Linear Algebra within 100 pages

Boyer Pascal

22 septembre 2024

Table des matières

1 Vector Spaces

In the following, $\mathbb{K}$ is a field that we may initially assume to be equal to $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$.

1.1 Generalities

Definition 1. *A $\mathbb{K}$-vector space is a triplet $(E, +, .)$ where*
- *$(E, +)$ is a commutative group,*
- *equipped with an external operation $(\lambda, e) \in \mathbb{K} \times E \mapsto \lambda.e \in E$ that satisfies the following properties :*
  - *for all $\lambda, \mu \in \mathbb{K}$ and for all $e \in E$, we have $(\lambda + \mu).e = \lambda.e + \mu.e$;*
  - *for all $\lambda \in \mathbb{K}$ and $e, f \in E$, we have $\lambda.(e + f) = \lambda.e + \lambda.f$;*
  - *for all $\lambda, \mu \in \mathbb{K}$ and $e \in E$, we have $(\lambda\mu).e = \lambda.(\mu.e)$;*
  - *for all $e \in E$, we have $1.e = e$.*

Remarque: These definitions also make sense in the case where $\mathbb{K}$ is simply a unital ring A, in which case we talk about an A-module, see §2.5.
Examples :
- $\mathbb{K}$ and more generally $\mathbb{K}^n$, $\mathbb{K}^{\mathbb{N}}$, or $\mathbb{K}^{(\mathbb{N})}$;
- $\mathbb{M}_{m,n}(\mathbb{K})$ and $\mathbb{K}[X]$;
- functions from X to $\mathbb{K}$, where X is any set ;
- the arbitrary product of a family of vector spaces is a vector space.

Definition 2. *Let E be a $\mathbb{K}$-vector space ; a subset $F \subset E$ is a* vector subspace *if and only if it is a subgroup stable under the external operation, i.e., if and only if F is non-empty and for all $f_1, f_2 \in F$ and for all $\lambda \in \mathbb{K}$, we have $f_1 + \lambda f_2 \in F$.*

Examples :
- $\mathbb{K}_n[X] \subset \mathbb{K}[X]$, the subset of polynomials of degree $\leq n$;
- the set of convergent sequences of $\mathbb{K}^{\mathbb{N}}$;
- $\mathbb{R} \subset \mathbb{C}$ is a $\mathbb{R}$-vector subspace but is not a $\mathbb{C}$-vector subspace.

Remarque: as previously, a vector subspace is a vector space, and this remark is usually used to test whether one is dealing with a vector space.
Remarque: the arbitrary intersection of a family of vector subspaces is a vector space, which allows the definition of the vector subspace generated by a subset $A \subset E$, denoted $< A >$.

Remarque: if $\mathbb{K}$ is an infinite field, any finite union of vector subspaces is a vector subspace if and only if they are all contained in a single one. In particular, a finite union of distinct hyperplanes is not a vector subspace.
Fundamental example : let $(e_i)_{i\in I}$ be an *arbitrary* family of elements of E, then $< \{e_i : \ i \in I\} >$ is the set of *finite support linear combinations* $\sum_{i\in I} \lambda_i e_i$.

Definition 3. *Let F and G be vector subspaces of a vector space E.* The sum $F+G$ *is the subspace $< F \cup G >$ generated by F and G. We say that F and G are* in direct sum *and we write $F \oplus G$ if $F \cap G = \{0\}$.*

Remarque: it is easily verified that $F + G = \{f + g : \ f \in F \text{ and } g \in G\}$; moreover, F and G are in direct sum if and only if the expression of an element $e \in F + G$ as $f + g$ is unique.

Definition 4. *We say that F and G are* complementary *if $E = F \oplus G$, i.e., if the sum $F + G$ is the whole space and they are in direct sum.*

Remarque: care should be taken not to confuse *complementary* with *supplementary* ; recall that the complement of a vector subspace is never a subspace since it does not contain the zero vector !
Exercise : show that subspaces $E_1, \cdots, E_n$ are in direct sum if and only if for all $i = 1, \cdots, n$

$$E_i \cap \Big(\sum_{1\leq k\neq i\leq n} E_k \Big) = \{0\}.$$

1.2 Dimension theory

Definition 5. *A family $\{(e_i)_{i\in I}\}$ of vectors in a vector space E is called* linearly independent *if for every family $(\lambda_i)_{i\in I} \in \mathbb{K}^I$*

$$\sum_{i\in I} \lambda_i e_i = 0 \Rightarrow \forall i \in I, \ \lambda_i = 0.$$

It is called spanning *if $< \{e_i : \ i \in I\} >= E$, i.e. if every vector in E can be written as a finite support linear combination of the e_i.*

Remarque: the family $(X^i)_{i\in\mathbb{N}} \in \mathbb{K}[X]$ is linearly independent and spanning.
Remarque: the family $(e_i)_{i\in I}$ is called *dependent* if it is not linearly independent, i.e. if there exists a non-zero family $(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}$ such that $\sum_{i\in I} \lambda_i e_i = 0$.

Definition 6. *A family $(e_i)_{i\in I}$ of vectors in E is called* a basis *if it is linearly independent and spanning.*

Theorem 7. *Incomplete Basis Theorem.*
Let $\{f_1, \cdots, f_p\}$ be a linearly independent family of vectors and $\{g_1, \cdots, g_q\}$ a spanning family of E. Then there exists an integer $n \geq p$ and a basis $\{e_1, \cdots, e_n\}$ of E such that $e_i = f_i$ for $1 \leq i \leq p$ and $e_j \in \{g_1, \cdots, g_q\}$ for $p + 1 \leq j \leq n$.

Preuve : Consider a maximal cardinal family of the form

$$(f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r})$$

which is linearly independent. We then show that it is also spanning by verifying that for all $1 \leq j \leq q$, the vector g_j belongs to the vector space generated by this family. If j is one of the i_k for $1 \leq k \leq r$, this is clear ; otherwise, by maximality, the family $(f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r}, g_j)$ is dependent, meaning there exists a non-zero family $(\lambda_i)_{1 \leq i \leq p+r+1}$ such that

$$\lambda_1 f_1 + \cdots + \lambda_p f_p + \lambda_{p+1} g_{i_1} + \cdots + \lambda_{p+r} g_{i_r} + \lambda_{p+r+1} g_j = 0.$$

Since the family $(f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r})$ is linearly independent, necessarily $\lambda_{p+r+1} \neq 0$, and thus $g_j \in \langle f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r} \rangle$. Therefore, $(f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r})$ is both linearly independent and spanning, so it is a basis.
Remarque: thus any space containing a finite spanning family admits a basis. In other words, any finite-dimensional vector space admits bases.

Lemme 8. *In a space with a spanning family of cardinality n, any family of non-zero vectors of cardinality $\geq n+1$ is necessarily dependent.*

Preuve : We proceed by induction on n. For $n = 1$ and v a basis, for f, g two non-zero vectors, there exist non-zero λ, μ such that $f = \lambda v$ and $g = \mu v$, so $\mu f - \lambda g = 0$, and thus (f, g) is dependent.

Assume the result holds up to rank $n-1$, and consider a family of $n+1$ non-zero vectors $(f_1, \cdots, f_{n+1})$ in a vector space admitting $(g_1, \cdots, g_n)$ as a spanning family. For each $i = 1, \cdots, n+1$, write

$$f_i = \lambda_{i,1} g_1 + \cdots + \lambda_{i,n} g_n.$$

Without loss of generality, assume $\lambda_{n+1,n} \neq 0$, and define for each $i = 1, \cdots, n$

$$\tilde{f}_i = \lambda_{n+1,n} f_i - \lambda_{i,n} f_{n+1} \in \langle g_1, \cdots, g_{n-1} \rangle.$$

By the induction hypothesis, the family $(\tilde{f}_1, \cdots, \tilde{f}_n)$ is dependent, meaning there exists a non-zero family $(\mu_1, \cdots, \mu_n)$ such that

$$\mu_1 \tilde{f}_1 + \cdots + \mu_n \tilde{f}_n = 0,$$

which provides the relation

$$\mu_1 \lambda_{n+1,n} f_1 + \cdots + \mu_n \lambda_{n+1,n} f_n - (\sum_{i=1}^{n} \mu_i \lambda_{i,n}) f_{n+1} = 0,$$

proving, as $(\mu_1 \lambda_{n+1,n}, \cdots, \mu_n \lambda_{n+1,n}, -\sum_{i=1}^{n} \mu_i \lambda_{i,n})$ is not zero, that the family $(f_1, \cdots, f_{n+1})$ is dependent.

Corollory 9. *Let E be a vector space with a basis of cardinality n. Then all bases of E have cardinality n.*

Preuve : Let $(e_1, \cdots, e_n)$ and $(f_1, \cdots, f_m)$ be two bases of E. Applying the previous lemma to the spanning family $(e_1, \cdots, e_n)$ (resp. $(f_1, \cdots, f_m)$) and the linearly independent family $(f_1, \cdots, f_m)$ (resp. $(e_1, \cdots, e_n)$), we deduce that $m \leq n$ (resp. $n \leq m$), and thus finally $n = m$.

Definition 10. *The cardinality of a basis (and therefore of any basis) of a vector space E is called its* dimension *; it is finite or infinite.*

Corollory 11. *Every vector subspace F of E has dimension less than or equal to that of E, with equality if and only if $F = E$.*

Preuve : It suffices to note that a basis of F is a linearly independent family of E and to apply the previous corollary.

Definition 12. *An* hyperplane *in a finite-dimensional vector space E is any subspace of dimension $n - 1$.*

Remarque: in infinite dimension, a hyperplane is a subspace such that E/F has dimension 1. The dimension of the quotient space E/F is called the *codimension* of F in E.
Remarque: the dimension of $E \times F$ is the sum of the dimensions of E and F. The vector space of linear maps $\mathcal{L}(E, F)$ from E to F has dimension $\dim E \cdot \dim F$.
Remarque: any linearly independent family has cardinality $\leq n$, with equality if and only if it is a basis.
Remarque: the dimension of $F + G$ is less than or equal to $\dim F + \dim G$, with equality if and only if F and G are in direct sum. More precisely, we have the rank formula.

Theorem 13. *Let F and G be subspaces of E, then*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

Preuve : Let $(e_1, \ldots, e_r)$ be a basis of $F \cap G$, which we complete to a basis $(e_1, \ldots, e_r, f_1, \ldots, f_p)$ (resp. $(e_1, \ldots, e_r, g_1, \ldots, g_q)$) of F (resp. of G). We then easily verify that $(e_1, \ldots, e_r, f_1, \ldots, f_p, g_1, \ldots, g_q)$ is a basis of $F + G$, which gives us the stated formula.

Let us conclude this section with a brief note on infinite dimension.

Proposition 14. *Let E be a $\mathbb{K}$-vector space, and let V, W_1, W_2 be subspaces such that $V \cap W_1 = \{0\}$ and $V + W_2 = E$. Then there exists a complement W of V contained in W_2 and containing W_1.*

Preuve : Consider the set $\mathcal{E}$ of subspaces of E containing W_1 and contained in W_2 ; $\mathcal{E}$ is not empty since $W_1 \in \mathcal{E}$. Moreover, $\mathcal{E}$ is partially ordered by inclusion and is inductive. Recall that this means every totally ordered chain admits an upper bound : here, for such a chain, an upper bound is simply given by the union, which is clearly a subspace.

By Zorn ?s Lemma, $\mathcal{E}$ admits a maximal element, which we denote by W. By definition, we have $W \cap V = \{0\}$ and $W_1 \subset W \subset W_2$. It remains to prove that $V + W = E$; every element $x \in E$ can be written as $x = v + w_2$ with $v \in V$ and $w_2 \in W_2$. If $w_2 \in W$, then we are done ; otherwise, consider the subspace generated by W and w_2, denoted by X. By the maximality of W, we must have $X \notin \mathcal{E}$, so there exists $0 \neq y \in X \cap V$; thus, $y = w + \lambda w_2 \in V$, and hence $y \in W \cap V$, which is a contradiction.
Remarque: The reader will note the essential use of Zorn ?s Lemma, which, let us recall, is equivalent to the Axiom of Choice. Thus, our proof is not constructive.

Corollory 15. *Every subspace V of E admits a complement.*

Corollory 16. *Every non-zero vector space admits a basis.*

Preuve : Consider the set $\mathcal{A}$ of free families in E ; this is clearly a non-empty set, partially ordered by inclusion and inductive. By Zorn ?s Lemma, it possesses a maximal element, which is therefore a maximal free family, and is thus necessarily a generating family, hence a basis.
Remarque: The reader may practice with $\mathbb{K}^{\mathbb{N}}$ by verifying that any basis is necessarily uncountable.

Corollory 17. ***(Incomplete Basis Theorem)***
Let $(e_i)_{i\in I}$ be a generating set of E. Let $J \subset I$ such that $(e_i)_{i\in J}$ is linearly independent. Then there exists $J \subset K \subset I$ such that $(e_i)_{i\in K}$ forms a basis.

Preuve : We consider the set $\mathcal{A}$ of linearly independent families $(e_i)_{i\in A}$ for $A \subset I$. This is a non-empty set, partially ordered by inclusion, and clearly inductive. By Zorn's Lemma, $\mathcal{A}$ has a maximal element K ; as before, $(e_i)_{i\in K}$ is linearly independent and spanning by the maximality of K.
Remarque: Finally, let us mention the case of Hilbert spaces, i.e., Hermitian spaces, in the sense of the section on bilinear algebra, which are complete, meaning all Cauchy sequences are convergent.

Definition 18. *A set $(e_i)_{i\in I}$ is called a* Hilbert basis *of a Hilbert space H if and only if :*

- *it is an orthonormal basis, i.e., $< e_i, e_j >= \delta_{i,j}$;*
- *the set is complete in the sense that for all $x \in H$, there exists a family of scalars $(\lambda_i)_{i\in I}$ such that $\sum_{i\in I} \lambda_i e_i = x$, i.e., the corresponding series in H converges to x.*

Remarque: The reader can easily verify that a basis in the Hilbert sense is not necessarily a basis in the classical sense, as seen for example in L^2 spaces.

1.3 Linear maps.

Definition 19. A linear map *or* a morphism f *from a vector space* E *to a space* F *is a map such that for all* $\lambda \in \mathbb{K}$ *and* $x, y \in E$, *we have* $f(x + \lambda y) = f(x) + \lambda f(y)$.

Remarque: A linear map from E to E is called an endomorphism. In the case where $F = \mathbb{K}$, it is called a *linear form.*
Remarque: Every linear map $f : E \to F$ satisfies $f(0) = 0$ and

$$f(\sum_{i=1}^{n} \lambda_i e_i) = \sum_{i=1}^{n} \lambda_i f(e_i).$$

Notation 1. *We denote by* $\mathcal{L}(E, F)$ *(resp.* $\mathcal{L}(E) = \mathcal{L}(E, E)$*) the set of morphisms from* E *to* F *(resp. the endomorphisms of* E*); it is a vector space of dimension* $\dim E . \dim F$.

Concerning the existence of linear maps, we have the following result.

Proposition 20. *Let* $(e_i)_{1 \leq i \leq n}$ *be a basis of* E. *For any set of* n *vectors* $\{f_1, \cdots, f_n\}$ *in* F, *there exists a unique linear map such that for all* $i = 1, \cdots, n$, *we have* $f(e_i) = f_i$.

Remarque: Thus, two linear maps are equal if and only if they coincide on a basis.

Definition 21. *For* $f \in \mathcal{L}(E, F)$, *we denote by* $\operatorname{Ker} f$ *the set of* $e \in E$ *such that* $f(e) = 0$*; it is a subspace of* E *called* the kernel *of* f.

Remarque: The image of f is also a subspace of F denoted $\operatorname{Im} f$. More generally, the direct or reciprocal image of a subspace is a vector subspace.

Proposition 22. *A linear map* f *is injective if and only if* $\operatorname{Ker} f = \{0\}$.

Remarque: f is surjective if and only if the image of a basis of E is a generating set of F. Thus, f is bijective, and we say that f is *an isomorphism*, if the image of a basis is a basis : this is then true for any basis.
Remarque: A linear map $f : E \to F$ where $\dim E = \dim F$ is injective if and only if it is surjective.

Notation 2. *We denote by* $GL(E)$ *the set of isomorphisms of* E, *also called* automorphisms. *It is a group under composition.*

Remarque: For $E = \mathbb{K}^n$, the vectors $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ define a basis called the *canonical* basis. Every vector space endowed with a basis $(e_i)_{1 \leq i \leq n}$ of cardinality n is isomorphic to $\mathbb{K}^n$ where $f : \mathbb{K}^n \to E$ is defined by $f(x_1, \cdots, x_n) = \sum_{i=1}^n x_i e_i$. In particular, two vector spaces of the same dimension are always isomorphic.

Theorem 23. *Let $f \in \mathcal{L}(E, F)$, then*

$$\dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im}(f).$$

Definition 24. *The dimension of* $\operatorname{Im} f$ *is called* the rank *of f ; it is denoted* $\operatorname{rg} f$.

1.4 Review on duality

Definition 25. *Given a vector space E, the set of linear forms on E is a vector space denoted by E^* and called* the dual *of E.*

Remarque: A basis $(e_i)_{1 \leq i \leq n}$ of E being fixed, the linear map $e_i^* \in E^*$ defined by $e_i^*(e_j) = \delta_{i,j}$ is a basis of E^* called the dual basis of $(e_i)_i$. One should be cautious with this notation since e_i^* depends on the entire basis $(e_i)_i$ and not just on the vector e_i alone.

Proposition 26. *Given a subspace $F \subset E$, the subset $F^\perp \subset E^*$ of linear forms vanishing on F is a subspace of dimension* $\dim E - \dim F$*, i.e., the dimension of $F^\perp$ is equal to the codimension of F.*

Definition 27. *Let $f \in \mathcal{L}(E, F)$, then we associate to it its* adjoint map *denoted by $f^* \in \mathcal{L}(F^*, E^*)$, defined by the formula*

$$y^* \in F_* \mapsto f^*(y^*) = y^* \circ f$$

in the sense that $f^*(y^*)$ is the linear form on E defined by $x \mapsto y^*(f(x))$.

Remarque: Note in particular that f and f^* have the same rank.

Proposition 28. *Let E be a finite-dimensional vector space ; then the bidual $(E^*)^*$ is canonically identified with E.*

Remarque: The map $E \to (E^*)^*$ is given by $x \mapsto (f \mapsto f(x))$.

2 Matrices

2.1 Generalities

Definition 29. A matrix with coefficients in $\mathbb{K}$ of size $m \times n$ *is a table* $(a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ *of scalars $a_{i,j} \in \mathbb{K}$ placed in the i-th row and the j-th column. We denote by $\mathbb{M}_{m,n}(\mathbb{K})$ the set of these matrices, which is endowed with a vector space structure by identifying it with $\mathbb{K}^{nm}$, i.e., coefficient by coefficient.*

Remarque: A row matrix (resp. column matrix) corresponds to the case where $n = 1$ (resp. $m = 1$); we also refer to them as row vectors (resp. column vectors). The rows (resp. columns) of a matrix are called its row vectors (resp. column vectors).
Remarque: The matrices $E_{i,j}$, whose coefficients are all zero except the one at index (i, j) equal to 1, form a basis of $\mathbb{M}_{m,n}(\mathbb{K})$.
Remarque: The matrix $(b_{i,j} = a_{j,i})_{i,j} \in \mathbb{M}_{n,m}(\mathbb{K})$ is called *the transpose matrix*, denoted as $B = {}^tA$ if $A = (a_{i,j})_{i,j}$.
Remarque: In the case where $m = n$, we refer to square matrices, and we denote $\mathbb{M}_n(\mathbb{K})$ for $\mathbb{M}_{n,n}(\mathbb{K})$. The elements $a_{i,i}$ of $A = (a_{i,j})_{i,j}$ are called *diagonal elements*. A matrix is called :

— *diagonal* if all its non-diagonal coefficients are zero; it is also called an *antidiagonal matrix* if $a_{i,j} = 0$ except when $i + j = n + 1$.
— *upper triangular* (resp. lower triangular) if all $a_{i,j}$ are zero for $i > j$ (resp. $j > i$).
— *tridiagonal* if $a_{i,j} = 0$ for all $|j - i| > 1$.

