and $(\mathbf{Z}/2\mathbf{Z})^{\mathcal{P}}$.
2. There is a canonical continuous group isomorphism $\mathbf{R}^* \xrightarrow{\sim} \{\pm 1\} \times \mathbf{R}$ identifying $\text{Aut}(\mathbf{R}^*)$ with $GL_{\mathbf{Q}}(\mathbf{R})$. In particular, $\text{Aut}(\mathbf{R}^*)$ is uncountable.
3. Any group continuous automorphisms of \mathbf{R}^* is of the form $t \mapsto \text{signe}(t)|t|^\chi$ for $\chi \in \mathbf{R}^*$ defining an isomorphism $\text{Aut}_c(\mathbf{R}^*) \xrightarrow{\sim} \mathbf{R}$.
4. If \mathbf{k} is finite of cardinality q , the group $\text{Aut}(\mathbf{k}^*)$ is cyclic of order $\varphi(q - 1)$ where φ is the Euler totient function.

Proof. We we'll denote by Ψ an element of $\text{Aut}(\mathbf{k}^*)$.

Proof of (1). The unique decomposition of an integer into a finite number of powers of prime numbers define an isomorphism $\mathbf{Q}^* \xrightarrow{\sim} \{\pm 1\} \times \mathbf{Z}^{(\mathcal{P})}$. Because -1 is the unique order 2 element, $\Psi(-1) = -1$ proving the two first assertions. Let X be a subset of $\mathbf{N} \simeq \mathcal{P}$ and let us chose σ_X some fixed point free bijection of X . We lift σ_X to a bijection of \mathbf{N} by the identity on the complement of X . Then $X \mapsto \sigma_X$ is a set injection $\mathcal{P}(\mathbf{N}) \hookrightarrow \text{Bij}(\mathbf{N})$. But the map 2-adic expansion $X \mapsto \sum_{i=1}^{\infty} (1_X(i) + 1)2^{-i}$ is a surjection $\mathcal{P}(\mathbf{N}) \rightarrow [0, 2]$ hence the uncountability. Finally, the map $X \mapsto 1_X$ is a bijection $\mathcal{P}(\mathcal{P}) \rightarrow (\mathbf{Z}/2\mathbf{Z})^{\mathcal{P}}$ (with inverse $(x_p) \mapsto \{p|x_p = 0\}$ giving the second uncountability statement.

⁵ \mathcal{P} is the set of prime number

Proof of (2). The continuous morphism $(\pm 1, x) \mapsto \pm e^x$ has inverse $y \mapsto (\text{sign}(y), \ln |y|)$ and is the wanted isomorphism. Any additive morphism of \mathbf{R} is \mathbf{Q} -linear giving the second point. The last one follows from the fact that a \mathbf{Q} -basis of \mathbf{R} is uncountable.

Proof of (3). By (2), an continuous isomorphism of \mathbf{R}^* is defined by a continuous isomorphism f of \mathbf{R} . If $F(x) = \int_0^x f(t)dt$, we get from $f(x+y) = f(x) + f(y)$ the formula $F(x+y) - F(y) = F(x) + xf(y)$ proving that f is C^1 on \mathbf{R} as linear combination of C^1 functions. Differentiating the previous relation, we get $f'(x+y) = f'(x)$ proving that f' is a constant $\chi \in \mathbf{R}$ and $f(x) = \chi x + f(0) = \chi x$. The condition $\chi \neq 0$ ensures that Ψ is bijective.

Proof of (4). This a direct consequence of the well-known fact (??) that for any field \mathbf{k} , a subgroup of \mathbf{k}^* is cyclic. In particular, so is \mathbf{k}^* if \mathbf{k} is finite. \square

23.2.5 Normal subgroups of $GL(V)$

We will explain the so-called Iwasawa to study normal subgroups of perfect groups G , or equivalently we will give a criterium of simplicity of $G/Z(G)$ where $Z(G)$ is the centrum of G .

Definition 23.2.5.1. *Let G be a group acting on a set X .*

1. *We say G acts primitively on X if:*

(a) *The action of G on X is transitive;*

(b) *The stabilizer G_x of a point $x \in X$ is a maximal subgroup of G .*

2. *We say G acts 2-transitively on X if for all $x_1, x_2, y_1, y_2 \in X$, $x_1 \neq x_2$, $y_1 \neq y_2$, there exists $g \in G$ such that $g \cdot x_1 = y_1$ and $g \cdot x_2 = y_2$.*

For instance, $SL(V)$ and $GL(V)$ act 2-transitively on $\mathbf{P}V$ if $\dim(V) \geq 2$.

Proposition 23.2.5.2 (Iwasawa criterium). *Suppose the group G acts primitively on X . If, for each $x \in X$, we are given a subgroup $T_x \subseteq G$ such that:*

1. *T_x is abelian;*

2. *$T_{g \cdot x} = gT_xg^{-1}$ for all $g \in G$ and $x \in X$;*

3. *$\bigcup_{x \in X} T_x$ generates G .*

Then any nontrivial normal subgroup of G acting on X contains $D(G)$.

Corollary 23.2.5.3. *If $\dim(V) \geq 2$, any normal nontrivial normal subgroup of $GL(V)$ (or $SL(V)$) contains $SL(V)$ unless \mathbf{k} is a field with 2 (or 8) elements.*

Proof. Take $X = \mathbf{P}(V)$ and $T_x \xrightarrow{\sim} \mathbf{k}$ be the group of transvections with vector x . It satisfies the hypotheses of the theorem, so a normal subgroup of $PSL(n, \mathbf{K})$, not reduced to $\{\text{Id}\}$, must contain $D(PSL(n, \mathbf{K})) = PSL(n, \mathbf{K})$ by Theorem 2.1.9. □

The action of $PSL(n, \mathbf{K})$ on $X = \mathbf{K}\mathbf{P}^{n-1}$ is 2-transitive, hence primitive. We apply Theorem 2.1.10 using for $x \in X$ the group of transvections with vector x . It satisfies the hypotheses of the theorem, so a normal subgroup of $PSL(n, \mathbf{K})$, not reduced to $\{\text{Id}\}$, must contain $D(PSL(n, \mathbf{K})) = PSL(n, \mathbf{K})$ by Theorem 2.1.9.