Matrices are not simply arrays of numbers but are introduced because they allow :

— the study of linear systems;
— the manipulation of endomorphisms of vector spaces.

Thus, for $f : E \to F$, an endomorphism between two vector spaces equipped with respective bases $(e_i)_{1\leq i\leq n}$ and $(f_j)_{1\leq j\leq m}$, we associate the matrix $A(f) = (a_{i,j})_{\substack{1\leq i\leq n\\ 1\leq j\leq m}}$ such that for all $1 \leq i \leq n$, we have

$$f(e_i) = \sum_{j=1}^{m} a_{i,j} f_j.$$

In other words, the column vectors of A are the $f(e_i)$ expressed in the basis $(f_j)_j$.
Remarque: As mentioned above, f is determined by its matrix $A(f)$, so we should be able to express the image $f(x)$ of any vector $x = \sum_{i=1}^{n} x_i e_i$.

Definition 30. *For any matrix $A \in \mathbb{M}_{m,n}(\mathbb{K})$ and any column vector $X \in \mathbb{M}_{n,1}(\mathbb{K})$, we define the column vector $Y = AX \in \mathbb{M}_{m,1}(\mathbb{K})$ by the formula :*

$$y_j = \sum_{k=1}^{n} a_{j,k} x_k.$$

For a matrix $B \in \mathbb{M}_{n,r}$, with column vectors denoted by $C_1, \cdots, C_r$, we define the matrix $M = AB \in \mathbb{M}_{m,r}(\mathbb{K})$ whose column vectors are the AC_i for $i = 1, \cdots, r$.

Proposition 31. *Let $f : E \to F$ and $A(f)$ be its matrix relative to bases $(e_i)_{1\leq i\leq n}$ and $(f_j)_{1\leq j\leq m}$ of E and F respectively. For any $x = \sum_{i=1}^{n} x_i e_i$,*

let X be the column vector $(x_{i,1})_{1\leq i\leq n}$. Then the coordinates of $f(x)$ in the basis $(f_j)_{1\leq j\leq m}$ are the coordinates $(y_{j,1})_{1\leq j\leq m}$ of the column vector $A(f)X$, i.e., $f(x)=\sum_{j=1}^{m} y_j f_j$.

Corollory 32. *For any $f : E \to F$ and $g : F \to G$ endomorphisms, where E, F, G are equipped with bases $(e_i)_{1\leq i\leq n}$, $(f_j)_{1\leq j\leq m}$, and $(g_k)_{1\leq k\leq r}$ respectively. Let $A(f), A(g)$, and $A(g\circ f)$ be the matrices associated with f, g, and $g\circ f$ relative to these bases. Then,*

$$A(g\circ f) = A(g)A(f).$$

Proposition 33. *If E, F are equipped with respective bases $(e_i)_i$ and $(f_j)_j$, then the matrix of f^* in the associated dual bases of F^* and E^* is the transpose of the matrix of f in the bases $(e_i)_i$ and $(f_j)_j$.*

In particular, since $\mathcal{L}(E)$ is an algebra, we deduce the following corollary.

Corollory 34. *The matrix multiplication defined above endows $\mathbb{M}_n(\mathbb{K})$ with an algebra structure.*

Definition 35. *Matrices in $\mathbb{M}_n(\mathbb{K})$ that correspond to automorphisms of E are called* invertible *; the set of these invertible matrices is denoted $GL_n(\mathbb{K})$.*

Remarque: A matrix is invertible if and only if its column vectors form a basis.

Definition 36. *Given a vector space E equipped with two bases $(e_i)_{1\leq i\leq n}$ and $(e'_i)_{1\leq i\leq n}$, the matrix that represents the change of basis from $(e_i)_i$ to $(e'_i)_i$, denoted by $P_{e_i\leftarrow e'_i}$, is the matrix in $\mathbb{M}_n(\mathbb{K})$ whose j-th column is given by the coordinates of e'_j in the basis $(e_i)_i$, i.e., $e'_j=\sum_{i=1}^{n} p_{i,j}e_i$.*

Remarque: The matrix $P_{e_i\leftarrow e'_i}$ can also be viewed as the matrix of the identity map $E \to E$, where the starting space is equipped with the basis $(e'_i)_i$ and the target space with the basis $(e_i)_i$. We deduce that :

- $P_{e_i\leftarrow e'_i}$ is invertible, with inverse $P_{e'_i\leftarrow e_i}$;
- if X' is the column vector of the coordinates of a vector e of E in the basis $(e'_i)_i$, then $X = P_{e_i\leftarrow e'_i}X'$ represents e in the basis $(e_i)_i$;
- if $A(f)$ is the matrix of $f : E \to F$ equipped with bases $(e_i)_i$ and $(f_j)_j$ of E and F respectively, then for bases $(e'_i)_i$ and $(f'_j)_j$, the matrix $A'(f)$ relative to these bases is $P^{-1}_{e_i\leftarrow e'_i}A(f)P_{f_j\leftarrow f'_j}$. In the particular case where $E = F$ and $A(f)$ and $A'(f)$ represent the matrix of f in the bases $(e_i)_i$ and $(e'_i)_i$ then $A'(f) = P^{-1}_{e_i\leftarrow e'_i}A(f)P_{e_i\leftarrow e'_i}$.

Examples : Given a matrix $A \in \mathbb{M}_{m,n}(\mathbb{K})$, multiplication on the left (resp. on the right) by the matrix :

- $T_{i,j}(\lambda)$, whose diagonal entries are all equal to 1, with all other entries being zero except $t_{i,j} = \lambda$, corresponds to modifying the rows (resp. columns) of A according to the rule $L_i \to L_i + \lambda L_j$ (resp. $C_i \to C_i + \lambda C_j$) ;

— $D_i(\lambda)$, a diagonal matrix whose diagonal entries are all equal to 1 except $d_{i,i} = \lambda$, corresponds to modifying the rows (resp. columns) of A according to the rule $L_i \to \lambda L_i$ (resp. $C_i \to \lambda C_i$) ;
— $P_{i,j} = I - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$ corresponds to modifying the rows (resp. columns) of A according to the rule $L_i \leftrightarrow L_j$ (resp. $C_i \leftrightarrow C_j$).

Remarque: For $i \neq j$, the matrices $T_{i,j}(\lambda)$ (resp. $D_i(\lambda)$) are called elementary transvection (resp. dilation) matrices relative to the canonical basis. The matrices $P_{i,j}$ are special cases of permutation matrices. These three types of matrices allow for performing elementary row and column operations on a matrix. We will revisit this topic in the study of linear systems.

2.2 Linear Systems.

Definition 37. *A* linear equation *in the variables $x_1, \cdots, x_n$ is an expression of the form*

$$L(x_1, \cdots, x_n) = b \text{ where } L(x_1, \cdots, x_n) = a_1 x_1 + \cdots + a_n x_n,$$

for which we seek solutions in $\mathbb{K}^n$. The equation is called homogeneous *when $b = 0$.*

Remarque: We can and should interpret $L(x_1, \cdots, x_n)$ as a linear form on $\mathbb{K}^n$ written in the canonical basis.

Definition 38. *A* linear system *of m equations with n variables is a collection of m linear equations :*

$$(S) = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m \end{cases}$$

which we aim to solve simultaneously. The system is called inconsistent *if it has no solutions, and* consistent *otherwise.*

Remarque: As suggested by the notation, we introduce the matrix $A_S = (a_{i,j}) \in \mathbb{M}_{m,n}(\mathbb{K})$ and write the system above in matrix form as $A_S X = B$, where X (resp. B) is the column vector of coordinates x_i (resp. b_i).

Definition 39. *Two linear systems (S) and (S') are said to be* equivalent *if they have the same set of solutions.*

Proposition 40. *Two linear systems (S) and (S') are equivalent if and only if there exists an invertible matrix $P \in GL_m(\mathbb{K})$ such that $A_S = PA_{S'}$ and $B = PB'$.*

Remarque: Using the fact that $GL_n(\mathbb{K})$ is generated by transvection and dilation matrices (and for convenience, we also include permutation matrices $P_{i,j}$), we should be able to manipulate system (S) to arrive at an equivalent system (S'). However, this process needs to be constructive, which is ensured by Gaussian elimination.

Definition 41. *Let (S) be a linear system, not necessarily homogeneous, which we write in matrix form as $A_S X = B$. We then introduce the matrix $\tilde{A}_S$ by appending the column B to the matrix A_S.*

Definition 42. *A matrix $M \in \mathbb{M}_{m,n}(\mathbb{K})$ is said to be in* row echelon form *if, below the first non-zero entry of each row, there are only zeros. It is called* reduced row echelon form *if the first non-zero entry of each row, called the* pivot*, is equal to* 1*, and each pivot is the only non-zero element in its column.*

Proposition 43. *For any matrix $M \in \mathbb{M}_{m,n}(\mathbb{K})$, there exists a unique matrix $P \in GL_m(\mathbb{K})$ such that PM is in reduced row echelon form.*

Remarque: The practical implementation of this result is called the **Gaussian elimination algorithm**.

Thus, given a linear system (S) with augmented matrix $\tilde{A}_S$, we apply Gaussian elimination to obtain the associated reduced row echelon form, for example, in the form

$$\begin{pmatrix} 0 & \cdots & \mathbf{1} & \bullet & \cdots & \bullet & \cdots & \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & \mathbf{1} & \bullet & 0 & 0 & \bullet & 0 & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \mathbf{1} & 0 & \bullet & 0 & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \mathbf{1} & \bullet & 0 & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \alpha \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

— If the last column contains a pivot, then the system is inconsistent, which in the example above corresponds to the case $\alpha \neq 0$.

— Otherwise, the positions of the pivots provide the indices of the *leading* variables, while the others are called *free* variables. The set of solutions is then an affine subspace of dimension equal to the number of free variables ; that is, for any values of the free variables, there is a unique solution for the leading variables, which can be found by solving the system from the bottom up.

Definition 44. *The system (S) is called a* Cramer system *if it has a unique solution.*

Remarque: In other words, (S) is a Cramer system if it is consistent and has no free variables, which allows us to prove the following proposition.

Proposition 45. *The system (S) is a Cramer system if and only if A_S is an invertible matrix, which in particular implies that $m = n$.*

Remarque: Linear systems and their resolution via Gaussian elimination are used, for example, to find the inverse of a matrix, or to provide linear equations for a subspace whose generating set is known.

2.3 Decompositions

The Gaussian elimination algorithm provides the following results :

(a) $SL_n(\mathbb{K})$ is generated by elementary transvection matrices. In dimension three or higher, all transvections are conjugate ; this implies the simplicity of $PSL_n(K)$. In dimension 2, for (e_1, e_2) a basis, we denote by τ_λ the transvection defined by $\tau_\lambda(x_1e_1 + x_2e_2) = (x_1e_1 + x_2e_2) + \lambda x_2e_1$. Every transvection is conjugate to a τ_λ, and τ_λ and τ_μ are conjugate if and only if $\frac{\lambda}{\mu} \in (K^\times)^2$.

It follows that $GL_n(\mathbb{K})$ is generated by elementary transvection matrices and dilation matrices $D_n(\lambda) = \mathrm{diag}(1, \cdots, 1, \lambda)$, where $\lambda \in \mathbb{K}^\times$.

(b) **Bruhat decomposition** : $T(n, \mathbb{C})\backslash GL(n, \mathbb{C})/T(n, \mathbb{C}) \simeq \mathfrak{S}_n$, where $T(n, \mathbb{C})$ denotes the set of upper triangular matrices. This result is interpreted in terms of flags : the set of equivalence classes of pairs of complete flags under the action of $GL(n, \mathbb{C})$ is in bijection with $\mathfrak{S}_n$.

(c) **LU decomposition** : When performing Gaussian elimination, we ask if pivoting requires row exchanges (in general, yes, especially to minimize rounding errors). The condition for avoiding row exchanges is that $\det M^{(k)}$ is invertible for all $1 \leq k \leq n$, where $M^{(k)}$ denotes the kth principal minor : the factorization is then unique if the diagonal of L is set to 1. To solve linear systems : we solve $LY = B$, then $UX = Y$ (note that computing L is particularly simple using matrices $T_{i,j}$; see Ciarlet, p.83). We can mention the particularly simple case of tridiagonal matrices (Ciarlet, p.85), which requires $8n - 6$ operations, or more generally band matrices, for which the LU decomposition is still of the same form (notion of static data structure).

(d) Let B be a matrix in $\mathbb{M}_{m,n}(\mathbb{R})$ and $c \in \mathbb{R}^m$. We seek u such that the Euclidean norm of $Bu - c$ is minimized (least squares method). We introduce the functional

$$2J(v) := ||Bv - c||^2 - ||c||^2 = ({}^tBBv, v) - 2({}^tBc, v),$$

which we seek to minimize. Since the symmetric matrix ${}^tBB$ is positive, the function J is convex, so the set of solutions coincides with that of the equation

$$J'(u) = {}^tBBu - {}^tBc = 0.$$

We are thus naturally led to solve equations of the form $AX = B$, where A is symmetric and positive definite, and we apply the **Cholesky decomposition** : $A = B^tB$, where B is lower triangular (the LU decomposition gives $A = LDV$, with, by uniqueness, $V = {}^tL$, and we take the square root of D).

(e) Any matrix is algorithmically similar to a **Hessenberg matrix** : in the QR method, if A is Hessenberg, then all the A_k are too, which shortens computation times.

(f) **QR factorization** : Q is unitary and R is upper triangular. If the diagonal coefficients of R are required to be positive or zero, and if A is invertible, then the factorization $A = QR$ is unique (on $\mathbb{R}$, Q is orthogonal). To construct them, one can use the Gram-Schmidt orthonormalization process, but this method should be avoided as it introduces rounding errors. The use of Householder matrices $H_u = I_n - 2\frac{u^tu}{||u||^2}$ is much more efficient.

(g) Given a symmetric matrix A, there exists a matrix P that is the product of $(n-2)$ Householder matrices, such that ${}^tPAP$ is tridiagonal (Ciarlet, p.120). Applying the Givens method, we obtain approximate eigenvalues : the characteristic polynomials of the principal minors form a Sturm sequence, which allows the roots to be located as precisely as desired, for example, using dichotomy (Ciarlet, p.123).

(h) QR method for eigenvalue localization : Let A be an invertible matrix with distinct eigenvalues, such that there exists a basis transformation matrix allowing for an LU decomposition. We then construct a sequence A_k defined recursively by $A_k = Q_kR_k = R_{k-1}Q_{k-1}$. Then A_k converges to the diagonal matrix of eigenvalues of A, ordered by decreasing magnitude.

2.4 Matrix Calculations in a Principal Ring

Let us recall that $\mathbb{M}_n(A)$ denotes the set of square matrices of size n with coefficients in A. In this section, A is a principal ring (ideally Euclidean, such as $\mathbb{Z}$ or $\mathbb{K}[X]$).

Lemme 46. *A matrix $M \in \mathbb{M}_n(A)$ is invertible if and only if* $\det M \in A^\times$.

Preuve : If the matrix M is invertible, then by applying the determinant to the equality $M.M^{-1} = I_n$, we obtain that the inverse of $\det M$ is $\det(M^{-1})$. Conversely, let $\tilde{M}$ denote the transpose of the cofactor matrix of M. From the equality $\tilde{M}.M = \det M$, we deduce that if $\det M \in A^\times$, then $(\det M)^{-1}\tilde{M}$ is the inverse of M.

The following lemma is the essential computational ingredient for the matrix calculations that follow.

Lemme 47. *Let $x, y \in A$ be non-zero, and let z be a greatest common divisor (gcd) of x and y, with $ux + vy = z$ a BÃ©zout relation. Then the*

matrix $M := \begin{pmatrix} u & v \\ -y/z & x/z \end{pmatrix}$*, with determinant* 1*, satisfies the equation*

$$M \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} z \\ 0 \end{pmatrix}.$$

We will use the following (n, n) matrix :

$$L_{j,k}(x,y) = \begin{pmatrix}
1 & 0 & \dots & & & & & & \dots & 0 \\
0 & \ddots & & & & & & & 0 & \vdots \\
 & & 1 & & & & & & & \\
\vdots & 0 & \dots & u & 0 & \dots & & v & \vdots & \\
 & \vdots & & & 1 & & & & & \\
 & & & & & \ddots & & & & \\
 & & & & & & 1 & & & \\
 & & & -y/z & 0 & \dots & & x/z & & \\
 & & & & & & & & 1 & \\
 & & & & & & & & \ddots & \\
0 & \dots & & & & & & & & 1
\end{pmatrix}$$

with u in (j, j), v in (j, k), $-y/z$ in (k, j), and x/z in (k, k).
Remarque: As usual, denoting by l_i the row with index i of a matrix M, the left multiplication of a matrix $M \in \mathbb{M}_{n,m}(A)$ by the matrix $L_{j,k}(x, y)$ replaces l_j and l_k with $\alpha l_j + \beta l_k$ and $\gamma l_j + \delta l_k$, respectively.

Proposition 48. *Let* $M \in \mathbb{M}_{n,m}(A)$*. Then there exists a matrix* $L \in SL_n(A)$ *such that* LM *is upper triangular.*

Preuve : The proof consists of applying Lemma 47 multiple times to create zeros below the main diagonal.
a) Let $M = (a_{ij})$ $(1 \le i \le n,\ 1 \le j \le m)$. We multiply M on the left by the matrix $L_1 = L_{1,2}(a_{1,1}, a_{2,1})$ so that the first column of $M_1 = L_1 M$ starts with $\begin{pmatrix} d \\ 0 \end{pmatrix}$, where $d = a_{11} \wedge a_{21}$.
b) Next, we multiply M_1 on the left by a matrix $L_2 = L_{1,3}(d, a_{3,1})$ to create a zero at position $(3, 1)$ and replace d with $d_1 = a_{1,1} \wedge a_{2,1} \wedge a_{3,1}$. We continue this process until we obtain the matrix $M_{n-1} = L_{n-1} \cdots L_1 M$, whose first column is $(d_{n-1}, 0, \dots, 0)$, where d_{n-1} is the gcd of the elements in the first column of M.
c) We proceed similarly with the second column, starting by multiplying on the left by a matrix $L_{2,3}$ so that the first row remains unchanged and only the rows $l_2, \dots, l_n$ are manipulated. In this way, we obtain a second column of the form $(a, b, 0, \dots, 0)$. Repeating this process for all the columns, we obtain an upper triangular matrix.

Definition 49. *A matrix $M \in \mathbb{M}_{n,m}(A)$ is called* reduced *if*

$$M = \begin{pmatrix} a_{1,1} & 0 & \dots & & 0 \\ 0 & a_{2,2} & 0 & \dots & \vdots \\ \vdots & & \ddots & & \\ & & & a_{n,n} & \dots \; 0 \end{pmatrix}$$

with

$$a_{i,i} \mid a_{i+1,i+1}, \quad 1 \leq i \leq \inf(n,m) - 1.$$

Remarque: The matrix M is represented with $n < m$. Note that the last $a_{i,i}$ may be zero, and all the off-diagonal elements are zero.