23.2.6 Additional exercises

Exercise(s) 23.2.6.1. ⁶ *Let \mathbf{k} be a field of characteristic different from 2, and $m \geq 3$. Let $V = \mathbf{k}^{2m}$, equipped with the standard alternating bilinear form b and $Sp_{2m}(\mathbf{k})$ the corresponding symplectic group. Let $s, t \in Sp_{2m}(\mathbf{k})$ be two involutions.*

1. *Prove that one can write a decomposition*

$$V = E_+(s) \oplus E_-(s),$$

where $E_+(s)$ and $E_-(s)$ are the eigenspaces of s corresponding to the eigenvalues 1 and -1 , respectively.

2. *Deduce a bijection between the set of involutions of $Sp_{2m}(\mathbf{k})$ and the set of non-degenerate subspaces of V .*

Define the type of an involution s as $(2r, 2m - 2r)$ if the dimension of $E_+(s)$ is $2r$. An involution of type $(2, 2m - 2)$ or $(2m - 2, 2)$ is called extremal. In this case, denote by $E_2(s)$ the eigenspace $E_{\pm}(s)$ of dimension 2.

3. *Considering maximal commutative families of conjugate involutions in $Sp_{2m}(\mathbf{k})$, prove that any automorphism of $Sp_{2m}(\mathbf{k})$ maps an extremal involution to an extremal involution.*

As for the general linear group, we define for such a minimal pair of involutions $\{s, t\}$ the "commutant" $\{s, t\}^{\diamond}$ as the set of extremal involutions of $Sp_{2m}(\mathbf{k})$ commuting with both s and t .

4. *Show that s and t form a minimal pair if and only if:*

(a) $st \neq ts$, and

(b) for all $s', t' \in \{s, t\}^{\diamond}$ with $s't' \neq t's'$, it holds that

$$\{s, t\}^{\diamond\diamond} = \{s', t'\}^{\diamond\diamond}.$$

⁶Thanks to O. Debarre for this exercise

5. Determine the maximal sets I of extremal involutions such that every pair of elements in I either forms a minimal pair or commutes. [If $\pm s, \pm t, \pm u$ distincts elements of I , observe $E_2(s) \cap E_2(t) \cap E_2(u)$ is a line V_1 or is zero and we define in this case $V_3 = E_2(u) \subseteq E_2(s) + E_2(t) =: V_3$. Show that I is of of the form $I_1(V_1) := \{v \text{ extremal involution} \mid V_1 \subseteq E_2(v)\}$ or $I_3(V_3) := \{v \text{ extremal involution} \mid E_2(v) \subseteq V_3\}$.

6. Prove that every automorphism of $\mathrm{Sp}_{2m}(\mathbf{k})$ is of the form

$$x \mapsto axa^{-1}$$

for some semi-linear transformations of \mathbf{k}^n preserving b .

Exercise(s) 23.2.6.2. Adapt the arguments of (23.2.6.1) to prove that automorphisms of $\mathrm{SO}_{2n}(\mathbf{R})$, $n \geq 3$ are inner automorphisms.

Exercise(s) 23.2.6.3. Show that the continuous isomorphisms of \mathbf{C}^* are given by $z \mapsto |z|^\chi (z/|z|)^{\pm 1}$ with $\chi \in \mathbf{C}^*$.

Part V

Supplements

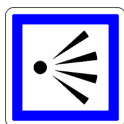
Chapter 24

Bilinear forms without symmetries



Herbert Westren Turnbull

24.1 Perspective



We chose this theme, beyond its mathematical importance, to demonstrate a somewhat unexpected link between the congruence of invertible matrices M and similarities of their co-squares (19.3.1.2) ${}^tM^{-1}M$ which is evidently hidden in the case of ε -symmetric forms. This subject has a long history (see [13]) for a history.

24.2 Introduction

We are interested in the congruence $A \mapsto {}^tPAP$ with $P \in GL_n(\mathbf{k})$ and $A \in M_n(\mathbf{k})$. For example, if J_d is a Jordan block, J_d and tJ_d are congruent via a diagonal congruence as long as t is non-zero. If P is a

permutation matrix, ${}^t\text{PAP}$ is deduced from A by permuting the rows and then the columns of the same index. Similarly, if $P = T_{i,j}(\lambda)$ is a matrix of transvection, we deduce ${}^t\text{PAP}$ from A by adding λ times the i -th column to the j -th column and then adding λ times the i -th row to the j -th row. We will talk about permitted congruences \equiv .

In particular, if X is a non-zero element of the kernel of A , we can reduce by permitted congruences to a first null column (write $X = \sum x_i e_i$ with $\ell = x_i \neq 0$, make the congruence associated with the transpositions $(1, i)$ if $i > 1$ then the permitted operations associated with $T_{1,j}(-a_{1,j}/\ell), j > 1$).

The complete classification¹ in the algebraically closed case is obtained in 24.6.0.6. We know that we cannot hope for a result on any field, even in dimension 1! As we will see, the existence of square roots is key to this classification, as in the complex symmetric case (21.8.0.1) or real (21.9.0.1). But as soon as they exist, classification is possible.

There is no difficulty in adapting to the sesquilinear case. We leave that to the interested reader. Addressing the bilinear case simply has the advantage of simplifying the notations.