Theorem 50. *Let $M \in \mathbb{M}_{n,m}(A)$. Then there exist $L \in SL_n(A)$ and $R \in SL_m(A)$ such that $M' = LMR$ is reduced.*

Remarque: The analogous statement for a field K is that any matrix $M \in \mathbb{M}_{n,m}(K)$ is equivalent to a matrix of the form $M' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Preuve : We have seen how manipulating rows allows us to bring the gcd of each column to the first row. We start by doing this for each column. Then, having operated on the rows, we proceed to operate on the columns to bring the gcd of the first row, and thus the gcd of all the coefficients of the matrix, to position $(1,1)$.

At this point, the coefficient $a_{1,1}$ is the gcd of all the $a_{i,j}$. We then proceed by manipulating the rows to obtain the first column equal to $(a_{1,1}, 0, \cdots, 0)$: note, in particular, that the first row is not modified. We then proceed similarly with the columns so that the first row is equal to $(a_{1,1}, 0, \cdots, 0)$. As before, we do not modify the first row, so at this point, the matrix is block diagonal with the first block of size 1 and the second of size $(n-1, m-1)$.

The proof is then concluded by induction.

2.5 Generalities on modules

In this section A is any commutative ring.

Definition 51. - *An A-module is a commutative group $(M,+)$ equipped with an application $A \times M \to M$, where ax denotes the image of (a,x), such that :*

1. *$\forall a \in A$ and $x, y \in M$, $a(x+y) = ax + ay$;*
2. *$\forall a, b \in A$ and $x \in M$, $(a+b)x = ax + bx$;*
3. *$\forall a, b \in A$ and $x \in M$, $1x = 1$ and $a(bx) = (ab)x$.*

- A submodule of an A-module M is a subgroup N of M stable by the action of A.

- A morphism of A-modules $M \to N$ is a morphism of additive groups $(M,+) \to (N,+)$ which is moreover A-linear.

Remarque: the notion of A-module is formally identical to that of K-vector space except that the external action is relative to a ring A rather than to a field K.
Examples : the usual constructions on vector spaces are generalized to the case of A-modules (quotient, sum, intersection, A-module generated...). We will use the following examples more specifically in the following.
- If $(G,+)$ is a commutative group, it is canonically equipped with a $\mathbb{Z}$-module structure, by defining, for $n \geq 0$, ng as $\overbrace{g+g+\cdots+g}^{n}$, and $(-1)g$ as $-g$.
- If V is a K-vector space and $u \in \mathcal{L}(V)$ an endomorphism of V, we equip V with a $K[X]$-module structure by setting for all $P \in K[X]$ and for all $\overrightarrow{v} \in V$, $P.\overrightarrow{v} := P(u)(\overrightarrow{v})$.
- The ring A is itself an A-module. The sub-A-modules of A are its ideals.

Definition 52. *A subset S of an A-module M is said*
- free *if the equality $\sum_{s\in S} a_s s = 0$ where the family $(a_s)_{s\in S}$ is assumed to have finite support, implies that for all $s \in S$, we have $a_s = 0$.*
- generator *if every element $m \in M$ can be written in the form $\sum_{s\in S} a_s s$ where the family $(a_s)_{s\in S}$ is finitely supported.*
- a basis *if S is both free and generating.*

We say that M is
- free *if M has a basis.*
- finitely generated *if it has a finite subset S that is generating.*
- torsion *if the set of elements $\lambda \in A$ that cancel M i.e. such that $\forall m \in M$ we have $\lambda m = 0$, is a non-zero ideal of A. This ideal is called* cancellator *of M and denoted* $\mathrm{Ann}(M)$.

Example : the annihilator of $M = A/I$ is the ideal I.
Remarque: the data of a basis of a free A-module of finite type is equivalent to the data of an isomorphism $A^n \longrightarrow M$.

Proposition 53. *Let M be a free A-module of finite type. Then all its bases have the same cardinality.*

Preuve : Let $(e_1, \cdots, e_n)$ be a basis of M and $\mathcal{M} \in A$ be a maximal ideal so that the quotient $A/\mathcal{M}$ is a field k. We denote $\mathcal{M}M = \{\sum_{i=1}^n m_i e_i : \ m_i \in \mathcal{M}\}$ so that $V := M/\mathcal{M}M$ is a k-vector space, i.e. a group with an external action of $A/\mathcal{M}$. Note also that $(\overline{e}_i)_{i=1,\cdots,n}$ is a basis of V. It is clearly a generating family. For freedom, $\sum_{i=1}^n \lambda_i \overline{e}_i = 0$ is also written $\sum_{i=1}^n \mu_i e_i = \sum_{i=1}^n m_i e_i$ where $\overline{\mu}_i = \lambda_i$ and the $m_i \in I$. The family $(e_i)_{i=1,\cdots,e_n}$ being free, we deduce that $\mu_i = m_i$ and therefore $\lambda_i = \overline{\mu}_i = 0$.

Thus n is the dimension of the vector space $M/\mathcal{M}M$ and is therefore the cardinal of any basis of the A-module M.

Proposition 54. *A submodule of a finitely generated module is finitely generated.*

Preuve : Let $f : A^n \twoheadrightarrow M$ be defined by the data of a generating family of cardinality n of M. For a submodule N of M, it suffices to show that $f^{-1}(N)$ is finitely generated, i.e. we are reduced to the case where $M = A^n$.

We then reason by induction on n : in the case $n = 1$, a submodule of A is an ideal which is therefore principal and therefore free of rank 1.
Let us then assume the result acquired up to rank $n-1$ and treat the case of n. Consider then the application $g : N \hookrightarrow A^n \twoheadrightarrow A$ where the second arrow is given by the first projection $(a_1, \cdots, a_n) \mapsto a_1$. The image of g is of the form $a_1 A$ and let $n_1 \in N$ be an antecedent then $N' = N \cap A^{n-1}$ where A^{n-1} is the kernel of the first projection. Note that $n_1 A \cap N' = (0)$ since if $g(\lambda n_1) = 0$ then $\lambda = 0$. Furthermore for $n \in N$, we can write $n = \lambda n_1 + (n - \lambda n_1)$ where $f(n) = \lambda a_1 = f(\lambda n_1)$ and therefore $n - \lambda n_1 \in N'$. In other words, we have $N = An_1 \oplus N'$ with $N' \subset A^{n-1}$. By induction N' is finitely generated and therefore N too.

2.6 Adapted Basis Theorem

Theorem 55. *Let N be a submodule of a free A-module L of finite type. Then N is a free submodule of finite type and there exists a basis, called* adapted, *$(f_1, \ldots, f_n)$ of L as well as elements $a_i \in A$, $1 \leq i \leq n$ such that :*

$$\begin{cases} a_1 \mid a_2 \mid \cdots \mid a_n, \\ \text{the } (a_i f_i) \text{ such that } a_i \neq 0 \text{ form a basis of } N. \end{cases}$$

Moreover, the sequence of ideals (a_i) satisfying these conditions is unique.

Preuve : According to the proposition 54, N is of finite type, let us then denote $(g_1, \ldots, g_m)$ a generating family of N and let us write the transition matrix M of the g_i for $1 \leq i \leq m$ in a basis $(e_1, \cdots, e_n)$ of L. According to 50, there exist $P \in SL_n A$ and $Q \in SL_m(A)$ such that $M' = PMQ$ is reduced with elements $a_{i,i}$ on the diagonal that we simply denote by a_i.
The matrix P (resp. Q) is interpreted as a matrix of change of basis of L (resp. of change of generating family of N). Let $(f_1, \cdots, f_n)$ be the new basis of L ; the matrix writing of M' is then interpreted by saying that $(a_1 f_1, \cdots, a_r f_r)$ is a generating family of N, where we denote by a_r the last of the non-zero a_i. We then note that this new generating family of N is free, i.e. N is also free with $M/N \simeq A/(a_1) \times \cdots \times A/(a_r) \times A^{n-r}$. Let us then show the uniqueness of (a_i). First, note that $n - r$ depends only on N. To do this, let us consider an irreducible p not dividing a_r so that for all $1 \leq i \leq n$, a_i is invertible modulo p and therefore for $M_i = A/(a_i)$ we have M_i/pM_i is zero. We thus see that $n - r$ is the dimension of the $A/(p)$ vector space $(M/N)/p(M/N)$. Let us then consider

$$\frac{A}{(a_1)} \times \cdots \times \frac{A}{(a_q)} \simeq \frac{A}{(a'_1)} \times \cdots \times \frac{A}{(a'_s)},$$

the (a_i) and the (a'_i) verifying the divisibility properties of the statement and are all non-zero. Then $(a_r) = \mathrm{Ann}(M/N) = (a'_s)$ and therefore $\frac{A}{(a_1)} \times \cdots \times \frac{A}{(a_{q-1})} \simeq \frac{A}{(a'_1)} \times \cdots \times \frac{A}{(a'_{s-1})}$. By proceeding in the same way, we identify step by step the a_{q-i} with the a'_{s-i} until obtaining $s = q$.

Definition 56. *We say that $m \in M$ is a* torsion element *if $m \neq 0$ and if there exists $\lambda \in A, \lambda \neq 0$, such that $\lambda m = 0$. The set of torsion elements of M is denoted M_t. If $M_t = \{0\}$, we say that M is* torsion-free.

Remarque: M_t is a submodule of M. Moreover, M is torsion if and only if $M = M_t$. With these notions the proof of the previous theorem is rewritten.

Theorem 57. *Let M be a finitely generated A-module, M_t its torsion submodule. Then there exists a free submodule $L \subset M$ of rank r such that $M = M_t \oplus L$ as well as elements $a_1|a_2|\cdots|a_q$ of A such that*

$$M_t \simeq A/(a_1) \oplus A/(a_2) \oplus \cdots \oplus A/(a_q).$$

Definition 58. *The rank of the free part of M is called the* rank *of M. The non-zero ideals (a_i) for $1 \leq i \leq q$ are the* invariant factors *of M.*

Remarque: the field of rationals $\mathbb{Q}$ is an example of a torsion-free and non-free $\mathbb{Z}$-module (if $q_1 = a/b$ and $q_2 = c/d$, are two non-zero rationals, we have the relation $bcq_1 - adq_2 = 0$). Thus $\mathbb{Q}$ is not a $\mathbb{Z}$-module of finite type.

Proposition 59. *Let M be a finitely generated A-module. The following conditions are equivalent :*

1. *the module M is indecomposable ;*
2. *$M \simeq A$, or there exists an irreducible element $p \in A$, an integer $\alpha > 0$ such that $M \simeq A/(p^\alpha)$.*

Preuve : 1. $\Rightarrow$ 2.
According to the theorem 57 we can assume $M = A/(a)$. If the element a has at least two irreducible factors, it follows from the Chinese lemma that M is not indecomposable.
2. $\Rightarrow$ 1.
Since the ring A is integral, it is clear that the A-module A is indecomposable.
If $\alpha > 0$, the submodules of $\tilde{M} = A/(p^\alpha)$ are generated by the images in $\tilde{M}$ of the elements p^γ for $\gamma \leq \alpha$. If M_1 and M_2 are two such submodules, we always have $M_1 \subset M_2$ or $M_2 \subset M_1$; they cannot therefore be in direct sum.

Definition 60. *Let M be an A-module, $p \in \mathcal{P}$ an irreducible element. We denote $M(p)$ the set of elements $x \in M$ of p-torsion, i.e. vanished by a power of p.*

Remarque: by applying the Chinese theorem to the 57 theorem, we obtain the canonical decomposition into indecomposable given by the following theorem.

Theorem 61. *Let M be a finitely generated A-torsion module, $(a) = \mathrm{Ann}(M)$ its annihilator. Then :*

1. *$M = \bigoplus_{p_i \in \mathcal{P},\, p_i | a} M(p_i)$ and $M(p_i) \neq (0)$ for each irreducible element p_i such that $p_i | a$*
2. *for each irreducible element $p_i \in \mathcal{P}$, $p_i | a$, there exists a sequence of integers $\nu_{i1} \leq \nu_{i2} \leq \cdots \leq \nu_{ik}$ unique such that :*

$$M(p_i) \simeq \prod_{j=1}^{k} A/(p_i^{\nu_{ij}})).$$

3. *the decomposition $M \simeq \prod_{i,j} A/(p_i^{\nu_{ij}})$ is the unique decomposition of M into a product of indecomposable modules.*

Example : let us take $A = \mathbb{Z}$, $M = M_t = \mathbb{Z}/96\mathbb{Z} \times \mathbb{Z}/72\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z}$. We have $96 = 2^5 \times 3$, $72 = 2^3 \times 3^2$, hence :

$$M \simeq \mathbb{Z}/32\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$$

by the Chinese theorem. We therefore have

$$\begin{array}{lcccccc} M(2) & \simeq & \mathbb{Z}/32\mathbb{Z} & \times & \mathbb{Z}/8\mathbb{Z} & \times & \mathbb{Z}/2\mathbb{Z} \\ M(3) & \simeq & \mathbb{Z}/9\mathbb{Z} & \times & \mathbb{Z}/3\mathbb{Z} & & \\ M(5) & \simeq & \mathbb{Z}/5. & & & & \end{array} \tag{1}$$

To find the decomposition into indecomposable (resp. into invariant factors), we read the table above in rows (resp. in columns), and we find $M = M(2) \oplus M(3) \oplus M(5)$ (resp. $a_3 = 32 \times 9 \times 5 = 1440$, $a_2 = 8 \times 3 = 24$, $a_1 = 2$, hence the decomposition $M \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/1440\mathbb{Z}$.
Remarque: by considering an abelian group as a $\mathbb{Z}$-module, we can give the classification of finite abelian groups of given order n. We proceed as follows next. Let G be a group of order n.

1. We write $n = p_1^{\nu_1} \ldots p_s^{\nu_s}$ with p_i primes, ν_i integers ;
2. we then have $G \simeq G(p_1) \oplus \cdots \oplus G(p_s)$;
3. for each integer i, $1 \leq i \leq s$ there exists a unique sequence (ν_{ij}) of integers > 0 such that $\sum_j \nu_{ij} = \nu_i$ and $G(p_i) \simeq \bigoplus_j \mathbb{Z}/p_i^{\nu_{ij}}\mathbb{Z}$;
4. two groups of order n are isomorphic if and only if all p_i and integers ν_{ij} are the same.

Example : let us give up to isomorphism all the abelian groups of order $108 = 2^2 \times 3^3$. Let $G = G(2) \oplus G(3)$ with $G(2)$ of order 4 and $G(3)$ of order 27. Up to isomorphism there are two possibilities for $G(2)$, $\mathbb{Z}/4\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, and three for $G(3)$, $\mathbb{Z}/27\mathbb{Z}$, $\mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ and $(\mathbb{Z}/3\mathbb{Z})^3$. There are therefore up to isomorphism six abelian groups of order 108.

2.7 Exercices

Exercice 1. *(Fresnel p.127) Let A be a matrix of $\mathbb{M}_n(K)$. We introduce the matrix N of $\mathbb{M}_{n^2,n+1}(K)$ where for $1 \leq k \leq n+1$, its k-th column vector is made up of the elements of the matrix A^{k-1} taken in an order prescribed once and for all. Show that there exists a matrix P of $SL_{n^2}(K)$ such that PN is of the form $\begin{pmatrix} M_1 & M_2 \\ 0 & M_3 \end{pmatrix}$, where M_1 is an "upper triangular" element of $\mathbb{M}_{d,d+1}(K)$ whose "diagonal" terms are all non-zero and give a way to compute the minimal polynomial*

Preuve : The existence of S follows from the Gauss pivot by row operations. Thus the d+1-first columns of $N' = SN$ are linked either ${}^t(u_0, \cdots, u_{d-1}, -1, 0, \cdots, 0)$ is a kernel vector of N' and therefore also of N because S is invertible. By noting C_i the columns of N, we therefore have $C_d = u_0C_0 + \cdots + u_{d-1}C_{d-1}$ and therefore $A^d - u_{d-1}A^{d-1} - \cdots - u_0\mathrm{Id} = 0$. Suppose that there exists a annihilating polynomial of degree less than or equal to $d-1$, so that the first $d-1$ columns of N are linked and therefore also those of N' which is not, so that $X^d - u_{d-1}X^{d-1} - \cdots - u_0$ is indeed the minimal polynomial of A.

Exercice 2. *Let M be a matrix of $\mathbb{M}_n(\mathbb{R})$; the letters L, L' (resp. U, U') denote lower (resp. upper) triangular matrices with diagonal coefficients equal to 1. The letter D denotes a diagonal matrix.*

(1) Show that we can put M in the form LU (resp. LDU) if and only if $\det M^{(k)} = 1$ (resp. $\det M^{(k)} \neq 0$) for all $1 \leq k \leq n$ where $M^{(k)}$ denotes the principal minor of order k.

(2) Write the matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ in the form $U'LU$. Find a matrix of $\mathbb{M}_2(\mathbb{R})$ that cannot be written in the form LDU where D denotes a diagonal matrix.

(3) Let M_k be the matrix obtained from M by substituting its last row for its row of index k. Let $\det M_k^{(k)} \neq 0$ for all $1 \leq k \leq n$. Show that we can put, in a unique way, M in the form $U'LU$ where U, U' have zero coefficients outside its diagonal and the last column.

(4) Let $M \in GL_n(\mathbb{R})$, show that there exists U' with $n-1$ zero coefficients on its last column such that the first k rows of $(U'M)^{(k+1)}$ are independent for all $k < n$.

(5) Show that any matrix M with determinant n can be put in the form $L'U'LU$ where L' has its zero coefficients outside its diagonal and its last row.

Preuve : (1) The proof is identical in the respective case, we treat the LU decomposition. Let us suppose by induction that $M^{(n-1)} = LU$. We have

$$M = \begin{pmatrix} LU & A \\ B & m \end{pmatrix} = \begin{pmatrix} L & 0 \\ BU^{-1} & 1 \end{pmatrix} \begin{pmatrix} U & L^{-1}A \\ 0 & m' \end{pmatrix}$$

where $m' = m - BU^{-1}L^{-1}A$. By passing to the determinant we have $m' = 1$ hence the result. For uniqueness if $LU = L'U'$ then $L'^{-1}L = U'U^{-1}$ which is necessarily equal to the identity.

(2) The decomposition :

$$\begin{pmatrix} \sin\theta & 1 \end{pmatrix} \begin{pmatrix} 1 & -\tan(\theta/2) \\ 0 & 1 \end{pmatrix}$$

is used to implement rotation of digital images. The matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ cannot be written in the form LDU.

(3) We denote by $(c_1, \cdots, c_{n-1}, 1)$ the coefficients of the last column of U'^{-1}. By multilinearity of the determinant we have

$$\det(U'^{-1}M)^{(k)} = \det M^{(k)} + \sum_{j=1}^{k} c_j M_j^{(k)}$$

The equations $\det(U'^{-1}M)^{(k)} = 1$ for $1 \leq k \leq n-1$ form an invertible triangular system because $\det M_k^{(k)} \neq 0$, so it has a unique solution and $(U'^{-1}M)^{(k)} = LU$ according to (1).

(4) We prove this property for all $k \leq r$ by induction on $r < n$ using that the matrices U' form a group. We denote $U_{k,k}$ as the identity matrix and $U_{k,l}$ $(k < l < n)$ as the matrix that by left multiplication adds row l to row k. We extract from M the upper left matrices $M^{(i,j)} \in \mathbb{M}_{i,j}(\mathbb{R})$.

For $r = 1$, $\mathrm{rg}M^{(n,2)}) = 2$ guarantees $\mathrm{rg}M^{(n-1,2)} \geq 1$. If the first row of $M^{(n-1,2)}$ is zero we denote $l > 1$ as the index of a non-zero row, otherwise we set $l = 1$. The matrix $U' = U_{1,l}$ is suitable, i.e. the first row of $(U_{1,l}M)^{(2)}$ is not zero.

Assume that M satisfies the property for all $k \leq r-1$ so that the first $r-1$ rows of $M^{(n-1,r+1)}$ are independent. Moreover, $\mathrm{rg}M^{(n-1,r+1)} \geq \mathrm{rg}M^{(n,r+1)} - 1 = r$. Choose $U' = U_{r,l}$ where $l = r$ if the r-th row of $M^{(n-1,r+1)}$ is independent of the previous ones and $l > r$ is the index of an independent row otherwise. The first r rows of $(U'M)^{(k+1)}$ are then independent for all $k \leq r$.

(5) Let U_0 be the matrix obtained by applying the previous question to M. We note that $L_0 = L'^{-1}$ is of the same form as L' and therefore that U_0 and L_0 commute. For any vector v independent of the $(n-1)$-first rows of M, we can choose L_0 such that the left multiplication of L_0 on M changes the last row of M into v. By construction of U_0, it is possible to choose the coefficients of v successively so that L_0U_0M satisfies the hypotheses of question (3). By denoting U_1 the matrix U' obtained by applying question (3) to L_0U_0M we obtain $U_0L_0M = L_0U_0M = U_1LU$. We conclude by setting $U' = U_0^{-1}U_1$.

Exercice 3. *Let $x = (n_1, \ldots, n_p) \in \mathbb{Z}^p$.*
a) Show that the following conditions are equivalent :

1. *GCD $(n_1, \ldots, n_p) = 1$*
2. *There exists $A \in SL_p(\mathbb{Z})$ such that $A^t x =^t (1, 0, \ldots, 0)$*
3. *The vector x is part of a basis of $\mathbb{Z}^p$.*

b) Let $p = 4$ and $x = (10, 6, 7, 11)$. Complete x in a basis of $\mathbb{Z}^4$.

Preuve : (i) implies (ii) : the result is proven by induction ; the first step of the calculation is as follows :

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \cdots & u & v \\ 0 & \cdots & -n_r/(n_r, n_{r-1}) & n_{r-1}/(n_r, n_{r-1} \end{pmatrix} \begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = \begin{pmatrix} n_1 \\ \vdots \\ (n_r, n_{r-1}) \\ 0 \end{pmatrix}$$

where u and v are the Bezout coefficients between n_{r-1} and n_r : $un_{r-1} + vn_r = (n_r, n_{r-1})$.

Consider $f : \mathbb{Z}^r \longrightarrow \mathbb{Z}$ defined by $f(a_1, \cdots, a_r) = a_1 n_1 + \cdots + a_r n_r$. The adapted basis theorem assures us the existence of a basis $(e_1, \cdots, e_n)$ of $\mathbb{Z}^n$ such that $\operatorname{Ker} \phi = \mathbb{Z}a_1 e_1 \oplus \cdots \oplus \mathbb{Z}a_r e_r$ with $a_i | a_{i+1}$ in $\mathbb{Z}$ that we call the invariant factors of $\mathbb{Z}^n / \operatorname{Ker} \phi$; we have moreover that the a_i are all equal to 1 for $1 \leq i < r$ and $a_r = 0$ (by a dimension argument). Thus if we denote A transposed from the passage matrix of the basis $(e_r, e_{r-1}, \cdots, e_1)$ in the canonical basis, we have $A \in SL_r(\mathbb{Z})$ and $A \begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

(ii) implies (iii) : it is obvious that the family $\begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $A^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdots, A^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ forms a base of $\mathbb{Z}^n$.

(iii) implies (i) : let $e_2, \cdots, e_r$ be a family that completes $\begin{pmatrix} n_1 \\ \vdots \\ n_r \end{pmatrix}$ in a basis and we denote by A the matrix of passage of this basis in the canonical basis ; we calculate the determinant of A by developing it with respect to

the first column so that it is divisible by the gcd of the n_i which is therefore equal to 1 because $\det A = \pm 1$.

Examples : the first case is simple because we have the relation $7-6=1$ so that the following matrix is of determinant -11

$$\begin{pmatrix} 10 & 6 & 7 & 11 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

so that the 4 column vectors of the transpose of the matrix above constitute a basis of $\mathbb{Z}^4$.
In the second example we have the Bezout relation : $1 = 6.6 - 2.10 - 15$. We therefore look for 6 coefficients a, b, c, d, e, f such that $cf - de = 6$, $af - be = 2$ and $ad - bc = -1$. For example the following matrix is suitable

$$\begin{pmatrix} 6 & 1 & 0 \\ 10 & 0 & -1 \\ 15 & 6 & 2 \end{pmatrix}$$

Exercice 4. *Let n be a positive integer and $G \subset \mathbb{Z}^n$ be a subgroup of rank n. Let $(g_1, \cdots, g_n)$ be a basis of G ; we denote by M the passage matrix of this basis in the canonical basis of $\mathbb{Z}^n$.*

(i) Show that the group $\mathbb{Z}^n/G$ is finite.

(ii) Show that card$((\mathbb{Z}^n/G)) = |\det M|$.

(iii) Let H be an abelian group generated by three elements h_1, h_2, h_3 subject to the relations

$$3h_1 + h_2 + h_3 = 0$$

$$25h_1 + 8h_2 + 10h_3 = 0$$

$$46h_1 + 20h_2 + 11h_3 = 0$$

Show that card$(H) = 19$ and then that $H \simeq \mathbb{Z}/19\mathbb{Z}$. What generalization does this suggest ?

(iv) Triangularize the matrix

$$\begin{pmatrix} 3 & 1 & 1 \\ 25 & 8 & 10 \\ 46 & 20 & 11 \end{pmatrix}$$

by multiplying on the right by a matrix of $SL_3(\mathbb{Z})$ that we will specify.

(v) Deduce an isomorphism $\varphi : H \simeq \mathbb{Z}/19\mathbb{Z}$ and specify the values of $\varphi(h_1), \varphi(h_2), \varphi(h_3)$.

Preuve :

(i) The adapted basis theorem provides a basis $(f_1, \cdots, f_n)$ of $\mathbb{Z}^n$ as well as integers $1 < a_1 | \cdots | a_n \neq 0$ such that $(a_1 f_1, \cdots, a_n f_n)$ is a basis of G. We then obtain $\mathbb{Z}^n/G \simeq \mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_n\mathbb{Z}$.

(ii) According to the above, we therefore have $\mathrm{card}(\mathbb{Z}^n/G) = \prod_{i=1}^n a_i$ which is therefore equal to $\det M$.

(iii) Let $M = \begin{pmatrix} 3 & 1 & 1 \\ 25 & 8 & 10 \\ 46 & 20 & 11 \end{pmatrix}$ so that $M \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ There then exist matrices $L, R \in GL_3(\mathbb{Z})$ such that $M = L\mathrm{diag}(a_1, a_2, a_3)R$ with $a_1|a_2|a_3$. Furthermore if we put $\begin{pmatrix} h'_1 \\ h'_2 \\ h'_3 \end{pmatrix} := R \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}$, H is also generated by h'_1, h'_2, h'_3 and the equation $L\mathrm{diag}(a_1, a_2, a_3) \begin{pmatrix} h'_1 \\ h'_2 \\ h'_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is equivalent to

$$\begin{cases} a_1 h'_1 = 0 \\ a_2 h'_2 = 0 \\ a_3 h'_3 = 0 \end{cases}$$

and therefore $H \simeq \mathbb{Z}/a_1\mathbb{Z} \times \mathbb{Z}/a_2\mathbb{Z} \times \mathbb{Z}/a_3\mathbb{Z}$, with $a_1.a_2.a_3 = \det M$. The statement suggests us to simply calculate $\det M$; we easily verify that it is equal to -19 (cf. (iv) below) as announced. We then obtain $a_1 = a_2 = 1$ and $a_3 = 19$.

In general, if the prime factorization of $\det M$ does not reveal any multiplicity (i.e. $p^2 \not| \det M$ for any prime p), then all the a_i are equal to 1 except the last one which is equal to $\det M$ and the quotient group is then cyclic.

(iv) We calculate

$$\begin{array}{ccc} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & -1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & -1 \end{pmatrix} \\ \begin{pmatrix} 3 & 1 & 1 \\ 25 & 8 & 10 \\ 46 & 20 & 11 \end{pmatrix} & \begin{pmatrix} 3 & 1 & 0 \\ 25 & 8 & 2 \\ 46 & 20 & -9 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 8 & -1 & 2 \\ 20 & 14 & -9 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 8 & 1 & 0 \\ 20 & -14 & -19 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & -1 & 0 \\ 1 & 3 & -1 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 2 \\ 1 & -3 & -5 \\ 0 & 0 & -1 \end{pmatrix} \end{array}$$

(v) We have $\begin{pmatrix} h'_1 \\ h'_2 \\ h'_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & -3 & -5 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}$ which is easily inverted (the matrix is "triangular") i.e. $h_3 = h'_3$, $h_2 + 2h_3 = h'_1$ i.e. $h_2 = h'_1 - 2h'_3$ and $h_1 - 3h_2 - 5h_3 = h'_2$ i.e. $h_1 = 3h'_1 + h'_2 - h'_3$. Since $\varphi(h'_1) = \varphi(h'_2) = 0$ and $\varphi(h'_3) = 1$, we obtain $\varphi(h_1) = -1$, $\varphi(h_2) = -2$ and $\varphi(h_3) = 1$.

3 Reduction of endomorphisms

As we saw above, to each endomorphism f, we associate matrices that depend on the choice of bases. We then seek to find bases so that the matrix is as simple as possible.
Remarque: we will use the language of the following section with orthogonal (resp. unitary), symmetric (resp. Hermitian) endomorphisms in the real (resp. complex) framework. We hope that the loss of logic of presentation of objects will be compensated by the practical aspect of finding in a single section, a set of results on reduction.

3.1 Equivalent matrices

Definition 62. *Two matrices $A, A' \in \mathbb{M}_{m,n}(\mathbb{K})$ (resp. of $\mathbb{M}_n(\mathbb{K})$) are said to be* equivalent *(resp.* similar*) if there exist two matrices $P \in GL_m(\mathbb{K})$ and $Q \in GL_n(\mathbb{K})$ (resp. $P \in GL_n(\mathbb{K})$) such that $A' = PAQ$ (resp. $A' = P^{-1}AP$).*

Remarque: A and A' are equivalent (resp. similar) if they represent the same morphism $f : E \to F$ (resp. $f : E \to E$) relative to different bases at the start and at the finish (resp. different bases but the same at the start and at the finish).

Proposition 63. *Any matrix $A \in \mathbb{M}_{m,n}(\mathbb{K})$ is equivalent to a matrix of the form*

$$\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & 0 & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 \end{pmatrix},$$

where the number of 1 *is equal to the rank of f.*

Remarque: thus the equivalence classes in $\mathcal{L}(E, F)$ are given by the rank.

Let us now consider the case $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$) and where the spaces considered are equipped with a scalar product (resp. Hermitian). We are then allowed to consider only orthonormal bases and therefore orthogonal (resp. unitary) change of base matrices.

Proposition 64. ***(Polar decomposition)*** *Let $A \in \mathbb{M}_m(\mathbb{R})$ (resp. $\mathbb{M}_m(\mathbb{C})$), we can then write A in the form $A = PU$ where $P \in \mathbb{M}_m(\mathbb{R})$ (resp. $P \in \mathbb{M}_m(\mathbb{C})$) is positive semi-definite of the same rank as A and $U \in \mathbb{M}_m(\mathbb{R})$ (resp. $\mathbb{M}_m(\mathbb{C})$) whose column vectors form an orthonormal family, i.e. $U^tU = I_m$ (resp. $UU^* = I_m$). The matrix P is uniquely determined as the unique*

positive square root of A^tA (resp. of AA^*) while U is uniquely determined if A is of rank m.

Remarque: we must see this decomposition as the generalization in the case $n = 1$ of the writing of a complex number in the form $\rho e^{i\theta}$.

Preuve : We treat the case of $\mathbb{R}$, the case of $\mathbb{C}$ being treated in a similar way. The matrix A^tA is positive symmetric and therefore diagonalizable in an orthonormal basis, i.e. there exists O orthogonal such that $OA^tA({}^tO) = \mathrm{diag}(\lambda_1, \cdots, \lambda_m)$ with the λ_i positive or zero. We then note

$$P = {}^tO\mathrm{diag}(\sqrt{\lambda}_1, \cdots, \sqrt{\lambda}_m)O$$

which is therefore positive semi-definite. Suppose that P is invertible, i.e. that A is of rank m, then $U = P^{-1}A$ verifies

$$U{}^tU = P^{-1}(A^tA){}^tP^{-1} = P^{-1}(P^2)P^{-1} = I_m.$$

If A is not of rank m, let $(A_i)_{i\in\mathbb{N}}$ be a sequence of matrices of $\mathbb{M}_m(\mathbb{R})$ of rank m converging to A. From the above, we can write uniquely $A_i = P_iU_i$: as $(U_i)_{i\in\mathbb{N}}$ belongs to the compact $O_m(\mathbb{R})$ and therefore admits a value of adherence $U \in O_n(\mathbb{R})$ so that the corresponding extracted sequence of P_i converges to $P := AU^{-1}$ necessarily symmetric positive semi-definite since the set of symmetric positive semi-definite matrices is clearly closed.

Remarque: by reasoning with extracted sequences as in the proof above, it is easy to demonstrate that the polar decomposition in the invertible case is a homeomorphism.

Remarque: in the previous proof we construct quite naturally the positive Hermitian matrix P while U is only given artificially by the formula $P^{-1}A$. To correct this injustice let us show the following corollary.

Corollory 65. *The application $M \mapsto \sqrt{\frac{1}{n}\mathrm{tr}(MM^*)}$ defines a norm $||-||$ on $\mathbb{M}_n(\mathbb{C})$ such that $||U|| = 1$ for any unitary matrix. For $A \in GL_n(\mathbb{C})$ with polar decomposition $A = SU$, the unitary matrix is uniquely determined by the following property :*

$$||A - U|| = \min_{U'\in\mathbb{U}_n(\mathbb{C})} ||A - U'||,$$

where $\mathbb{U}_n(\mathbb{C})$ denotes the set of unitary matrices of $\mathbb{M}_n(\mathbb{C})$.

Remarque: another way to state the result is to say that U is the orthogonal projection of A onto the unit ball of matrices M such that $||M|| \leq 1$ which is u

Corollory 66. ***(singular values)***
Let $A \in \mathbb{M}_m(\mathbb{C})$ be of rank k, we can then write it in the form

$$A = VDW^*,$$

where $V \in \mathbb{M}_m(\mathbb{C})$ and $W \in \mathbb{M}_m(\mathbb{C})$ are unitary and $D = \mathrm{diag}(d_1, \cdots, d_m)$ with $d_1 \geq d_2 \geq \cdots \geq d_k > d_k = \cdots = d_m = 0$. The numbers $d_{i,i}$ are the positive square roots of the eigenvalues of AA^* and are therefore uniquely determined : they are called the singular values of A.

Remarque: V and W are not uniquely determined, we can just say that the columns of V (resp. W) are eigenvectors of AA^* (resp. A^*A).

Preuve : We start from the polar decomposition $A = PU$ and we diagonalize $P = VDV^*$ in orthonormal basis where D is as in the statement. We then set $W = U^*V$ so that $A = VDW^*$.

Remarque: singular values appear in the conditioning of a matrix in numerical analysis. Let us briefly recall what this is about. Let A be an invertible matrix and B a column matrix, we seek to solve the equation $AX = B$ with unknown X. From a practical point of view, B may undergo a small perturbation δ_B due for example to rounding and we seek to control the perturbation δ_X induced on X, i.e. $A(X + \delta_X) = B + \delta_B$, or $A\delta_X = \delta_B$. We choose a norm subordinate for example to the classical Euclidean norm $|| - ||_2$ so that $|||A|||_2$ is equal to the largest eigenvalue of A^*A, i.e. to the largest singular value. We then have, using

$$||A\delta_X||_2 \leq |||A|||_2.||\delta_X||_2, \qquad ||A^{-1}B||_2 \leq |||A^{-1}|||_2.||B||_2 \tag{2}$$

we deduce

$$\frac{|||\delta_X|||_2}{|||X|||_2} \leq |||A|||_2.|||A^{-1}|||_2 \frac{|||\delta_B|||_2}{|||B|||_2}.$$

The conditioning of A relative to the norm $|| - ||_2$ is then the quantity $|||A|||_2.|||A^{-1}|||_2$ which is therefore the quotient $\frac{\mu_n}{\mu_1}$ of the largest singular value by the smallest. Using that in finite dimension, the unit sphere is compact, the inequalities of (2) are optimal, i.e. the cases of equality exist, so that the previous upper bound is optimal.

Let us illustrate the phenomenon :

— we take $A \in GL_2(\mathbb{R})$ symmetric positive definite so that the singular values are equal to the eigenvalues.

— Let e_1 and e_2 be the eigenvectors associated to $\lambda_1 \leq \lambda_2$ and assume λ_2 large and λ_1 small.

— Let $B = e_2$ be such that $X = \frac{1}{\lambda_2} e_2$ which is small, and

— $\delta_B = \lambda_1 e_1$, which is therefore small,

so that $\delta_X = e_1$ is large.

Remarque: we can also perturb the matrix A while keeping B : again the perturbation of X is controlled by the conditioning of A.

3.2 Eigenvectors and eigenspaces

Definition 67. *A vector $v \in E$ is said to be* proper *by an endomorphism f if it is non-zero and if there exists a scalar $\lambda \in \mathbb{K}$ such that $f(v) = \lambda v$;*

the scalar λ is then called an eigenvalue. *We denote by* Spec$f$ *the set of eigenvalues of f : it is the spectrum of f.*

Remarque: an eigenvector defines a stable line by f ; more generally a subspace F of E is said to be *stable* by f if $f(F) \subset F$. If we take a basis of F that we complete with a basis of E, the matrix of f in this basis will be block triangular, i.e. of the form $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ with $A \in \mathbb{M}_{n_1}(K)$ and $B \in \mathbb{M}_{n_2}(K)$ with $n_1 + n_2 = n$.
Examples : the kernel $\operatorname{Ker} f$ and the image $\operatorname{Im} f$ are subspaces stable by f.

Proposition 68. *Let E be a vector space with a basis $(e_i)_{1 \leq i \leq n}$. There then exists a unique application* $\det_{(e_i)_i} : E^n \to \mathbb{K}$ *which is* alternating multilinear *and such that* $\det_{(e_i)_i}(e_1, \cdots, e_n) = 1$.

Definition 69. *For $E = \mathbb{K}^n$ equipped with the canonical basis and $\mathbb{M}_n(\mathbb{K})$ identified via its column vectors to E^n, the application of the previous proposition defines* $\det : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$ *and is called* the determinant.

Remarque: by elementary operations on the column vectors of a matrix, we show that $\det A \neq 0$ if and only if $A \in GL_n(\mathbb{K})$ as well as the following corollary.

Corollory 70. *For $A, B \in \mathbb{M}_n(\mathbb{K})$ we have* $\det(AB) = \det A . \det B$.