24.3 Existence of a Decomposition

The following result on matrix decomposition allows us to reduce the study of bilinear forms to that of non-degenerate forms. Strangely, it is quite recent, due to P. Gabriel [15]. We give a simplified version essentially derived from [21] (for the existence part) and [14] (for the uniqueness part). This lemma is obvious in the \pm -symmetric (or hermitian) case because of the coincidence of the kernels of M and tM . The intersection $\text{Ker}(M) \cap \text{Ker}({}^tM)$ clearly appears in the proof of the following result.

Lemma 24.3.0.1. *Every matrix is (algorithmically) congruent to a block diagonal matrix $\text{diag}(A, J_{\underline{d}})$ where $A \in \text{GL}_r(\mathbf{k})$, $r = \text{rank}(A)$ and $J_{\underline{d}} = \text{diag}(J_{d_i})$ is the block diagonal Jordan matrix of size d_i associated with a partition $\underline{d} = (d_i)$ of $n - r$.*

PROOF. *It is sufficient to show that every matrix is (algorithmically) congruent to a block diagonal matrix $\text{diag}(A, J(\underline{\delta}))$ where $A \in \text{GL}_r(\mathbf{k})$, $r = \text{rank}(A)$, $\underline{\delta} \in (\mathbf{k} - \{0\})^{n-r}$, and $J(\underline{\delta}) = \text{diag}(\delta_i J_{d_i})$ is the block diagonal Jordan matrix of size d_i associated with a partition $\underline{d} = (d_i)$ of $n - r$. Indeed, a diagonal congruence leads to the desired form.*

We proceed by induction on $r = \text{rank}(M)$. We can assume $0 < r < n$ and the theorem proven at rank $r - 1$. From the previous remark, we can therefore assume after permitted operations that the first column of M is null ($e_1 \in \text{Ker}(M)$):

$$M \equiv \begin{pmatrix} 0 & L_{1,n-1} \\ 0 & M'_{n-1,n-1} \end{pmatrix}.$$

¹With the system of representatives of Turnbull and Aitken

If $e_1 \in \text{Ker}({}^tM)$ i.e. if $L_{1,n-1} = 0$, then $M = \text{diag}(0, M')$ with M' of rank $r - 1$ and we conclude by induction. Otherwise, a coefficient $\delta_1 = L_{1,j}, j > 1$ is non-null. Since the first column is null and therefore invariant by row operations, we can reduce by permitted congruences to $j = 2$. Using δ_1 as a pivot, we reduce by permitted congruences first to

$$M \equiv \begin{pmatrix} 0 & \delta_1 & 0_{1,n-2} \\ 0 & \gamma & L'_{1,n-2} \\ 0 & C_{n-2,1} & M'_{n-2,n-2} \end{pmatrix}.$$

The permitted congruence subtracting from the 2-nd row γ/δ_1 times the first does not change the second column since the first column is null!. Likewise for rows with index $> n - 1$. Thus, we have a permitted congruence

$$M \equiv \begin{pmatrix} 0 & \delta_1 & 0_{1,n-2} \\ 0 & 0 & L_{1,n-2} \\ 0 & 0_{n-2,1} & M''_{n-2,n-2} \end{pmatrix}.$$

If $L_{1,n-2} = 0$, we apply the induction hypothesis to $M''_{n-2,n-2}$.

If $L_{1,n-2} \neq 0$, one of the coefficients δ_2 of $L_{1,n-2} = 0$ is non-null: it is found in a column of \tilde{M} with index $j \geq 3$. By permitted operation, we can assume $j = 3$ without changing the first two columns of \tilde{M} so that we have

$$M \equiv \begin{pmatrix} 0 & \delta_1 & 0 & 0_{1,n-3} \\ 0 & 0 & \delta_2 & L'_{1,n-3} \\ 0 & 0 & C_{n-3,1} & M_{n-3,n-3} \end{pmatrix}$$

As before, by permitted operations on the rows with indices > 2 , we can assume that the coefficient of index $(2,3)$ is the only non-null coefficient in its column and therefore

$$M \equiv \begin{pmatrix} 0 & \delta_1 & 0 & 0_{1,n-3} \\ 0 & 0 & \delta_2 & L_{1,n-3} \\ 0 & 0 & 0_{n-3,1} & M''_{n-3,n-3} \end{pmatrix}.$$

If $L_{1,n-3} = 0$, we apply the induction hypothesis to $M''_{n-3,n-3}$. Otherwise, we iterate the process which is clearly finite thus completing the sought existence. Subject to multiplying by a diagonal matrix on the right and left, the δ s can be modified by multiplication by non-null squares hence the second point.

Remark(s) 24.3.0.2. Note that in the usual symmetric (or alternating, or Hermitian) case, considering a basis obtained by completing a basis of the kernel of M immediately gives the lemma, with A being congruent to the matrix of the form Ψ induced by M on $\mathbf{k}^n/\text{Ker}(\Psi)$. This easily yields a statement of uniqueness in this case. Even replacing, for example, the kernel by the left kernel $L(V)$ of $V = (\mathbf{k}^n, \Psi)$, the

problem in the general case, as we shall see, is that Ψ only factorizes over a certain subspace $L^2(V)/L(V)$ of the quotient $V/L(V)$.

24.4 The Typical Bilinear Space

Note that at this stage, it is not clear that neither \underline{d} nor the congruence class of A are determined by the congruence class of M .

We simply denote V_M as the bilinear space \mathbf{k}^n equipped with the bilinear form $(X, Y) \mapsto {}^tXMY$ (with matrix M in the canonical basis)

We simply note V_d as the space $V_{J_d} = \mathbf{k}[T]/(T^d)$ associated with the standard Jordan block of size $d \geq 1$ seen as the matrix of multiplication by T in the base of monomials as usual. Hence, $(T^i, T^j) = \delta_{j+1, i}$. $V_d = \{0\}$ if $d \leq 0$ because J_d is the empty matrix.