Definition 71. The characteristic polynomial *of an endomorphism $f \in \mathcal{L}(E)$ is the determinant $\chi_A(X) := \det(A(f) - XI_n) \in \mathbb{K}[X]$ where $A(f)$ is the matrix of f in any basis of E.*

Remarque: the fact that χ_A is independent of the basis comes from the fact that $\det(P^{-1}AP) = \det A$ according to the previous corollary.
Remarque: in the case where χ_A is totally decomposed (for example if $\mathbb{K} = \mathbb{C}$), the product of the eigenvalues is equal to $(-1)^n$ times the constant coefficient of χ_A and therefore to the determinant of A.

Lemme 72. *On $\mathbb{C}$, the norm of the product of the eigenvalues of a matrix $A \in \mathbb{M}_n(\mathbb{C})$ is equal to the product of the absolute values of its singular values.*

Remarque: to the corollary **??**, we will give refinements of this equality.
Preuve : It suffices to note that the determinant of a unitary matrix is necessarily of modulus 1, since $U^*U = I_n$ implies that $|\det U| = 1$.
Remarque: the formation of the characteristic polynomial $A \mapsto \chi_A(X)$ is clearly continuous since it is polynomial. Let us recall that, cf. the corollary **??**, the roots depend continuously on their polynomial, so that the eigenvalues depend continuously on the matrix. The following proposition quantifies this property.

Proposition 73. ***(Gershgorin disks)*** *The eigenvalues of* $A = (a_{i,j})_{1\leq i,j\leq n} \in \mathbb{M}_n(\mathbb{C})$ *belong to the union of the closed disks centered at* $a_{i,i}$ *and of radius* $\sum_{j\neq i} |a_{i,j}|$.

Preuve : The result follows directly from Hadamard's lemma applied to $A-\lambda \mathrm{Id}$. Recall that this lemma states that if for all $1 \leq i \leq n$, we have $|a_{i,i}| > \sum_{j\neq i} |a_{i,j}|$ then A is invertible. Indeed, let X with coordinates $(x_i)_{1\leq i\leq n}$ in the kernel of A and let i_0 be such that $|x_{i_0}|$ is maximal among the $|x_i|$. From the equality $a_{i_0,i_0}x_{i_0} = -\sum_{j\neq i_0} a_{i_0,j}x_j$ we deduce the upper bound $|a_{i_0,i_0}x_{i_0}| \leq |x_{i_0} \sum_{j\neq i_0} |a_{i_0,j}|$ and therefore $x_{i_0} = 0$ or $X = 0$.

From a practical point of view, let us examine the perturbation undergone by the eigenvalues when the matrix is perturbed. Let us start with the simple example given by the companion matrix of $X^{100} - 10^{-100}$ whose eigenvalues are of modulus equal to 1. This matrix is thus a very weak perturbation of the Jordan block of size 100 whose eigenvalues are all zero. In conclusion, a perturbation of order 10^{-100} leads us to a perturbation of order 0.1 which is enormous. Let us try to quantify this phenomenon : to do so, we consider a matrix norm $||.||$ such that for any diagonal matrix $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$, we have $||D|| = \max_i |\lambda_i|$. For example, we can take the norms $|||.|||_1$, $|||.|||_2$ or $|||.|||_\infty$.

Proposition 74. *Let* A *be a diagonalizable matrix with* $\mathrm{Spec}A = \{\lambda_1, \cdots, \lambda_n\}$. *Then*

$$\mathrm{Spec}(A + \delta_A) \subset \bigcup_{i=1}^{n} \big\{ z \in \mathbb{C} : \ |z - \lambda_i| \leq \gamma(A)||\delta_A|| \big\},$$

where

$$\gamma(A) = \inf\{||P||.||P^{-1}|| : \ P^{-1}AP = \mathrm{diag}(\lambda_1, \cdots, \lambda_n).$$

Remarque: Thus the control of the eigenvalues is given by the conditioning of the passage matrices and not by the conditioning of the starting matrix.

Preuve : Let P diagonalize A, i.e.

$$P^{-1}AP = D := \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$$

and let λ be an eigenvalue of $A + \delta_A$. If $\lambda \in \{\lambda_1, \cdots, \lambda_n\}$, there is nothing to show, otherwise $(D - \lambda I_n)$ is invertible and we can write

$$\begin{aligned} P^{-1}(A + \delta_A - \lambda I_n)P &= D - \lambda I_n + P^{-1}\delta_A P \\ &= (D - \lambda I_n)(I_n + (D - \lambda I_n)^{-1}P^{-1}\delta_A P). \end{aligned}$$

The matrix $(I_n+(D-\lambda I_n)^{-1}P^{-1}\delta_A P)$ is not invertible, so -1 is an eigenvalue of $(D-\lambda I_n)^{-1}P^{-1}\delta_A P)$, so that according to **??** its norm is ≥ 1, which gives :

$$1 \leq ||(D - \lambda I_n)^{-1}P^{-1}\delta_A P)|| \leq ||(D - \lambda I_n)^{-1}||.||P^{-1}||.||\delta_A||.||P||.$$

As by hypothesis $||(D - \lambda I_n)^{-1}|| = \frac{1}{\min|\lambda_i - \lambda|}$, there exists at least one index i such that

$$|\lambda_1 - \lambda| \leq ||P||.||P^{-1}||.||\delta_A||.$$

Remarque: In particular if A is normal, the passage matrix is orthogonal and its conditioning is equal to 1. In other words when a normal matrix is perturbed, its eigenvalues are perturbed in the same proportion. The most interesting case is certainly the one where A and δ_A are both symmetric, we refer to the §4.8 for a more precise study in this situation.

Definition 75. *The eigensubspace E_λ (resp. characteristic $E(\lambda)$) associated with an eigenvalue λ is* $\mathrm{Ker}(f - \lambda\mathrm{Id})$ *(resp.* $\mathrm{Ker}(f - \lambda\mathrm{Id})^n$ *where n is the dimension of E, or more simply the multiplicity of λ in the minimal polynomial μ_f).*

Remarque: the dimension of the characteristic subspace is equal to the multiplicity of λ in the characteristic polynomial.

Lemme 76. ***(kernel lemma)***
If $P = P_1P_2$ is a polynomial that cancels f with $P_1 \wedge P_2 = 1$ then $E =$ $\mathrm{Ker}\, P_1(f) \oplus \mathrm{Ker}\, P_2(f)$.

Preuve : We start from a Bezout relation $UP_1 + VP_2 = 1$ so that for all $e \in E$, we have $e = e_1 + e_2$ with $e_1 = UP_1(f)(e)$ and $e_2 = VP_2(f)(e)$. We then have

$$P_2(f) = (e_1) = U(f) \circ P_1P_2(f)(e_1) = 0 \text{ and } P_1(f) = V(f) \circ P_1P_2(f)(e) = 0$$

and therefore $e_1 \in \mathrm{Ker}\, P_2(f)$ and $e_2 \in \mathrm{Ker}\, P_1(f)$ and $E = \mathrm{Ker}\, P_1(f) + \mathrm{Ker}\, P_2(f)$. Furthermore if $e \in \mathrm{Ker}\, P_1(f) \cap \mathrm{Ker}\, P_2(f)$ then $e_1 = e_2 = 0$ and $e = e_1 + e_2 = 0$.
Remarque: it is important to note that the projectors on each of these stable subspaces parallel to the other, are polynomials in f.

3.3 Minimal Polynomial

The theory of the reduction of an endomorphism is totally controlled by a series of polynomials associated with it, called *its similarity invariants*. We will see that the characteristic polynomial is the product of the similarity invariants. In general, the answer to a question about an endomorphism is expressed using the full power of the similarity invariants. The largest of the similarity invariants is given by the following definition where we recall that given an endomorphism f and a polynomial $P(X) = \sum_i a_iX^i \in \mathbb{K}[X]$, $P(f)$ denotes the endomorphism $\sum_i a_if^i$ where f^i denotes $f \circ \cdots \circ f$.

Definition 77. *For $f \in \mathcal{L}(E)$, the set I_f of polynomials $P \in \mathbb{K}[X]$ such that $P(f)$ is the zero endomorphism, is an ideal of $\mathbb{K}[X]$; this ring being principal, there exists a unique unitary polynomial μ_f, called* minimal polynomial *of f, such that $I_f =< \mu_f >$.*

Remarque: since E is finite-dimensional, the family $\mathrm{Id}, f, f^2, \cdots, f^{n^2+1}$ is related so that I_f is not the zero ideal and thus μ_f is not the zero polynomial.

Theorem 78. ***(Cayley-Hamilton)***
The characteristic polynomial χ_f belongs to I_f, i.e. $\chi_f(f)$ is the zero endomorphism.

3.4 Trigonalization

Definition 79. *An endomorphism is said to be* trigonalizable *if there exists a basis in which its matrix is upper triangular.*

Remarque: another way to state this property is to require that there exists a complete flag

$$\{0\} = F_0 \subset F_1 \subset \cdots \subset F_n = E$$

with $\dim F_i = i$, stable by f, i.e. for all $i = 0, \cdots, n$ we have $f(F_i) \subset F_i$.

Theorem 80. *An endomorphism f is trigonalizable if and only if χ_f is split on $\mathbb{K}$.*

Remarque: thus if K is algebraically closed all endomorphisms are trigonalizable.

We are now interested in the question of simultaneous trigonalization, i.e. given a subset $\mathcal{E}$ of $\mathcal{L}(E)$, we look for flags, if possible complete, stable by all $u \in \mathcal{E}$. We recall that if u is an endomorphism of a vector space E and if $F \subset E$ is a subspace stable by u then u induces an endomorphism $\overline{u}$ of E/F.

Given a subset $\mathcal{E}$ of $\mathcal{L}(E)$ and $G \subset F \subset E$ of subspaces stable by all elements u of $\mathcal{E}$, the set of quotients of $\mathcal{E}$ for $\{G \subset F\}$ is by definition $\{\overline{u} \in \mathcal{L}(F/G) : \ u \in \mathcal{E}\}$.

Definition 81. *A property P will be said to be stable by quotients if for any set $\mathcal{E} \subset \mathcal{L}(E)$ consisting of elements satisfying P then the quotient set of $\mathcal{E}$ for $\{G \subset F\}$ is also consisting of elements of $\mathcal{L}(F/G)$ satisfying P.*

General principle : let $\mathcal{P}$ be a set of properties stable by quotients and verifying the following property : for any $\mathcal{E} \subset \mathcal{L}(E)$ consisting of elements verifying $\mathcal{P}$ with $\dim E > 1$, $\mathcal{E}$ is reducible i.e. there exists a non-trivial vector subspace F of E stable by all elements of $\mathcal{E}$. Then $\mathcal{E}$ is triangularizable.
Example : for E a finite-dimensional $\mathbb{C}$-vector space, any commutative subset of $\mathcal{L}(E)$ is triangularizable. Indeed, commutativity is clearly a property stable by quotient. The reducibility property will then follow from the following facts :

— every endomorphism admits an eigenvalue and
— every eigensubspace of A is stable under any matrix B commuting with A.

Remarque 82. using the following facts :

— the eigenvalues of a triangular matrix are its diagonal terms ;

— the diagonal terms of a product (resp. sum) of a triangular matrix are the products $t^{(1)}_{i,i}t^{(2)}_{i,i}$ (resp. $t^{(1)}_{i,i} + t^{(2)}_{i,i}$) of their diagonal terms.

we deduce that if $\{A_1, \cdots, A_k\}$ are simultaneously diagonalizable and if p is a noncommutative polynomial (i.e. a linear combination of words) in $\{A_1, \cdots, A_k\}$ then

$$\sigma(p(A_1, \cdots, A_k)) \subset p(\sigma(A_1), \cdots, \sigma(A_k))$$

where $p(\sigma(A_1), \cdots, \sigma(A_k))$ denotes the set of $p(\lambda_1, \cdots, \lambda_k)$ where for all $i = 1, \cdots, k$, $\lambda_i \in \sigma(A_i)$. In the following we will give the converse to this result.

If $\mathcal{A}$ is a subalgebra of $\mathcal{L}(E)$ the set $\mathcal{A}.x := \{Ax : \ A \in \mathcal{A}\}$, where $x \in E$, is a stable subspace under $\mathcal{A}$. If $\mathcal{A}.x = E$, we say that x is *a cyclic vector* for $\mathcal{A}$. The determination of the subalgebras of $\mathcal{L}(E)$ that have non-trivial invariant subspaces is settled by the following theorem which deals with the reducibility part of the general principle stated above in the case of subalgebras of $\mathcal{L}(E)$.

Theorem 83. ***(Burnside's)***
Any proper subalgebra of $\mathcal{L}(E)$ is reducible.

Preuve : Let $\mathcal{A}$ be an irreducible subalgebra of $\mathcal{L}(E)$; since every endomorphism is a sum of endomorphisms of rank 1, we will show that every endomorphism of rank 1 belongs to $\mathcal{A}$.

Let us first show that $\mathcal{A}$ contains an element of rank 1. Let $u_0 \in \mathcal{A}$ be nonzero of minimal rank; if this rank is strictly greater than 1, then there exist vectors x_1 and x_2 such that $(u_0(x_1), u_0(x_2))$ is linearly independent. Since $\{u \circ u_0(x_1) : \ u \in \mathcal{A}\} = E$, there exists $u_1 \in \mathcal{A}$ such that $u_1 \circ u_0(x_1) = x_2$ and therefore $(u_0 \circ u_1 \circ u_0(x_1), u_0(x_1))$ is free. Then let λ be such that the restriction of $u_1 \circ u_0 - \lambda\mathrm{Id}$ to $u_0(E)$ is not invertible; $(u_0 \circ u_1 - \lambda\mathrm{Id})u_0$ is non-zero because the image of x_1 is non-zero, and $(u_0 \circ u_1 - \lambda\mathrm{Id}) \circ u_0$ is of rank strictly smaller than that of u_0, hence the contradiction and therefore u_0 is of rank 1.

For y_0 in the image of u_0, we consider the linear form φ_0 defined by $u_0(x) = \varphi_0(x)y_0$. Let $u \in \mathcal{L}(E)$ be defined by $u(x) = \varphi(x)y$ where $y \in E$ and $\varphi \in E^*$. Let us then show that u belongs to $\mathcal{A}$. For $v \in \mathcal{A}$, we have $u_0 \circ v \in \mathcal{A}$ and $u_0 \circ v(x) = \varphi_0(v(x))y_0$. Let then $F' \subset E^*$, the set of linear forms φ such that $x \mapsto \varphi(x)y_0$ belongs to $\mathcal{A} : F'$ is clearly a subspace of E^*. If this subspace were strict, there would exist $x_0 \neq 0$ such that $\varphi(x_0) = 0$ for all $\varphi \in F'$ (a finite-dimensional vector space is reflexive). The contradiction then follows from the fact that $\varphi_0(v(x_0)) = 0$ for all $v \in \mathcal{A}$ implies that x_0 is zero because $\{v(x_0) : \ v \in \mathcal{A}\} = E$. So let $v_1 \in \mathcal{A}$ be such that $\varphi = \varphi_0 \circ v_1$.

Similarly, as $y_0 \neq 0$, then $\{v(x_0) : \ v \in \mathcal{A}\} = E$ and so for all $y \in E$ let $v_2 \in \mathcal{A}$ be such that $v_2(y_0) = y$ and so $u = v_2 \circ u_0 \circ v_1$.

Corollory 84. *The only two-sided ideals of $\mathcal{L}(E)$ are $\{0\}$ and $\mathcal{L}(E)$.*

Preuve : Let $\mathcal{I}$ be a two-sided ideal of $\mathcal{L}(E)$ not reduced to 0. It is then sufficient to show that $\mathcal{I}$ is irreducible. If $u \neq 0$ belongs to $\mathcal{I}$, for all $0 \neq x \in E$, there exists $v \in \mathcal{L}(E)$ such that $u \circ v(x) \neq 0$. Let $y \in E$ and $w \in \mathcal{L}(E)$ be such that $w \circ u \circ v(x) = y$. We have $w \circ u \circ v \in \mathcal{I}$ so that any vector $x \neq 0$ is cyclic for $\mathcal{I}$ and therefore $\mathcal{I}$ is irreducible.

Corollory 85. *Let E be a finite-dimensional $\mathbb{C}$-vector space then every algebra automorphism ϕ of $\mathcal{L}(E)$ is inner, i.e. there exists $P \in GL(E)$ such that for all $A \in \mathcal{L}(E)$, $\phi(A) = PAP^{-1}$.*

Preuve : Let $A_0 \in \mathcal{L}(E)$ be an idempotent of rank 1, $\phi(A_0)$ is then an idempotent, let us show that it is also of rank 1. The set $\{A_0BA_0 : \ B \in \mathcal{L}(E)\}$ is a vector subspace of $\mathcal{L}(E)$ of dimension 1 : we can identify it with $\mathcal{L}(\operatorname{Im} A_0)$. Its image by ϕ, $\{\phi(A_0)C\phi(A_0) : \ C \in \mathcal{L}(E)\}$ is therefore also a subspace of $\mathcal{L}(E)$ of dimension 1 identified with $\mathcal{L}(\operatorname{Im} \phi(A_0))$ so that $\phi(A_0)$ is of rank 1. Since all idempotents of rank 1 are similar to $\operatorname{diag}(1, 0, \cdots, 0)$, even if it means composing ϕ by $A \mapsto PAP^{-1}$, we can assume that $\phi(A_0) = A_0$.

Let x_0 be a direction vector of $\operatorname{Im} A_0$ and let $P \in \mathcal{L}(E)$ be defined by $P(Bx_0) = \phi(B)x_0$: if $B_1x_0 = B_2x_0$ then as $A_0x_0 = x_0$, we have $(B_1 - B_2)A_0 = 0$ and therefore $(\phi(B_1) - \phi(B_2))A_0 = 0$ so that $\phi(B_1)x_0 = \phi(B_2)x_0$ and P is well defined and obviously linear. Suppose that $\phi(B)x_0 = 0$ so that $\phi(B)\phi(A_0) = \phi(BA_0) = 0$ and therefore $BA_0 = 0$ or $Bx_0 = 0$ which proves the injectivity of P and since we are in finite dimension $P \in GL(E)$.

Let then be $A \in \mathcal{L}(E)$, we have $P(AB)x_0 = \phi(AB)x_0 = \phi(A)\phi(B)x_0 = \phi(A)PBx_0$ and therefore $PAy = \phi(A)Py$ for all $y = Bx_0$. When B describes $\mathcal{L}(E)$, y describes E and therefore $PA = \phi(A)P$ for all $A \in \mathcal{L}(E)$ hence the result.

Corollory 86. *Any algebra of nilpotent endomorphisms is triangularizable.*

Preuve : The property of being nilpotent is stable by quotient as moreover there exist elements of $\mathcal{L}(E)$ that are not nilpotent, any algebra consisting of nilpotent endomorphisms is, by Burnside's theorem, reducible. Triangularization then follows from the general principle stated above.

Theorem 87. *If $\mathcal{A}$ is a subalgebra of $\mathcal{L}(E)$ then $\mathcal{A}$ is triangularizable if and only if every commutator $BC - CB$ with $B, C \in \mathcal{A}$ is nilpotent.*

Preuve : If $\mathcal{A}$ is triangularizable then by the remark 82, we have $\sigma(BC - CB) = \{0\}$ and thus $BC - CB$ is nilpotent. Conversely the property of having nilpotent commutators is stable by quotient and since there exist non-nilpotent commutators in $\mathcal{L}(E)$, by Burnside's theorem, $\mathcal{A}$ is reducible ; the triangularization then follows from the general principle.

Remarque: we can thus see triangularizability as a generalization of commutativity ; we relax the condition of being zero for a commutator, by requiring that it be nilpotent.

Theorem 88. ***((McCoy)*** *The pair $\{A, B\}$ is triangularizable if and only if $p(A, B)(AB - BA)$ is nilpotent for any commutative polynomial p in A and B.*

Preuve : The direct meaning follows from the remark 82. For the converse according to the general principle it suffices to show that the algebra $\mathcal{A}$ generated by A, B is reducible as soon as $\dim E > 1$. If $AB = BA$ it is clear, otherwise let $x \in E$ be such that $(AB - BA)x \neq 0$ and $C \in \mathcal{L}(E)$ verifying $C(AB - BA)x = x$. If $\mathcal{A}$ were irreducible then by Burnside's theorem it would be equal to $\mathcal{L}(E)$ and therefore $C \in \mathcal{A}$ and therefore of the form $p(A, B)$. The contradiction then follows from the fact that $C(AB - BA)$ is not nilpotent.

Corollory 89. *A unitary subalgebra $\mathcal{A}$ of $\mathcal{L}(E)$ is triangularizable if and only if $\mathcal{A}/\mathrm{Rad}\mathcal{A}$ is commutative, where*

$$\mathrm{Rad}\mathcal{A} = \{A \in \mathcal{A} : \sigma(AB) \subset \{0\}\ \forall B \in \mathcal{A}\}$$

is the radical of A, i.e. the intersection of all maximal right (or left) ideals of $\mathcal{A}$.

Preuve : If $\mathcal{A}$ is triangularizable and if $B, C \in \mathcal{A}$ then for all $A \in \mathcal{A}$ according to the remark 82, we have $\sigma\Big((BC - CB)A\Big) = \{0\}$ and therefore $BC - CB \in \mathrm{Rad}\mathcal{A}$ i.e. $\mathcal{A}/\mathrm{Rad}\mathcal{A}$ is commutative.

Conversely if $\mathcal{A}/\mathrm{Rad}\mathcal{A}$ is commutative then $BC - CB \in \mathrm{Rad}\mathcal{A}$ for all $B, C \in \mathcal{A}$ so that $\mathcal{A}$ is triangularizable according to the corollary 86.

We will focus on vector subspaces of $\mathcal{L}(E)$ stable under certain non-associative multiplications such as for example the *Lie algebras* stable under the Lie bracket $[A, B] = AB - BA$, or the Jordan algebras stable under $(A, B) \mapsto AB + BA$. The most famous result is Engel's theorem below.

Theorem 90. *Let $\mathcal{N}$ be a subset of the nilpotent cone of $\mathcal{L}(E)$ verifying the following property : for all $A, B \in \mathcal{N}$ there exists a noncommutative polynomial p in $A$and B such that $AB + p(A, B)A \in \mathcal{N}$. Then $\mathcal{N}$ is triangularizable.*

Preuve : We reason by induction on the dimension n of E ; the case $n = 1$ being trivial, let us therefore assume the result acquired up to rank n and let us treat that of $n + 1$. Let $\mathcal{F}$ be the set of subspaces of E that are intersections of kernels of elements of $\mathcal{N}$, i.e. of the form $V_{\mathfrak{S}} = \cap_{A \in \mathfrak{S}} \mathrm{Ker}\, A$ where $\mathfrak{S} \subset \mathcal{N}$, and let $K \in \mathcal{F}$ be of minimal non-zero dimension. Let us then denote $\mathcal{N}_0 = \{A \in \mathcal{N} :\ Ax = 0\ \forall x \in K\}$; the set $\overline{\mathcal{N}_0} \subset \mathcal{L}(E/K)$

verifies the hypotheses of the statement so that according to the recurrence hypothesis $\overline{\mathcal{N}_0}$ is triangularizable and therefore $\mathcal{N}_0$ also because its elements are zero on K.

It is then sufficient to show that $\mathcal{N}_0 = \mathcal{N}$. Otherwise, let $B \in \mathcal{N}$ and $x \in K$ be such that $Bx \neq 0$. If K were a stable subspace of the nilpotent endomorphism B, there would exist $x_0 \in K$ such that $Bx_0 = 0$ and $\operatorname{Ker} B \cap K$ would be a non-zero element of $\mathcal{F}$ of dimension strictly smaller than that of K which is not by hypothesis. Let then be $x_1 \in K$ and $A_1 \in \mathcal{N}_0$ such that $A_1Bx_1 \neq 0$ and let p_1 be a noncommutative polynomial in A_1 and B such that $B_1 = A_1B + p_1(A_1, B)A_1 \in \mathcal{N}$. Since $B_1x_1 \neq 0$ as before K is not a stable subspace of B_1 ; let then $A_2 \in \mathcal{N}_0$ and $x_2 \in K$ such that $A_2B_1x_2 \neq 0$. Let p_2 be such that $B_2 = A_2B_1 + p_2(A_2, B_1)A_2 \in \mathcal{N}$. Continuing the process, we construct

$$\{A_1, A_2, \cdots, A_{n+1}\} \subset \mathcal{N}_0, \qquad \{B_1, B_2, \cdots, B_n\} \subset \mathcal{N}$$

and vectors $\{x_1, \cdots, x_{n+1}\}$ such that

$$A_{n+1}A_n \cdots A_2A_1Bx_{n+1} = A_{n+1}B_nx_{n+1} \neq 0.$$

Since $\mathcal{N}_0$ is triangularizable, any product of $n+1$ of its elements is zero so that $A_{n+1} \cdots A_1 = 0$ which contradicts $A_{n+1} \cdots A_1Bx_{n+1} \neq 0$. We therefore deduce that $\mathcal{N}_0 = \mathcal{N}$ which is trigonalizable.

Corollory 91. *A set $\mathcal{N}$ of nilpotent elements of $\mathcal{L}(E)$ that satisfies one of the following properties is trigonalizable :*

- — ***Jacobson's theorem** : for all $A, B \in \mathcal{N}$, there exists a scalar c such that $AB - cBA \in \mathcal{N}$;*
- — ***Engel's theorem** : for all $A, B \in \mathcal{N}$, $[A, B] = AB - BA \in \mathcal{N}$;*
- — *for all $A, B \in \mathcal{N}$, $AB + BA \in \mathcal{N}$.*

Corollory 92. *Let $\mathcal{E} \subset \mathcal{L}(E)$ be stable by the Lie bracket. Then $\mathcal{E}$ is triangularizable if and only if all its commutators are nilpotent.*

Preuve : The direct meaning follows from the remark 82. Conversely, as the property is clearly stable by quotient according to the general principle, it suffices to show that $\mathcal{E}$ admits a stable subspace. Let $\mathcal{N}$ be the set of commutators of $\mathcal{E}$; if $\mathcal{N} = \{0\}$ then $\mathcal{E}$ is commutative and therefore admits a stable subspace. Otherwise, according to the previous corollary, $\mathcal{N}$ is trigonalizable so that $K = \cap_{N \in \mathcal{N}} \operatorname{Ker} N$ is a non-zero subspace stable by all elements of $\mathcal{E}$. Indeed, for $x \in K$ and $B \in \mathcal{E}$, for all $A \in \mathcal{N}$, we have $Ax = 0$ and $(AB - BA)x = 0$ so that $ABx = 0$ is $Bx \in K$, hence the result.

Theorem 93. ***(Levitzki)*** *Any semigroup S of nilpotent elements of $\mathcal{L}(E)$ is triangularizable.*

Preuve : Since nilpotence is a stable property by quotient, it suffices according to the general principle to show the reducibility of S as soon as $\dim E > 1$. The trace is a linear form that vanishes on S and therefore on the algebra generated by S which is simply the vector space generated by S. The result follows then from Burnside's theorem and from the fact that there exist elements of $\mathcal{L}(E)$ with non-zero traces.
Remarque: we could also have deduced this result from Jacobson's theorem for $c = 0$.

Theorem 94. ***(Kolchin)*** *If every element of the semigroup S is unipotent then S is triangularizable.*

Preuve : Since unipotency is a stable property by quotient it suffices according to the general principle to show the reducibility of S as soon as $\dim E > 1$. If all the commutators are zero then S is abelian and therefore reducible. Otherwise let $C = AB - BA \neq 0$ and let $\mathcal{A}$ be the algebra generated by S. The bilateral ideal generated by C is then contained in the kernel of the linear trace form : indeed XCY is written as a linear combination of product DCE with $D, E \in S$ so that $\mathrm{tr}(DCE) = \mathrm{tr}(DABE) - \mathrm{tr}(DBAE) = n - n = 0$. Since $\mathcal{L}(E)$ does not admit a non-zero proper ideal, we deduce that $\mathcal{A}$ is not equal to $\mathcal{L}(E)$ and is therefore reducible according to Burnside's theorem.

If $\mathcal{E} \subset \mathcal{L}(E)$ is triangularizable then the trace function is permutable on $\mathcal{E}$, i.e. for all $A_1, \cdots, A_m \in \mathcal{E}$ and for all permutation $\sigma \in \mathfrak{S}_m$, we have

$$\mathrm{tr}(A_1 A_2 \cdots A_m) = \mathrm{tr}(A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(m)}).$$

The converse is true and follows simply from theorem 87.

Proposition 95. *Let $\mathcal{E} \subset \mathcal{L}(E)$, then $\mathcal{E}$ is triangularizable if and only if the trace is permutable on $\mathcal{E}$.*

Preuve : For all $A, B, C \in \mathcal{E}$, we have $\mathrm{tr}\Big((AB - BA)C\Big) = 0$. In particular we deduce that for all $k > 0$, $\mathrm{tr}(AB - BA)^k = 0$ and therefore $AB - BA$ is nilpotent hence the result according to 87.

Corollory 96. *If S is a semigroup that verifies one of the following properties, is triangularizable :*

(i) ***Kaplansky*** *: the trace is constant on S ;*

(ii) the trace is multiplicative on S.

Furthermore, all in the first situation the diagonal terms in such a triangularization depend only on S and are either equal to 0 or 1 ; in the second there exists j that depends only on S such that all diagonal terms in a triangularization, except the one in (j, j), are zero.

Preuve : (i) Triangularizability follows from the previous proposition. Then let $A \in S$ so that the $\mathrm{tr}A^k$ are constant for all $k \geq 1$ equal to c. Let $\lambda_1, \cdots, \lambda_m$ be the non-zero eigenvalues of A; it is then sufficient to show that for all $\mathrm{tr}A = m$. Indeed, according to Newton's relations, the set $\{\lambda_1, \cdots, \lambda_m\}$ is uniquely determined by the $S_k = \lambda_1^k + \cdots + \lambda_m^k$ for $k = 1, \cdots, m$; if all the S_k are equal to 1 then $\lambda_1 = \cdots = \lambda_m = 1$ is trivially suitable. Let us denote σ_i the usual symmetric functions of the λ_j so that we have

$$S_{m+1} - S_m\sigma_1 + S_{m-1}\sigma_2 + \cdots + (-1)^m S_1 = 0\sigma_n$$

$$S_m - \sigma_1 S_{m-1} + \cdots + (-1)^m m\sigma_m = 0$$

so that since $\sigma_m \neq 0$, we obtain $c = m$.

Thus the eigenvalues of A are 0 or 1; if the diagonal terms of $A, B \in S$ were not equal then we would have $\mathrm{tr}(ST) < \mathrm{tr}S$ which is not the case.

(ii) Triangularizability follows from the previous proposition. Let $A \in S$ have eigenvalues $\lambda_1, \cdots, \lambda_n$. By hypothesis we have $\sum_i \lambda_i^k = (\sum_i \lambda_i)^k$. If $\sum_i \lambda_i = 0$ then all λ_i are zero hence the result. If $\sum_i \lambda_i = b \neq 0$ then for all $k \geq 1$, $\sum_i (a_i/b)^k = 1$ so that according to the proof of (i), there exists a unique j such that $\lambda_j \neq 0$. Trace multiplicativity implies that this j is the same for all $A \in S$.

Corollory 97. *Let G be a subgroup of $GL(E)$; the following properties are then equivalent :*

(i) G is triangularizable ;

(ii) for all $g \in G$, the trace is constant on $gD(G)$;

(iii) the trace is constant on $D(G)$;

(iv) for all $A \in D(G)$, $\sigma(A) = \{1\}$, i.e. A is unipotent.

Preuve : (i) $\Rightarrow$ (ii) : according to the previous proposition, the trace is permutable on G so that for all $g \in G$ and $h \in D(G)$, we have $\mathrm{tr}(gh) = \mathrm{tr}(g\mathrm{Id}) = \mathrm{tr}g$.

(ii) $\Rightarrow$ (iii) : immediate

(iii) $\Leftrightarrow$ (iv) : follows from the Kaplansky theorem proved above using that the elements of G are invertible.

(iv) $\Rightarrow$ (i) : since property (iv) is clearly stable by quotient, it suffices to show that G is reducible. By Kolchin's theorem, $D(G)$ is triangularizable : let $\{0\} = F_0 \subset F_1 \subset \cdots \subset F_{n-1} \subset F_n = E$ be the corresponding complete flag. Assume $D(G) \neq \{\mathrm{Id}\}$ because otherwise G is commutative and the result is clear. Let F be the vector subspace generated by $\bigcup_{h \in D(G)} (h - \mathrm{Id})(E)$. As for all $h \in D(G)$, $(h - \mathrm{Id})(E) \subset F_{n-1}$, F is a strict subspace of E which is also invariant under G : indeed, let $g \in G$ then for all $h \in D(G)$, we have $g(h - \mathrm{Id}) = (ghg^{-1} - \mathrm{Id})g$ and as $ghg^{-1} \in D(G)$, we have $g\Big((h - \mathrm{Id})(E)\Big) \subset (ghg^{-1} - \mathrm{Id})(E) \subset F$, hence the result.

Corollory 98. *Let G be a subgroup of $GL(E)$ such that for all $A, B, C \in G$, $\sigma(ABC) = \sigma(BAC)$ then G is triangularizable.*

Remarque: since $\sigma(AB) = \sigma(BA)$, the property of the statement amounts to saying that the spectrum is permutable, i.e. for all $\{A_1, \cdots, A_m\} \subset G$, and for any permutation $\sigma \in \mathfrak{S}_m$, we have $\sigma(A_1 A_2 \cdots A_m) = \sigma(A_{\sigma(1)} \cdots A_{\sigma(m)})$. The reader will note that this property is weaker than that of permutability of the trace because here we just require that the sets of eigenvalues are equal, without taking into account the multiplicities.

Preuve : The result follows directly from the implication (iv) $\Rightarrow$ (i) in the previous corollary. Indeed for all $h \in D(G)$, we have $\sigma(h) = \sigma(\mathrm{Id}) = \{1\}$ so that h is unipotent.

3.5 Iterated kernels

Let E be a $\mathbb{K}$-vector space of finite dimension n and $u \in \mathcal{L}(E)$. For all $\lambda \in \mathbb{K}$ and $r \geq 1$, we note

$$K_r(\lambda) := \mathrm{Ker}(u - \lambda \mathrm{Id})^r \quad \text{and} \quad I_r(\lambda) := \mathrm{Im}(u - \lambda \mathrm{Id})^r,$$

and we note $dK_r(\lambda) := \dim_{\mathbb{K}} K_r(\lambda)$ and $dI_r(\lambda) := \dim_{\mathbb{K}} I_r(\lambda)$. We also set $dK_0(\lambda) = 0$ and $dI_0(\lambda) = n$.

Proposition 99. *The sequence $dK_r(\lambda)$ (resp. $dI_r(\lambda)$) is first strictly increasing (resp. decreasing) then stationary from an index r_0 (resp . the same index r_0). Moreover the sequence*

$$\delta_r(\lambda) := dK_r(\lambda) - dK_{r-1}(\lambda)$$

for $r \geq 1$ is decreasing up to rank r_0 then stationary equal to 0.

Preuve : If we consider $u - \lambda \mathrm{Id}$, we assume $\lambda = 0$ and we simply note K_r for $K_r(0)$. Then let r be such that $K_r = K_{r+1} \subset K_{r+2}$. For $x \in K_{r+2}$ we have $u(x) \in K_{r+1} = K_r$ and therefore $u^{r+1}(x) = 0$, i.e. $x \in K_{r+1}$ and therefore $K_{r+2} = K_{r+1}$, which shows the first part of the statement since with the rank theorem $dK_r(\lambda) + dI_r(\lambda) = n$.

As for δ_r, consider the endomorphism $K_r \longrightarrow K_{r-1}/K_{r-2}$ which sends x to the image of $u(x) \in K_{r-1}/K_{r-2}$. Its kernel is clearly K_{r-1} so that we have an injection

$$K_r/K_{r-1} \hookrightarrow K_{r-1}/K_{r-2}$$

and therefore $\delta_r \geq \delta_{r-1}$, hence the result.

Proposition 100. *Let $u \in \mathcal{L}(E)$ whose characteristic polynomial is split. We denote*

$$E = \bigoplus_{\lambda \in \mathrm{Spec}(u)} E(\lambda)$$

the decomposition of E into characteristic subspaces. For each $\lambda \in \mathrm{Spec}u$, there exists a basis of $E(\lambda)$ in which the matrix of u is block diagonal of the type

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

where the block sizes are given by the lengths of the rows of the array of Young, cf. the figure **??***, associated with u and λ whose columns are of size $\delta_r(\lambda)$.*

One way to construct this Young tableau is as follows : we take a vector e_1 of $K_r - K_{r-1}$ and we denote for $i = 1, \cdots, r-1$, $e_{i+1} = u^i(e_1)$. If $\dim K_r/K_{r-1} > 1$, we choose a vector $e_{r+1} \in K_r$ such that the images of e_1, e_{r+1} in K_r/K_{r-1} are free and we set for $i = 1, \cdots, r-1$, $e_{r+1+i} = u^i(e_{r+1})$. We continue the process until we obtain a basis $e_1, e_{r+1}, \cdots, e_{kr+1}$ of K_r/K_{r-1}. We then choose a vector $e_{(k+1)r+1}$ of K_{r-1} such that the images of $u(e_1), \cdots, u_(e_{kr+1}), e_{(k+1)r+1}$ form a free family of K_{r-1}/K_{r-2} and we set for all $i = 1, r-2$, $e_{(k+1)r+1+i} = u(e_{(k+1)r+1})$. We continue this process until we exhaust all K_i. In parallel, we fill the Young table as in figure **??** in which the image of e_1 is a basis of $\mathrm{Ker}\, u^6/\mathrm{Ker}\, u^5$, the images of $u(e_1), e_7, e_{12}$ form a basis of $\mathrm{Ker}\, u^5/\mathrm{Ker}\, u^4$, the images of $u^4(e_1), u^3(e_7), u^3(e_{12}), e_{17}$ form a basis of $\mathrm{Ker}\, u^2/\mathrm{Ker}\, u$ and

$$u^5(e_1), u^4(e_7), u^4(e_{12}), u(e_{17}), e_{19}$$

form a basis of $\mathrm{Ker}\, u$.

Definition 101. *The house of u is the set of Young tableaux of u for $\lambda \in \mathrm{Spec}(u)$.*

Remarque: over an algebraically closed field, the conjugacy class of an endomorphism corresponds to its house in the sense of the previous definition. An interesting application of this result is the following Brauer theorem.

Theorem 102. ***(Brauer)*** *Let K be any field and for $\sigma \in \mathfrak{S}_n$, we denote by $M(\sigma)$ the associated permutation matrix, i.e. defined by $M(\sigma)(e_i) = e_{\sigma(i)}$ for all $i = 1, \cdots, n$. Then σ and σ' are conjugate in $\mathfrak{S}_n$ if and only if $M(\sigma)$ and $M(\sigma')$ are similar in $GL_n(K)$.*

Preuve : The direct meaning is obvious since if $\sigma' = \tau\sigma\tau^{-1}$ then $M(\sigma') = PM(\sigma)P^{-1}$ with $P = M(\tau)$.