We have (for example) the right kernel

$$R(V) = \{v \in V \mid \Psi(V, v) = \{0\}\}$$

and the left kernel

$$L(V) = \{v \in V \mid \Psi(v, V) = \{0\}\}$$

which is an invariant by isomorphism of bilinear spaces. We then immediately have,

$$R(V_M) = \text{Ker}(M) \text{ and } R(V_M) = \text{Ker}({}^tM).$$

We also define

$$R^2(V) = \{v \in V \mid \Psi(R(V), v) = \{0\}\}.$$

This is a subspace that contains $R(V)$ and Ψ induces a form on $R^2(V)/R(V)$ making it a bilinear space.

Exercise(s) 24.4.0.1. Show that $R^2(V)/R(V)$ is the largest subspace of $V/R(V)$ over which Ψ is quotiented.

For $M = J_d$ and any $d \geq 1$, we have

- $R(V_d) = \langle T^{d-1} \rangle \times_{T^{1-d}} V_1$ and $L(V_d) = \langle 1 \rangle \simeq V_1$
- $\dim R(V_d) \cap L(V_d) = \delta_{1, d}$
- $R^2(V_d) = \langle 1, T, \dots, T^{d-3}, T^{d-1} \rangle$
- $R^2(V_d)/R(V_d) = T^2\mathbf{k}[T]/(T^{d-1}) \times_{T^{-2}} \mathbf{k}[T]/(T^{d-3}) = V_{d-2}$.

For A invertible, we have

- $R(V_A) = \{0\}$

- $R^2(V_A) = V_A$
- $R^2(V_A)/R(V_A) = V_A$.

We say that two subspaces V_1, V_2 of a bilinear space V are in orthogonal direct sum if they are in direct sum and if $(V_1, V_2) = (V_2, V_1) = \{0\}$, and we then write $V = V_1 \overset{\perp}{\oplus} V_2$. Then

$$R(V_1 \overset{\perp}{\oplus} V_2) = R(V_1) \overset{\perp}{\oplus} R(V_2) \text{ and } R^2(V_1 \overset{\perp}{\oplus} V_2) = R^2(V_1) \overset{\perp}{\oplus} R^2(V_2)$$

24.5 Uniqueness

With the previous notations, the lemma leads to the existence of an isomorphism of bilinear spaces

$$V_M \simeq V_A \overset{\perp}{\oplus} \overset{\perp}{\oplus}_{d \in \underline{d}} V_d.$$

Ultimately, we want to recover the isomorphism classes of V_A (thus of congruence of A) and of $V_{\underline{d}}$ and even the partition \underline{d} . Specifically, with obvious notations, we want to show

Lemma 24.5.0.1. *If there are isomorphisms of bilinear spaces*

$$V_A \overset{\perp}{\oplus} \overset{\perp}{\oplus}_{d \in \underline{d}} V_d \simeq V_{A'} \overset{\perp}{\oplus} \overset{\perp}{\oplus}_{d' \in \underline{d}'} V_{d'}$$

with A and A' invertible, then A and A' are congruent and $\underline{d} = \underline{d}'$.

PROOF. *To do this, we will proceed by induction on $n + \max(\underline{d})$ simply by calculating the bilinear spaces $R(V_M)$ and $R^2(V_M)/R(V_M)$ (which are invariants by isomorphisms of bilinear spaces, let's recall). We apply the formulas from the previous section that give*

1. $R(V_M) \simeq \overset{\perp}{\oplus}_{\substack{d \in \underline{d} \\ d \geq 1}} V_d$
2. $R^2(V_M)/R(V_M) = V_A \overset{\perp}{\oplus} \overset{\perp}{\oplus}_{\substack{d \in \underline{d} \\ d \geq 3}} V_{d-2}$

We thus have $R \simeq R'$ and $R^2/R \simeq R'^2/R'$.

- *If $\dim(R^2/R) < n$, the induction hypothesis ensures A and A' are congruent and $\underline{d}, \underline{d}'$ coincide on their elements $d, d' \geq 3$. But then $\dim R \cap L = \text{Card}\{d \in \underline{d} | d = 1\}$ and $\dim R(V_d) = \text{Card}\{d \in \underline{d} | d \geq 1\}$ show that the values 1 and 2 appear with the same weight in \underline{d} and \underline{d}' such that in fine $\underline{d} = \underline{d}'$.*
- *Assume $\dim(R^2/R) = n$.*
 - *If $\max(\underline{d}) \geq 3$, we can also apply the induction hypothesis and conclude as before.*
 - *Otherwise, if $\max(\underline{d}) \leq 2$ then $R^2/R = V_A = V_{A'}$ showing that A and A' are congruent in this case. But in this case also, considering R and $R \cap L$ shows as before $\underline{d} = \underline{d}'$ since there are no elements $d \in \underline{d} \cup \underline{d}'$ such that $d \geq 2$.*

From lemmas 24.3.0.1 and 24.5.0.1 we immediately deduce

Theorem 24.5.0.2. *Every matrix $M \in M_n(\mathbf{k})$ is (algorithmically) congruent to a block diagonal matrix $\text{diag}(A, J_{\underline{d}})$ where $A \in GL_r(\mathbf{k})$, $r = \text{rank}(A)$ and $J_{\underline{d}} = \text{diag}(J_{d_i})$ is the block diagonal Jordan matrix of size d_i associated with a partition $\underline{d} = (d_i)$ of $n - r$. Furthermore, the congruence class of A and \underline{d} are uniquely determined by the congruence class of M .*

Remark(s) 24.5.0.3. *The reader will adapt these results without any difficulty to the sesquilinear case.*

24.6 Classification: Algebraically Closed Case

Here we assume \mathbf{k} is algebraically closed (or simply that every element is a square). The following result is old, probably due to Turnbull as early as 1936 ([30]). Here I revisit the approach from [19].

Proposition 24.6.0.1. *Two invertible matrices A, B are congruent if and only if the asymmetries² $A^{-1}A$ and ${}^tB^{-1}B$ of the associated forms are similar, i.e., have the same similarity invariants.*

Note that this is consistent with the classification of quadratic forms, alternating (or Hermitian, cf. remark supra) in the algebraically closed case.

PROOF. *If M is invertible, we set $M' = {}^tM^{-1}M$ (called the cosquare in the literature).*

If ${}^tPAP = B$ with P invertible, we have $P^{-1}A'AP$ and $B'B$, hence the direct sense.

Conversely, suppose that $A'A$ and $B'B$ are similar. Then considering the pencils $A + {}^tAT$ and $B + {}^tBT$, there exist according to 5.5.0.4, P, Q invertible such that

$$PAQ = B \text{ and } P^tAQ = {}^tB \text{ and thus } {}^tQA^tP = B.$$

We deduce

$$\begin{aligned} PAQ &= {}^tQA^tP \\ XA &= B^tX \text{ with } X = Q'P \text{ and thus} \\ \Xi A &= B^t\Xi \text{ for any } \Xi \in \mathbf{k}[X] \end{aligned}$$

As X is invertible and \mathbf{k} algebraically closed, we then choose $\Xi \in \mathbf{k}[X]$ such that $\Xi^2 = X$ (8.3.2.1). We then have

²see 19.3.1.2.

$$\begin{aligned}
 B &= {}^tQA{}^tP \\
 &= {}^tQA{}^tXQ \\
 &= {}^tQA({}^t\Xi)^2Q \\
 &= {}^tQ(A{}^t\Xi){}^t\Xi Q \\
 &= {}^tQ(\Xi A){}^t\Xi Q \\
 &= {}^t({}^t\Xi Q)A({}^t\Xi Q)
 \end{aligned}$$

Exercise(s) 24.6.0.2. *Retrace the result 21.13.0.1.*

We now need to find a family of matrices not congruent pairwise that describes the possible similarity invariants of cosquares. There are inevitably restrictions since the determinant of a cosquare is 1. If P is a unitary polynomial such that $P(0) \neq 0$, we denote

$$P^*(T) = \frac{T^{\deg(P)}P(1/T)}{P(0)}$$

its polynomial (unitary) of inverses. If A is invertible with similarity invariants $\underline{P} = (P_i)$, we note $\underline{P}^* = (P_i^*)$.

Lemma 24.6.0.3. *Let \underline{P} be the invariants of A invertible.*

1. *The family of similarity invariants of A^{-1} is \underline{P}^* .*
2. *If A is a cosquare, then $\underline{P} = \underline{P}^*$.*
3. *The invariants*

PROOF. *Let P be a similarity invariant of A . As A is invertible, T is invertible in $V = \mathbf{k}[T]/(P)$ (with inverse $(P(T) - P(0))/(TP(0))$). But $V_{T^{-1}}$ is a cyclic $\mathbf{k}[T^{-1}]$ module (generated by any monomial annihilated by P^* which is therefore its minimal for dimension reasons, hence (1)). As we have already noted, ${}^tBT - B$ is equivalent to $\text{Id} - {}^tB^{-1}B$ such that the similarity invariants of $A = {}^tB^{-1}B$ are the invariant factors of ${}^tBT - B$, thus coincide with those of its transpose $TB - {}^tB$ therefore with the similarity invariants of $B^{-1}{}^tB = A^{-1}$.*

Designate by Λ_{\pm} the set of (unordered) pairs $\bar{\lambda} = \{\lambda, \lambda^{-1}\}$ -thus with $\lambda \neq \pm 1$. The similarity invariants of a cosquare with spectrum are therefore

$$P_i(T) = \prod_{\Lambda_{\pm}} [(T - \lambda)(T - \lambda^{-1})]^{v_{\bar{\lambda},i}} (T - 1)^{v_{i,+}} (T + 1)^{v_{i,-}}$$

so that the associated Jordan blocks are (with a minor abuse of notation)

$$\text{diag}(\lambda \text{Id} + J_{v_{\bar{\lambda},i}}, \lambda^{-1} \text{Id} + J_{v_{\bar{\lambda},i}}, \text{Id}_{v_{i,+}}, -\text{Id}_{v_{i,-}})$$

In other words, the Jordan reduction of a cosquare has two blocks $\pm \text{Id}$ of possibly different sizes³ and blocks of type

$$(i) \quad \text{diag}(\lambda \text{Id} + J_d, \lambda^{-1} \text{Id} + J_d) \text{ with } \lambda \neq \pm 1$$

Remark(s) 24.6.0.4. *If β is the asymmetry of the bilinear space V and λ, μ are eigenvalues, it is easily verified that the (sums of) characteristic spaces $V[\mathbb{T} - \lambda] + V[\mathbb{T} - \lambda^{-1}]$ and $V[\mathbb{T} - \mu]$ are orthogonal as long as $\lambda' \notin \{\lambda, \lambda^{-1}\}$. This is a general phenomenon, even if the field is not algebraically closed, by replacing characteristic spaces with primary components associated with unitary irreducibles P, P^* on one hand and $Q \notin \{P, P^*\}$ on the other. Thus, the Jordan decomposition, with blocks suitably grouped, corresponds to an orthogonal decomposition!*

It remains to exhibit a family of matrices whose cosquares are as in i. We will look at the representatives of [19] (other classical choices exist, see for example [20]).