Conversely, suppose that $M(\sigma)$ and $M(\sigma')$ are similar. Let $V^\sigma = \mathrm{Ker}(M(\sigma) - \mathrm{Id})$ denote the space of invariants under $M(\sigma)$, i.e. the set of $v = \sum_{i=1}^n \lambda_i e_i$

such that λ_i depends only on the orbit of i under the action of σ. Thus $\dim V^\sigma$ is equal to the number of orbits under σ. Let then $\sigma = c_1 \cdots c_r$ be the decomposition into cycles with disjoint supports of σ and let us denote, for $k = 1, \cdots, n$, $n_k(\sigma)$ the number of cycles c_i of lengths k. We then have

$$\dim V^\sigma = \sum_{k=1}^{n} n_k(\sigma) = \sum_{k=1}^{n} n_k(\sigma') = \dim V^{\sigma'}.$$

Similarly, as $M(\sigma^r) = M(\sigma)^n$, we also have $\dim V^{\sigma^r} = \dim V^{(\sigma')^r}$ and since, if c is a cycle of length k then c^r is written as the product of $k \wedge r$- cycles with disjoint supports all of the same length $\frac{k}{k \wedge r}$, we deduce that for all r, we have

$$\sum_{k=1}^{n} (k \wedge r) n_k(\sigma) = \sum_{k=1}^{n} (k \wedge r) n_k(\sigma'),$$

which is written as a matrix $SX = SX'$ where $S := (i \wedge j)_{1 \leq i,j \leq n}$ is the matrix of gcd and X (resp. X') is the column matrix of $n_k(\sigma)$ (resp. $n_k(\sigma')$). Let $A = (a_{i,j})_{1 \leq i,j \leq n}$ be the matrix defined by $a_{i,j} = 1$ if j divides i and 0 otherwise. The relation $\sum_{d|m} \varphi(d) = m$ is then written matrixally in the form

$$A \mathrm{diag}(\varphi(1), \cdots, \varphi(n))^t A = S$$

so that the matrix S is invertible and therefore $X = X'$, i.e. σ and σ' have $\ll$ the same $\gg$ decomposition into cycles with disjoint supports and are therefore conjugate.

3.6 Cyclic Endomorphisms

Let E be a finite-dimensional K-vector space.

Definition 103. *An endomorphism $f \in \mathcal{L}(E)$ is said to be cyclic if and only if there exists $v \in E$ such that $E = \{P(f)(v) : \ P \in K[X]\}$.*

Remarque: if n is the dimension of E then $E = \{P(f)(v) : \ P \in K_{n-1}[X]\}$ and since $(1, X, \cdots, X^{n-1})$ is a basis of $K_{n-1}[X]$, we deduce that $\big(v, f(v), \cdots, f^{n-1}(v)\big)$ is a basis of V. In this basis the matrix of f is of the form

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & a_0 \\ 1 & 0 & & \vdots & a_1 \\ 0 & \ddots & \ddots & & \vdots \\ \cdots & & 1 & 0 & a_{n-2} \\ 0 & \cdots & \cdots & 1 & a_{n-1} \end{pmatrix}.$$

We easily check that the characteristic polynomial of this matrix is $P(X) = X^n - a_{n-1}X^{n-1} - \cdots - a_0$ and we say that the previous matrix is *the companion matrix* of $P(X) = X^n - a_{n-1}X^{n-1} - \cdots - a_0$.

Lemme 104. *The minimal polynomial of a cyclic endomorphism is equal to its characteristic polynomial.*

Preuve : Since $v, f(v), \cdots, f^{n-1}(v)$ is a free family, any polynomial Q of degree $\leq n-1$ then verifies $Q(f)(v) \neq 0$, so that the minimal polynomial is of degree $\geq n$. Since moreover it divides the characteristic polynomial, which is of degree n, we deduce that it is equal to it.

Proposition 105. *Let $g \in \mathcal{L}(E)$ be such that $gf = fg$; there then exists $P \in K[X]$ such that $g = P(f)$.*

Remarque: in other words the commutant of a cyclic endomorphism is $\{P(f) : \ P \in \mathbb{K}[X]/(\pi_f)\}$.

Preuve : We write $g(v) = \alpha_0 v + \cdots + \alpha_{n-1} f^{n-1}(v)$ and we set $Q(X) = \alpha_0 + \cdots + \alpha_{n-1} X^{n-1}$. To verify the equality $g = Q(f)$ it suffices to verify it on the basis $(v, f(v), \cdots, f^{n-1}(v))$, i.e.

$$g(f^i(v)) = f^i(g(v)) = f^i(Q(f)(v)) = Q(f)(f^i(v)),$$

hence the result.

3.7 Similarity invariants

Let V be a finite-dimensional K-vector space and $f \in \mathcal{L}(E)$. The new idea is then to consider V equipped with the endomorphism f, as a $K[X]$-module as follows, and to use the structure theorem of finitely generated modules over a principal ring, cf. the §**??**.

Definition 106. *We equip V with a $K[X]$-module structure by setting $X.v := f(v)$ and by linearity for all $P \in K[X]$, we have $P.v = P(f)(v)$. We will denote V_f the V space equipped with this $K[X]$-module structure.*

Proposition 107. *Two endomorphisms f and g are similar if and only if the two structures of $K[X]$-module induced on V are isomorphic, i.e. $V_f \simeq V_g$.*

Preuve : Let $h \in \mathcal{L}(E)$ be such that $g = hfh^{-1}$ then for $v \in V_f$, we have $h(X.v) = hf(v) = g(h(v) = X.h(v)$, i.e. h induces an isomorphism of $K[X]$-modules : $V_f \simeq V_g$.

Conversely if $h : V_f \simeq V_g$ then $h(X.v) = X.h(v)$, i.e. $h(f(v) = g(h(v))$ and therefore $hf = gh$ let $g = hfh^{-1}$ and therefore f and g are similar.

On the general theory of modules over a principal ring, cf. the theorem 57, we deduce the following theorem which must be understood as a decomposition of the space into a direct sum of cyclic subspaces.

Theorem 108. *There exists a unique sequence of non-constant polynomials $P_1(X)|\cdots|P_r(X)$ such that*

$$V_f \simeq K[X]/(P_1) \times \cdots \times K[X]/(P_r). \tag{3}$$

In particular we have $P_r = \pi_f$ and $P_1 \cdots P_r = \chi_f$.

Example : if $V_f \simeq K[X]/(X^n)$ then the matrix of f in the basis of image $(1, X, \cdots, X^{n-1})$ is the Jordan matrix $J_n(0)$. Thus if P_r is totally decomposed, by applying the Chinese lemma, we pass from the decomposition (3) to that of Jordan.

Proposition 109. ***(Dunford decomposition)*** *Let $f \in \mathcal{L}(E)$ such that its characteristic polynomial is separable. Then f is written uniquely in the form $d + n$ where*

— *n is nilpotent and d is semi-simple,*

— *d and n commute.*

Furthermore d and n are polynomials in f.

Remarque: a polynomial is separable if all its irreducible factors are. Moreover, an irreducible polynomial is separable if and only if its roots are simple in its decomposition field which is then a Galois extension of the starting field. Finally, let us recall that in characteristic zero any polynomial is separable.

Remarque: since the characteristic polynomial and the minimal polynomial of f have the same irreducible factors, they are either both separable or not.

Preuve : Let us start with the existence : using (3) and the Chinese lemma, it suffices to treat the case where f is cyclic with $\chi_f = \pi^r$ for π an irreducible polynomial of $K[X]$. Let $L = \mathrm{Dec}_K(\pi)$ be a decomposition field of π over K so that the extension L/K is Galois. We write

$$L[X]/(\pi^r) \simeq L[X]/(X - \lambda_1)^r \times \cdots \times L[X]/(X - \lambda_n)^r,$$

where the λ_i are the roots of π in L, i.e. $\pi(X) = \prod_{i=1}^n (X - \lambda_i)$. On each component $L[X]/(X - \lambda_i)^r$, we set $d_i = \lambda_i \mathrm{Id}$ and $n_i = f - \lambda_i$ which is nilpotent of order r. We then note, according to the Chinese lemma, P *the* monic polynomial of degree $< nr$ such that $P(X) \equiv \lambda_i \mod (X - \lambda_i)^r$ for all $i = 1, \cdots, n$. Let then be $\sigma \in \mathrm{Gal}(L/K)$ which permutes the λ_i : in particular we note that $\sigma(P)$ verifies the same congruences so that $\sigma(P) = P$ and therefore finally $P(X) \in K[X]$ and therefore the decomposition $f = d + n$ is well defined on K where d, and therefore also n, is a polynomial in f.

Finally, consider the uniqueness problem : let $d' + n' = d + n$ with $d'n' = n'd'$ and where d, n is defined as before, i.e. d and n are polynomials in f. Thus d' commutes with d and therefore d, d', and therefore also $d - d'$, are simultaneously diagonalizable in a finite extension. Similarly n' commutes with n and therefore $n - n'$ is nilpotent and semi-simple, since equal to $d' - d$ which imposes that $d' - d = 0 = n - n'$, i.e. $d = d'$ and $n = n'$, hence the result.

Counterexample : consider $K = \mathbb{F}_p((T))$ and the K-vector space $E = K[X]/(X^p - T)$ equipped with the endomorphism f defined by the multiplication by X. Note that $X^p - T$ being irreducible E is a field such that for all $Q \in K[X]$, the endomorphism $Q(f)$ thus equal to the multiplication by

$Q(X)$ is either zero or an isomorphism. Thus if the previous statement were valid, the nilpotent endomorphism $n = Q(f)$ is necessarily zero and therefore f would be semi-simple, i.e. diagonalizable in $E' := K'[X]/(X^p - T)$ where $K' = \mathbb{F}_p((T^{\frac{1}{p}}))$. But in K' we have $X^p - T = (X - T^{\frac{1}{p}})^p$ and therefore f seen as an endomorphism of E' admits a unique eigenvalue : if it were diagonalizable it would therefore be a homothety, which is clearly not the case.

A flaw in the previous proof is that it requires knowledge of the roots of the minimal polynomial π_f of f and is therefore not constructive, i.e. one cannot program a computer to compute the Dunford decomposition based on the previous proof. We now propose, by adapting the classical Newton method, an algorithmic construction of d by noting that it must cancel the polynomial $P(X)$ defined, in zero characteristic, by

$$P(X) := \frac{\pi_f(X)}{\pi_f(X) \wedge \pi_f'(X)}.$$

We thus introduce the sequence $(f_n)_{n\in\mathbb{N}}$ defined by recurrence

$$f_0 = f, \qquad f_{n+1} = f_n - P(f_n)P'(f_n)^{-1}.$$

— Let us first verify that this sequence is well defined, i.e. that $P'(f_n)$ is an invertible matrix. To do this, we reason by recurrence and add the following property to the recurrence hypothesis :

$$P(f_n) = P(f)^{2^n} Q_n(f), \quad Q_n \in K[X],$$

so that $P(f_n)$ is nilpotent since by $P(f)$ is : indeed for r greater than the greatest multiplicity of a root of π_f, the polynomial P^r is divisible by π_f and therefore $P(f)^r$ is zero.

— At rank $n = 0$, by setting $Q_0 = 1$ we have $P(f) = P(f)^{2^0} Q_0(f)$. We then write a Bezout relation between P and P' which by hypothesis are coprime which gives $U(f)P(f) + V(f)P'(f) = \mathrm{Id}$. Since $P(f)$ is nilpotent, we deduce that $\mathrm{Id} - U(f)P(f)$ is invertible and therefore also $P'(f)$.

— Let us therefore assume the result acquired up to rank n, so that f_{n+1} is well defined : we will note in passing that since $P'(f_n)$ is a polynomial in f_n, we have $P(f_n)P'(f_n)^{-1} = P'(f_n)^{-1}P(f_n)$. We then calculate $P(f_{n+1})$ using Taylor's formula :

$$\begin{aligned} P(f_{n+1}) &= P(f_n) + (f_{n+1} - f_n)P'(f_n) + (f_{n+1} - f_n)^2 Q(f_n) \\ &= (f_{n+1} - f_n)Q(f_n) \\ &= P(f_n)^2 \big(P'(f_n)^{-2} Q(f_n) \end{aligned}$$

which is therefore of the form $P(f)^{2^{n+1}} Q_{n+1}(f)$, according to the induction hypothesis and using that f_n is a polynomial in f.

— So for n large enough, $P(f_n)$ is zero and the sequence f_n becomes stationary equal to d which by construction is a polynomial in f and verifies $P(d) = 0$. Thus d is semi-simple. Moreover we have

$$d - f = (f_n - f_{n-1}) + (f_{n-1} - f_{n-2}) + \cdots + (f_1 - f_0)$$

where each of the $f_{i+1} - f_i = -P(f_i)P'(f_i)^{-1}$ is a polynomial in f and nilpotent. Thus these nilpotent endomorphisms commute between them two by two and their sum is therefore nilpotent, hence the result.

3.8 Stable subspaces

A subspace $W \subset V$ is stable by f if and only if W is a sub-$K[X]$-module of V_f.

Lemme 110. *Let f be a cyclic endomorphism. The subspaces stable by f are the $\Im P(f) = \operatorname{Ker} \frac{\chi_f}{P}(f)$ where P describes the divisors of χ_f.*

Remarque: in particular a cyclic endomorphism admits only a finite number of stable subspaces. We refer to the exercise **??** for the converse.

Preuve : Any stable subspace of $V_f \simeq K[X]/(\chi_f)$ are its submodules and therefore to the ideals $(P(X))$ for P a divisor of χ_f, hence the result.

Definition 111. *An endomorphism is said to be* indecomposable *if the space cannot be decomposed into a direct sum of two strict stable subspaces. It is said to be* semi-simple *if any stable subspace admits a stable supplementary.*

Proposition 112. *An endomorphism f is indecomposable if and only if it is cyclic and its characteristic polynomial is the power of an irreducible.*

Preuve : According to the decomposition (3) into cyclic subspaces, the endomorphism must necessarily be cyclic. Moreover, according to the Chinese lemma, the characteristic polynomial must be the power of an irreducible.

Conversely, suppose that f is cyclic with χ_f the power of an irreducible. If we had $V = V_0 \oplus V_1$ then for $W_0 \simeq K[X]/(P_0)$ (resp. $W_1 \simeq K[X]/(P_1)$) a cyclic subspace of V_0, we should have, according to the Chinese theorem $P_0 \wedge P_1 = 1$ and $P_0 P_1$ which divide χ_f which is not.

Proposition 113. *An endomorphism f is semi-simple if and only if it is square-factorless.*

Preuve : Suppose that f is semi-simple and, by contradiction, that $\pi_f = P^2 Q$. The subspace $W = \operatorname{Ker} P(u)$ is stable and therefore admits a stable supplementary W'. Let us then show that $PQ(f)$ is zero on W and W' and is therefore divisible by π_f which will be contradictory. Clearly $QP(f)$ vanishes on W and for $w' \in W'$, we have $P(f) \circ (PQ(f))(w') = 0$ and therefore

$PQ(f)(w') \in W$: but since W' is stable by f, we also have $PQ(f)(w') \in W'$ and therefore since $W \cap W' = \{0\}$, $PQ(f)(w') = 0$.

Now suppose that π_f is irreducible so that V is an E-vector space where $E = K[X]/(\pi_f)$ is a field. A stable subspace of V is then an E-vector space which therefore admits a supplementary E which is a K-vector space stable by f and therefore f is indeed semi-simple.

Finally, let f be such that $\pi_f = \prod_i P_i$ is square-factorless. The decomposition (3) is also written in the form $V = \bigoplus_i V_i$ where $V_i = \operatorname{Ker} P_i(f)$ is an E_i-vector space where $E_i = K[X]/(P_i)$ is a finite extension of K. Let W be a stable subspace of V and let $W_i = W \cap V_i = \operatorname{Ker} P_i(f)_{|W}$ so that by the kernel lemma

$$W = \bigoplus_i W_i.$$

Since $f_{|V_i}$ is semisimple, we denote by W_i' a stable supplementary of W_i in V_i so that $W' := \bigoplus_i W_i'$ is a stable supplementary of W.
Remarque: over an algebraically closed field, semisimple is equivalent to the more classical notion of diagonalizable endomorphism in the sense of the following definition.

Definition 114. *An endomorphism is said to be* diagonalizable *if there exists a basis of eigenvectors, i.e. if there exists a basis in which its matrix is diagonal.*

Remarque: on $\mathbb{R}$ a semi-simple endomorphism is similar to a block diagonal matrix whose blocks are either of dimension 1 or of dimension 2 of the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Theorem 115. *The endomorphism f is diagonalizable if and only if its minimal polynomial is split with simple roots.*

Remarque: for example if the characteristic polynomial is split with simple roots then π_f also; more generally if P is a annihilating polynomial of f split with simple roots then μ_f will be too, since μ_f divides any annihilating polynomial.

We have seen that in the case of a cyclic endomorphism f, the commutant of f is $\{Q(f) : Q \in K[X]\}$. Consider the case $V_f \simeq K[X]/(P) \oplus K[X]/(PQ)$ and the endomorphism $g = A_1(f_1) \oplus A_2(f_2)$ where f_1 (resp. f_2) is the restriction of f to the first factor $K[X]/(P)$ (resp. $K[X]/(PQ)$). We then choose A_2 such that $A_2 \not\equiv A_1 \mod P$. If there exists $B \in K[X]$ such that $g = B(f)$ then $B \equiv A_1 \mod P$ and $B \equiv A_2 \mod PQ$ which is not since $A_2 \not\equiv A_1 \mod P$. We have thus constructed an endomorphism g commuting with f but which is not a polynomial in f. In particular we deduce that if f is such that its commutant is $\{Q(f) : Q \in K[X]\}$ then f is a cyclic endomorphism.

Proposition 116. *With the previous notations, we assume g commutes with f and that any stable subspace F by f is also stable by g. Then g is of the form $Q(f)$ for $Q \in K[X]$.*

Preuve : Consider the decomposition of

$$V_f = K[X]/(P_1) \oplus K[X]/(P_2) \oplus \cdots \oplus K[X]/(P_r)$$

into a cyclic subspace. From the cyclic case, since the restriction g_i of g to each of the factors $V_i := K[X]/(P_i)$ commutes to that f_i of f, we deduce that there exist polynomials $Q_1, \cdots, Q_r$ such that $g_i = Q_i(f_i)$ and it is a question of showing that $Q_r \equiv Q_i \mod P_i$ for all $i = 1, \cdots, r$. Let v_i be a vector generating the cyclic space V_i and let $w_i = v_i + \frac{P_r}{P_i}(f)(v_r)$ so that the cyclic space generated by w_i is isomorphic to $K[X]/(P_i)$. As by hypothesis g stabilizes this cyclic space, its restriction is of the form $Q(f)$, i.e.

$$\begin{aligned} g(w_i) &= Q(f)(w_i) = Q(f)(v_i) + Q(f)\big(\tfrac{P_r}{P_i}(f)(v_r)\big) \\ &= Q_i(f)(v_i) + Q_r(f)\big(\tfrac{P_r}{P_i}(f)(v_r)\big), \end{aligned}$$

and therefore, in $V_i \oplus V_r$,

$$Q \equiv Q_i \mod P_i \text{ and } Q \equiv Q_r \mod P_i,$$

hence the result.