Lemma 24.6.0.5. *Let $d > 0$, $V = \mathbb{T}^{-d} \mathbf{k}_{\leq 2d}[\mathbb{T}] / \mathbf{k}.1$ and $\lambda \neq \pm 1$. Let $a \in \text{End}_{\mathbf{k}}(V)$ be defined by*

$$\begin{aligned} a(\mathbb{T}^i) &= \mathbb{T}^{-i} \text{ if } i < 0 \\ &= \lambda \mathbb{T}^{-i} + \mathbb{T}^{-i+1} \text{ if } i > 0. \end{aligned}$$

Then, the Jordan reduction of the cocarré of a is $\text{diag}(\lambda \text{Id} + J_d, \lambda^{-1} \text{Id} + J_d)$.

PROOF. *Let A be the matrix of a in the basis $\mathbb{T}^i, 0 < |i| \leq d$ and P the permutation matrix $i \mapsto -i$. We then notice that the matrix of $(\mathbb{T}^t A - A)P$ is equal to*

$$\text{diag}((\mathbb{T} - \lambda) \text{Id} + J_d, (\lambda \mathbb{T} - 1) \text{Id} + \mathbb{T} J_d) \approx \text{diag}((\mathbb{T} - \lambda) \text{Id} + J_d, (\mathbb{T} - \lambda^{-1}) \text{Id} + \mathbb{T} J_d).$$

Let us consider the principal ring $\mathbf{k}[\mathbb{T}, \mathbb{T}^{-1}]$. As the subdiagonal of each of the two blocks is 1 or \mathbb{T} , the first principal minor of order $d-1$ obtained by deleting the first row and column is 1 or \mathbb{T} and is therefore invertible. We deduce that the invariant factors of these matrices in $\mathbf{k}[\mathbb{T}^{-1}, \mathbb{T}]$, defined up to an invertible is $(1, \dots, (\mathbb{T}\lambda)^d)$ or $(1, \dots, (\mathbb{T} - \lambda^{-1})^d)$. The invariant factors of these matrices in $\mathbf{k}[\mathbb{T}]$ are the same as in $\mathbf{k}[\mathbb{T}^{-1}, \mathbb{T}]$, up to an invertible of $\mathbf{k}[\mathbb{T}^{-1}, \mathbb{T}]$ of the form \mathbb{T}^i . But since they are coprime with \mathbb{T} by hypothesis, they are therefore indeed $(1, \dots, (\mathbb{T}\lambda)^d)$ or $(1, \dots, (\mathbb{T} - \lambda^{-1})^d)$.

Of course, the matrix of the standard non-degenerate alternating bilinear form (20.5) of rank $2d$ has cocarré $-\text{Id}_{2n}$. To distinguish 1 from -1 , we consider the characteristic different from 2.

³The size of the block $-\text{Id}$ is necessarily even in characteristics different from 2: this is the non-degenerate alternating case!

Theorem 24.6.0.6. *Let \mathbf{k} be algebraically closed of characteristic different from 2 and V any bilinear space, r the rank of the bilinear form. Then, there exists a unique partition \underline{d} of $\dim(V) - r$, there exists an orthogonal decomposition $V = W \oplus V_{\underline{d}}$ with W non-degenerate, unique up to bilinear isomorphism. Moreover, W decomposes into a direct orthogonal sum of its symmetric part, its antisymmetric part and non-degenerate spaces of matrices $\text{diag}(\lambda \text{Id} + J_d, \lambda^{-1} \text{Id} + J_d), \lambda \neq \pm 1$. This decomposition is unique up to bilinear isomorphism.*

Chapter 25

Index et bibliography

Index

- adjoint, 160, 220, 249
- adjoint,
 - euclidean, 160, 220
- affine chart, 233
- affine,
 - frame, 241
 - morphism, 241
 - space, 241
 - subspace, 242
 - universal envelop, 242
- algorithm,
 - Gauss, 262
- angle,
 - obtuse, 206
- anisotropic,
 - space, 254
- associated, 137
- basis,
 - ante-dual, 105
 - dual, 104
- bicommutant, 80
- bilinear form,
 - asymmetry, 248
 - degenerate, 247
 - kernel, 247
- co-square, 248
- cokernel, 37
- commutant, 41, 79
- commutative diagram, 44
- commutator, 29
- complex of modules, 42
- content, 141
- convex,
 - hull, 200
- cosquare, 308
- cross product, 153
- decomposition,
 - QR, 155
 - complex QR, 218
 - Complex Iwasawa , 218
 - Iwasawa, 155
- decomposition,LU, 155
- derived subgroup, 29
- determinant trick, 130
- diagram, 44
- diagram,
 - Hasse diagram, 121
- dilatation, 15
- duality bracket, 103
- duality,
 - contravariance, 109
 - convention of biduality, 107
 - differential, 104
 - Jacobian, 104
 - orthogonal, 104
 - polar, 104
 - transpose, 108
- elementary divisors, 63
- ellipsoid, 167
- ellipsoid,
 - John's, 212
 - Loewner's, 176

- endomorphism,
 - cyclic, 77
 - absolutely semisimple, 97
 - diagonalizable, 75
 - normal in the complex case, 161
 - real normal, 161
 - semisimple, 97
- exact sequence, 42
- extremal involution, 286
- extremal point, 209
- factorial, 139, 142
- form,
 - bilinear, 246
 - sesquilinear, 246
- functor, 48
- functoriality,
 - of the cokernel, 44
 - of the kernel, 46
- gauge function, 202
- Gauss,
 - elimination, 27
- Gaussian,
 - elimination, 59
 - pivot, 59
- GCD, 140
- general position, 232
- generalized Euclidean division, 68
- hermitian,
 - scalar product, 216
 - space, 216
- homogeneous coordinates, 232
- homography, 230
- Householder algorithm, 171
- hyperbolic (plane), 264
- identity of the median, 151
- inductive set, 15
- inequality of Cauchy-Schwarz,
 - complex, 217
- inequality,
 - Hadamard, 156
 - Vadamard, 218
- inequality, Cauchy-Schwarz,
 - real, 20, 151
- integer
 - algebraic, 143
- integer (element), 131
- inversion,
 - Euclidean, 179
 - orthogonal, 280
- involution,
 - extremal, 290
 - minimal pair, 291
- irreducibility of Φ_n over \mathbf{Q} , 144
- irreducible, 133
- irreducibles,
 - existence, 133
 - of $\mathbf{R}[T]$, 141
 - uniqueness of the decomposition into, 139
- isotropic,
 - space, 254
 - totally space, 254
 - vector, 254
- Jordan-Chevalley decomposition, 98
- LCM, 140
- lemma
 - of Zorn, 16
- lemma,
 - Farkas, 204
 - Farkas' Lemma, 108
 - five, 46
 - Hensel, 98
 - Morse, 197

- Nakayama, 130
- of Euclid, 138
- map,
 - σ -linear, 246
- matrix,
 - complex normal matrix, 161
 - Gram, 156, 218
- median formula, 217
- module, 34
- module,
 - V_a , 41
 - torsion, 34
 - associated with an endomorphism, 41
 - cyclic, 40
 - monogenic, 40
 - noetherian, 131
 - quotient, 37
 - semi-simple, 94
- monoid, 137
- morphism,
 - Frobenius, 96
- Noetherian,
 - Hilbert's basis theorem, 134
 - ring, 132
- noetherian,
 - module, 131
- order,
 - \leq on partitions, 120
 - \leq on types, 120
 - \preceq on partitions, 121
 - \preceq on types, 120
- orientation, 22
- orientation,
 - direct basis, 22
 - positively oriented basis, 22
- orthogonal group
 - special $SO(q)$, 278
- orthogonal group,
 - of a quadratic form $O(q)$, 278
- partition,
 - of an integer, 81
 - dual, 84
- pencil, 308
- perfect group, 29
- Pfaffian, 256
- polar decomposition,
 - complex, 225
 - real, 178
- polar form, 260
- polynomial
 - cyclotomic, 142
- primitive, 141
- projective completion, 243
- projective space, 230
- projectors,
 - spectral, 113
- Pythagoras theorem, 149, 217
- quadratic form, 260
- quadratic form,
 - index, 266
 - orthogonal, 248, 253
- quadratic pencil, 273
- quaternion, 185
- quotient, 37
- reduction,
 - Jordan, 80
 - real isometries, 164, 222
 - complex quadratic pencil, 274
 - Frobenius, 78
 - Hermitians, 223
 - of complex normal endomorphisms, 221
 - of real normal endomorphisms, 162

- real self-adjoints, 166
- real skew-adjoints, 168
- reflection,
 - Euclidean, 179
 - orthogonal, 280
- ring,
 - Noetherian, 132
 - Noetherian UFD, 140
 - UFD or factorial, 139, 142
- semi-simple,
 - module, 94
- semisimple,
 - endomorphism, 97
- sesquilinear, 216
- signature, 269
- similarity invariants, 73, 74
- similitude,
 - orthogonal, 280
 - Euclidean, 181
- simplicity of $\text{SO}(3, \mathbf{R})$, 180
- skew-field, 186
- space,
 - anisotropic, 254
 - bilinear, 247
 - characteristic, 112
 - sesquilinear, 247
 - stable, 41
 - totally isotropic, 254
- supporting hyperplane, 206
- Sylvester's inertia, 269
- symplectic group, 278
- theorem,
 - of Krein-Milman, 210
 - of Sylvester, 269
 - spectral, 166
 - spectral for Hermitian endomorphisms, 223
- transvection, 14, 28, 109
- type, 120
- UFD, 139, 142
- unitary group, 278
- universal property,
 - of the cokernel, 48
 - of the kernel, 48
 - of the product of modules, 48
 - of the sum of modules, 48
- Witt's simplification, 267

Bibliography

- [1] J. M. Almira. “Müntz type theorems. I”. In: *Surv. Approx. Theory* 3 (2007), pp. 152–194. ISSN: 1555-578X.
- [2] David Bartl. “A very short algebraic proof of the Farkas Lemma”. In: *Math Meth Oper Res* (2012), pp. 101–104.
- [3] H. Bass, J. Milnor, and J.-P. Serre. “Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$)”. In: *Inst. Hautes Études Sci. Publ. Math.* 33 (1967), pp. 59–137. ISSN: 0073-8301,1618-1913. URL: http://www.numdam.org/item?id=PMIHES_1967__33__59_0.
- [4] Anthony J. Bevelacqua. “A family of non-Euclidean PIDs”. In: *Amer. Math. Monthly* 123.9 (2016), pp. 936–939. ISSN: 0002-9890,1930-0972. DOI: 10.4169/amer.math.monthly.123.9.936. URL: <https://doi.org/10.4169/amer.math.monthly.123.9.936>.
- [5] Garrett Birkhoff and John Von Neumann. “The Logic of Quantum Mechanics”. In: *Annals of Mathematics* 37.4 (1936), pp. 823–843. (Visited on 06/28/2024).
- [6] Errett Bishop. *Foundations of constructive analysis*. McGraw-Hill Book Co., New York-Toronto-London, 1967, pp. xiii+370.
- [7] Armand Borel. *Linear algebraic groups*. Second. Vol. 126. Graduate Texts in Mathematics. Springer-Verlag, New York, 1991, pp. xii+288. ISBN: 0-387-97370-2. DOI: 10.1007/978-1-4612-0941-6. URL: <https://doi.org/10.1007/978-1-4612-0941-6>.
- [8] N. Bourbaki. *Algebra. Chapters 4–7*. Springer-Verlag, Berlin, 2007.
- [9] N. Bourbaki. *Éléments de mathématique. Algèbre. Chapitres 1 à 3*. Hermann, Paris, 1970.
- [10] N. Bourbaki. *Topological vector spaces. Chapters 1–5*.
- [11] Sunil K. Chebolu, Dan McQuillan, and Ján Mináč. “Witt’s cancellation theorem seen as a cancellation”. In: *Expo. Math.* 35.3 (2017), pp. 300–314. ISSN: 0723-0869,1878-0792. DOI: 10.1016/j.exmath.2016.10.003. URL: <https://doi.org/10.1016/j.exmath.2016.10.003>.
- [12] Karine Chemla and Shuchun Guo. *The Nine Chapters: A Mathematical Classic of Ancient China and its Commentaries*. Dunod, Paris, 2005.

- [13] Fernando De Terán. “Canonical forms for congruence of matrices and T-palindromic matrix pencils: a tribute to H. W. Turnbull and A. C. Aitken”. In: *SeMA J.* 73.1 (2016), pp. 7–16. ISSN: 2254-3902,2281-7875. DOI: 10.1007/s40324-015-0052-y. URL: <https://doi.org/10.1007/s40324-015-0052-y>.
- [14] Dragomir Z. Doković and Fernando Szechtman. “An elementary proof of Gabriel’s theorem on degenerate bilinear forms and its generalization”. In: *J. Algebra* 279.1 (2004), pp. 121–125. ISSN: 0021-8693,1090-266X. DOI: 10.1016/j.jalgebra.2004.05.014. URL: <https://doi.org/10.1016/j.jalgebra.2004.05.014>.
- [15] Peter Gabriel. “Appendix: degenerate bilinear forms”. In: *J. Algebra* 31 (1974), pp. 67–72. ISSN: 0021-8693. DOI: 10.1016/0021-8693(74)90005-2. URL: [https://doi.org/10.1016/0021-8693\(74\)90005-2](https://doi.org/10.1016/0021-8693(74)90005-2).
- [16] Murray Gerstenhaber. “On dominance and varieties of commuting matrices”. In: *Ann. of Math. (2)* 73 (1961), pp. 324–348. ISSN: 0003-486X. DOI: 10.2307/1970336. URL: <https://doi.org/10.2307/1970336>.
- [17] Daniel R. Grayson. “SK1 of an interesting principal ideal domain”. In: *Journal of Pure and Applied Algebra* 20 (1981), pp. 157–163. URL: <https://api.semanticscholar.org/CorpusID:123294985>.
- [18] David Hilbert. “Ueber die Theorie der algebraischen Formen”. In: *Math. Ann.* 36.4 (1890), pp. 473–534. ISSN: 0025-5831,1432-1807. DOI: 10.1007/BF01208503. URL: <https://doi.org/10.1007/BF01208503>.
- [19] W. V. D. Hodge and D. Pedoe. *Methods of algebraic geometry. Vol. II.* Cambridge Mathematical Library. Book III: General theory of algebraic varieties in projective space, Book IV: Quadrics and Grassmann varieties, Reprint of the 1952 original. Cambridge University Press, Cambridge, 1994, pp. x+394. ISBN: 0-521-46901-5. DOI: 10.1017/CB09780511623899. URL: <https://doi.org/10.1017/CB09780511623899>.
- [20] Roger A. Horn and Vladimir V. Sergeichuk. “Canonical matrices of bilinear and sesquilinear forms”. In: *Linear Algebra Appl.* 428.1 (2008), pp. 193–223. ISSN: 0024-3795,1873-1856. DOI: 10.1016/j.laa.2007.07.023. URL: <https://doi.org/10.1016/j.laa.2007.07.023>.
- [21] Khakim Ikramov. “On the congruent selection of Jordan blocks from a singular square matrix”. In: *Numerical Analysis and Applications* 11 (July 2018), pp. 204–207. DOI: 10.1134/S1995423918030023.
- [22] Nathan Jacobson. *Basic algebra. I.* Second. W. H. Freeman and Company, New York, 1985, pp. xviii+499. ISBN: 0-7167-1480-9.
- [23] Felix Klein. *Le programme d’Erlangen.* Collection “Discours de la Méthode”. Considérations comparatives sur les recherches géométriques modernes, Traduit de l’allemand par H. Padé, Préface de J. Dieudonné, Postface de François Russo. Gauthier-Villars Éditeur, Paris-Brussels-Montreal, Que., 1974, pp. xiv+72.

- [24] J. Milnor. “Whitehead torsion”. In: *Bull. Amer. Math. Soc.* 72 (1966), pp. 358–426. ISSN: 0002-9904. DOI: 10.1090/S0002-9904-1966-11484-2. URL: <https://doi.org/10.1090/S0002-9904-1966-11484-2>.
- [25] Emmy Noether. “Idealtheorie in Ringbereichen”. In: *Math. Ann.* 83.1-2 (1921), pp. 24–66. ISSN: 0025-5831,1432-1807. DOI: 10.1007/BF01464225. URL: <https://doi.org/10.1007/BF01464225>.
- [26] Daniel Perrin. *Cours d’algèbre*. Ellipses, Paris, 1988, p. 212. ISBN: 2-85929-011-7.
- [27] Hjalmar Rosengren. *Proof of the duality of the dominance order on partitions*. <https://math.stackexchange.com/q/3429855>. 2020.
- [28] Walter Rudin. *Real and complex analysis*. Third. McGraw-Hill Book Co., New York, 1987, pp. xiv+416. ISBN: 0-07-054234-1.
- [29] Jean-Pierre Serre. *Cours d’arithmétique*. Vol. No. 2. Le Mathématicien [The Mathematician]. Deuxième édition revue et corrigée. Presses Universitaires de France, Paris, 1977, p. 188.
- [30] H. W. Turnbull and A. C. Aitken. *An introduction to the theory of canonical matrices*. Dover Publications, Inc., New York, 1961, pp. xiii+200.
- [31] B. L. van der Waerden. *Modern Algebra. Vol. II*. Frederick Ungar Publishing Co., New York, N. Y., 1950, pp. v+227.