3.9 Exercises

Exercice 1. *Show that if an open set of $\mathbb{M}_n(\mathbb{C})$ contains the diagonal matrices and is stable by similarity, then it is equal to $\mathbb{M}_n(\mathbb{C})$ as an integer.*

Preuve : Let F be the complementary closed set ; if it were nonempty it would contain a matrix $M = S + N$, its Dunford decomposition, and would also contain its semi-simple part S, which is in the adherence of the similarity class of M, hence the contradiction.

Exercice 2. *Show that over an algebraically closed field, two matrices A and B are similar if and only if, for all $\lambda \in K$ and for all $k \geq 0$, we have $\mathrm{rg}(A - \lambda I)^k = \mathrm{rg}(B - \lambda I)^k$.*

Preuve : From the Jordan form, we recall that for $k \geq 1$, $d_k(A) := \dim \mathrm{Ker}(A - \lambda I)^k - \dim \mathrm{Ker}(A - \lambda)^{k-1}$ is equal to the number of Jordan blocks for the eigenvalue λ that are of size larger than k. By the rank theorem, for all $k \geq 1$, we have $d_k(A) = d_k(B)$ so that A and B have the same Jordan reductions and are therefore similar.

Exercice 3. *Show that the eigenvalues of $A = (a_{i,j})_{1 \leq i,j \leq n}$ are in the union of the closed disks centered at $a_{i,i}$ and of radius $\sum_{j \neq i} |a_{i,j}|$ (these are the Gershgorin disks).*

Preuve : The result follows directly from Hadamard's lemma applied to $A - \lambda\mathrm{Id}$. Recall that this lemma says that if for all $1 \leq i \leq n$, we have $|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|$ then A is invertible. Indeed, let X be of coordinates $(x_i)_{1 \leq i \leq n}$ in the kernel of A and let i_0 be such that $|x_{i_0}|$ is maximal among the $|x_i|$. From the equality $a_{i_0,i_0} x_{i_0} = -\sum_{j \neq i_0} a_{i_0,j} x_j$ we deduce the upper bound $|a_{i_0,i_0} x_{i_0}| \leq |x_{i_0} \sum_{j \neq i_0} |a_{i_0,j}|$ and so $x_{i_0} = 0$ or $X = 0$.

Exercice 4. *Let A be a matrix verifying $(A - I)^2(A - 2I) = 0$. Calculate $\exp(A)$ in the form of a polynomial in A.*

Preuve : The kernel lemma allows us to decompose the space $E = \mathrm{Ker}(A - I)^2 \oplus \mathrm{Ker}(A - 2I)$. We consider the projector q (resp. p) on $\mathrm{Ker}(A - I)^2$ (resp. $\mathrm{Ker}(A - 2I)$) parallel to $\mathrm{Ker}(A - 2I)$ (resp. $\mathrm{Ker}(A - I)^2$). From the equality $p + q = \mathrm{Id}$, we obtain $\exp(A) = \exp(A)p + \exp(A)q$. Now $\exp(A)p = e^2 \exp(A - 2I) = e^2 \sum_{n \geq 0} \frac{(A-2I)^n}{n!} \circ p = e^2 p$ because $(A - 2I) \circ p = 0$. Similarly we have $\exp(A)q = eAq$. From Bezout's identity $1 = (X - 1)^2 - X(X - 2)$, we deduce that $p = -A(A - 2I)$ and $q = (A - I)^2$ and therefore

$$\exp(A) = -e^2 A(A - 2I) + eA(A - I)^2$$

Exercice 5. *Show that the set of diagonalizable matrices of $\mathbb{M}_n(\mathbb{C})$ is connected and dense. What is its interior ? Is the latter still connected ?*

Preuve : Connectivity : we have a surjective application $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ on the set of diagonalizable matrices : we send $(P, (a_1, \cdots, a_n))$ on $P\mathrm{diag}(a_1, \cdots, a_n)P^{-1}$. The set $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ being connected, the same is true of the set of diagonalizable matrices.

We recall that $GL_n(\mathbb{C})$ is connected : let P_1, P_2 be two invertible matrices. We consider the polynomial $\det(P_1 z + (1 - z)P_2)$. The complement of the (finite) set of zeros of this polynomial is connected ; we then consider a path that connects 0 to 1 in this complement, which provides a path from P_1 to P_2 in $GL_n(\mathbb{C})$.

Density : let A be a complex matrix that we trigonalize $PAP^{-1} = T$. Let then $\epsilon_1, \cdots, \epsilon_n$ be small such that the $t_{i,i} + \epsilon_i$ are all distinct. The matrix $T + \mathrm{diag}(\epsilon_1, \cdots, \epsilon_n)$ is then diagonalizable because it has n distinct eigenvalues.

Interior : given a diagonalizable matrix A with a multiple eigenvalue ; $P^{-1}AP = \mathrm{diag}(a_1, a_2, \cdots, a_n)$ with $a_1 = a_2$, then $A + PE_{1,2}P^{-1}$ is no longer diagonalizable.

Conversely, if A is diagonalizable with distinct eigenvalues, given that the characteristic polynomial depends continuously on A, and that the roots of a polynomial depend continuously on its coefficients, we deduce that if A' is close to A, it will also have n distinct eigenvalues and will therefore be diagonalizable.

Moreover, the set of matrices with distinct eigenvalues is still connected. Indeed, it is the complement of the zeros of the polynomial in n^2 variable defined as the discriminant of the characteristic polynomial.

Exercice 6. *Characterize matrices that are squares : treat the case of* $\mathbb{C}$ *then of* $\mathbb{R}$*.*

Preuve : Let us already note that the problem comes down to treating the nilpotent case and the invertible case. Indeed M can be written in the form $P\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}P^{-1}$ with A nilpotent and B invertible. If we have $X^2 = M$ then M and X commute so that $X = P\begin{pmatrix} X_1 & X_3 \\ X_4 & X_2 \end{pmatrix}P^{-1}$ with $AX_3 = X_3B$ and $AX_4 = X_4B$; we then deduce that $\chi_A(A)X_3 = X_3\chi_A(B) = 0$. Now χ_A and χ_B are coprime and therefore $\chi_A(B)$ is invertible and therefore $X_3 = 0$. Similarly we have $X_4 = 0$.

The nilpotent case is treated in the questions on the lesson *Nilpotent Endomorphisms*. Let us then treat the invertible case.
Complex case : we reduce as before to the case $A = \lambda I + N$ with N nilpotent. We write $A = \lambda(I + \frac{N}{\lambda})$ which has as square root $\sqrt{\lambda}(I + \frac{N}{\lambda})^{1/2}$ where $(I + \frac{N}{\lambda})^{1/2}$ is defined by the series which is finite because N is nilpotent.
Real case : $A = X^2$ with A and X real. The case of strictly positive eigenvalues is treated as above. As for negative eigenvalues, they can only come from the pure imaginary eigenvalues of X ; the latter being real the Jordan blocks associated with $\lambda \in \mathbb{C}$ are the same as those associated with $\bar{\lambda}$. Moreover for N nilpotent kernel of dimension 1, we have $(\lambda\mathrm{Id} + N)^2 = \lambda^2\mathrm{Id} + N'$ with $N' = 2\lambda N + N^2$ nilpotent kernel of dimension 1 (write N in Jordan form) so that N' is similar to N. Thus $J_n(\lambda)^2$ is similar to $J_n(\lambda^2)$ and we therefore note that the Jordan blocks of A relative to the negative eigenvalues are, for each dimension, even in number.

Then all that remains is to treat the real matrices A without real eigenvalue : A is then similar to a direct sum of matrices of the form $J_n(\lambda) \oplus J_n(\bar{\lambda})$. We then write $J_n(\lambda) = X^2$. The block diagonal matrix $\mathrm{diag}(X, \bar{X})$ is similar to a real matrix : thus A is similar to the square of a matrix itself similar to a real matrix so that A is similar to the square of a real matrix ; it is then classical that we can take the real passage matrix and therefore A is a square.

Exercice 7. *Show that any matrix is effectively similar to a Hessenberg matrix, i.e. such that all the terms below its subdiagonal are zero.*

Preuve : The method is that of the Gauss pivot : we operate, on the left, on the rows without touching the first one so as to obtain a first column whose terms are all zero except possibly the first two. We then apply the same transformation on the right : as on the left we had not modified the

first row, on the right we do not modify the first column. We then reason by recurrence.

Exercice 8. *Show that the application $A \mapsto A\exp(A)$ is surjective on $\mathbb{C}$.*

Preuve :

— The most difficult is the case $n = 1$: it turns out that apart from $z \neq 0$ which has only one antecedent, all the other complexes each have an infinity of them by $f(z) = z\exp z$.

— For a nilpotent matrix : according to Jordan, we reduce to a single full Jordan block J_n. We reason by analysis and synthesis. We then notice that if $J_n = A\exp A$ then A is nilpotent with $\dim \mathrm{Ker}\, A = 1$ which imposes that A is similar to J_n : $A = PJ_nP^{-1}$. In summary, we simply note that, according to Jordan, $J_n + J_n^2 + \frac{J_n^3}{2} + \cdots + \frac{J^{n-1}}{(n-2)!}$ and J_n are similar.

— All that remains is to treat the invertible case where we reduce to $M = \lambda I_n + J_n$ and where we write $\lambda = \mu\exp(\mu)$ with $\mu \neq 0$ and even $\mu \neq -1$. We then have $f(\mu I_n + N) = f(\mu)I_n + f'(\mu)N + N^2p(N)$ where $p(N)$ is a polynomial in N. Since $f'(\mu) \neq 0$, we can then proceed as in the nilpotent case.

Exercice 9. *Let V be a $\mathbf{C}$ finite-dimensional vector space $n \geq 1$ and $u \in \mathrm{End}_{\mathbf{C}}(V)$. We then equip V with its $\mathbf{C}[X]$-module structure associated with u.*

(a) We assume that V has no non-trivial submodules. Show that $n = 1$. We now replace $\mathbf{C}$ by $\mathbf{R}$. Is the statement still true ?

(b) We denote by $P_u = P_1.P_2^2 \cdots P_l^l$ the characteristic polynomial of u, where the P_i are pairwise coprime, without square and unitary factors. Verify that such a writing is possible and is unique.

(c) With the notations of b), we further assume that V is a direct sum of sub-$\mathbf{C}[X]$-modules of dimension 1 (as $\mathbf{C}$-vector spaces). Compute the similarity invariants of u.

(d) Under the hypothesis of c), we are further given an element $v \in \mathrm{End}_{\mathbf{C}}(V)$ such that $v \circ u = u \circ v$. Show that there exists a basis of V where u and v are simultaneously diagonalizable.

Preuve : We first note that u has a unique eigenvalue because otherwise, for λ_1 and λ_2 distinct eigenvalues, $W = \mathrm{Ker}(u - \lambda_1 Id)$ and $W' = \mathrm{Ker}(u - \lambda_2 Id)$ would have an intersection reduced to the zero vector. Let then λ be the unique eigenvalue of u (on $\mathbb{C}$, an endomorphism always has at least one eigenvalue). We then note that $\mathrm{Ker}(u - \lambda Id)$ is of dimension 1, because otherwise for x_1 and x_2 non-collinear eigenvectors, $W = \mathbb{C}x_1$ and $W' = \mathbb{C}x_2$ would have an intersection reduced to the zero vector. We therefore deduce that u admits a unique invariant factor equal to its minimal polynomial and its characteristic polynomial, namely $(X - \lambda)^n$.

Exercice 10. *Let E be a vector space over $\mathbb{R}$ of dimension n, u an endomorphism of minimal polynomial P. We assume that $P = P_1P_2$, P_1 and P_2 being non-constant unitary polynomials coprime. We denote by E_u the space E equipped with the $\mathbb{R}[X]$-module structure defined by the endomorphism u.*

(a) Show that for $i = 1, 2$,

$$E_i = \{x \in E \mid P_i(u)(x)) = 0\}$$

are submodules of E_u.

(b) Show that $E_u = E_1 \oplus E_2$.

(c) Show that P_1 is the minimal polynomial of $u_{|E_1}$.

Preuve : (a) On $\mathbb{C}$ every endomorphism has an eigenvalue and therefore an eigenvector v so that $\mathbb{C}v$ is a stable subspace not reduced to the zero vector so that by hypothesis it is equal to the entire space which is therefore of dimension 1.

On $\mathbb{R}$, the statement is false, it suffices to consider in $\mathbb{R}^2$, a rotation matrix of angle $0 < \theta < \pi$.

(b) We decompose P_u, which by convention is unitary, into products of irreducible factors $Q_1^{\alpha_1} \cdots Q_r^{\alpha_r}$ and we note that P_i is defined as the product of the Q_j such that $\alpha_j = i$.

(c) As a $\mathbb{C}[X]$-module, V is of the form $(\mathbb{C}[X]/(X - \lambda_1))^{\alpha_1} \oplus \cdots \oplus (\mathbb{C}[X]/(X - \lambda_r))^{\alpha_r}$, where the λ_i are the eigenvalues of u and α_i their multiplicity in the characteristic polynomial. With the notations of (b), we have $P_i = \prod_{j/\alpha_j=i}(X - \lambda_j)$. The invariant factors are of the form $\mu_1|\mu_2|\cdots|\mu_l$ where each of the μ_j is of the form $\prod_{i\in I_j}(X-\lambda_i)$ where I_j is a certain subset of $\{1, \cdots, r\}$ such that $I_j \subset I_{j+1}$. Thus the elements of I_1 are repeated l times, those of I_2 are repeated $(l-1)$ times and generally those of I_j are repeated $(l+1-j)$ times. We therefore deduce that $l = \max_i\{\alpha_i\}$ then that I_j is the set of i such that $\alpha_i \geq l + 1 - j$ so that the invariant factors are $P_l, P_lP_{l-1}, P_lP_{l-1}P_{l-2}, \cdots, P_l \cdots P_1$.

Exercice 11. *For $n > 1$, we denote $J_n \in \mathbb{M}_n(\mathbb{C})$ the nilpotent matrix whose coefficients are all zero, except those of the first superdiagonal $j_{i,i+1}$ for $1 \leq i < n$ which are equal to 1. Consider the following matrices, written in blocks :*

(a) $A_1 = \mathrm{diag}(aI_3, bI_2, cI_1)$;

(b) $A_2 = \mathrm{diag}(I_3, I_2 + J_2, I_2 + J_2, I_3 + J_3, I_3 + J_3, 2I_2, 2I_3 + J_3, 3I_2, 3I_2 + J_2)$;

(c)
$$A_3 = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \ddots & \vdots & a_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & a_{n-1} \\ 0 & \cdots & 0 & 0 & a_n \end{pmatrix}$$

Determine in each case :

(i) the similarity invariants ;

(ii) the minimal and characteristic polynomials ;

(iii) the sequence of dimensions of the kernels $\mathrm{Ker}(A_i - \alpha)^k$ *where* α *is an eigenvalue.*

Preuve : We denote $V = \mathbb{C}^n$ the vector space in question, which we equip with the structure of $A = \mathbb{C}[X]$-module defined by the matrix to be studied ; we denote $a_r(X)|\cdots|a_1(X)$, its similarity invariants. The minimal polynomial is then $a_1(X)$ and the characteristic polynomial is the product of the similarity invariants. For any eigenvalue $\alpha \in \mathbb{C}$, we denote $K_\alpha^i = \mathrm{Ker}(u - \alpha Id)^i$ where u is the endomorphism associated with the matrix in question in the canonical basis ; we also denote r_α the index i such that $K_\alpha^{i-1} \neq K_\alpha^i = K_\alpha^j$ for all $j \geq i$; $K_\alpha^{r_\alpha}$ is called the characteristic subspace associated with α. The integer r_α is the multiplicity of α in $a_1(X)$ while its dimension is the multiplicity of α in the product of the a_i. We denote $\delta_\alpha^i = \dim K_\alpha^i - \dim K_\alpha^{i-1}$; starting from the Jordan form it is easy to see that δ_α^i is equal to the number of a_k divisible by $(X-\alpha)^i$. We thus note that the number r of similarity invariants is equal to the maximum of the dimensions of the eigensubspaces.

(a) The A-module V is clearly isomorphic to $(A/(X-a))^3 \times (A/(X-b))^2 \times A/(X-c)$; we then calculate the similarity invariants via the Chinese theorem as in the previous sheet which gives : $(X-a)$, $(X-a)(X-b)$ and $(X-a)(X-b)(X-c)$. The matrix being diagonalizable, the eigensubspaces are the characteristic subspaces, i.e. all the δ_α^i are zero.

(b) Similarly we have

$$V \simeq (A/(X-1))^3 \times (A/(X-1)^2)^2 \times (A/(X-1)^3)^2 \times \\ (A/(X-2))^2 \times A/(X-2)^3 \times (A/(X-3))^2 \times A/(X-3)^2$$

the similarity invariants given as usual by application of the Chinese theorem are then

$$(X-1)^3(X-2)^3(X-3)^2, \quad (X-1)^3(X-2)(X-3), \quad (X-1)^2(X-2)(X-3), \\ (X-1)^2, \quad (X-1), \quad (X-1), \quad (X-1).$$

With the notations introduced above, we have $\delta_1^1 = 7$ (resp. $\delta_2^1 = 3$, resp. $\delta_3^1 = 3$), then $\delta_1^2 = 4$ (resp. $\delta_2^2 = 1$, resp. $\delta_3^2 = 1$), and $\delta_1^3 = 2$ (resp. $\delta_2^3 = 1$, resp. $\delta_3^3 = 0$) all δ_i^k being zero for $k > 3$. We then obtain $\dim K_1^1 = 7$ (resp. $\dim K_2^1 = 3$, resp. $\dim K_3^1 = 3$), $\dim K_1^2 = 11$ (resp. $\dim K_2^2 = 4$, resp. $\dim K_3^2 = 4$) and $\dim K_1^3 = 13$ (resp. $\dim K_2^3 = 5$, resp. $\dim K_3^3 = 4$) with $r_1 = 3$ (resp. $r_2 = 3$, resp. $r_3 = 2$).

(c) The eigensubspace associated with the eigenvalue 0 (resp. 1) is of dimension greater than or equal to 1 (resp. $n-2$) ; in $\mathbb{C}$, the last eigenvalue

is determined via the trace of the matrix which we know is equal to the sum of the eigenvalues counted with multiplicity (indeed any complex matrix is trigonalizable) ; thus we have $1.0 + (n-2).1 + x = n - 2 + a_n$ so that the last eigenvalue is a_n. If $a_n \neq 0, 1$ then the sum of the dimensions of the eigensubspaces associated with the eigenvalues $0, 1, a_n$ is n so that the matrix is diagonalizable and therefore

$$V \simeq A/(X) \times A/(X - a_n) \times (A/(X-1))^{n-2}$$

and the similarity invariants are

$$a_1(X) = (X-1)X(X-a_n), \quad a_2(X) = X - 1, \quad \cdots \quad a_{n-2}(X) = X - 1.$$

If $a_n = a_1 = 0$, we are in the same situation, because the kernel of the matrix is then of dimension 2 because its rank is obviously $n - 2$; the similarity invariants are then

$$a_1(X) = X(X-1), \quad a_2(X) = X(X-1), \quad a_3(X) = \cdots = a_{n-2}(X) = X-1.$$

In the case where $a_n = 0$ and a_1 non-zero, we then have $r_0 = 2$ with $\dim K_0^1 = 1$ so that

$$V \simeq A/(X^2) \times (A/(X-1))^{n-2}$$

or

$$a_1(X) = X^2(X-1), \quad a_2(X) = \cdots = a_{n-2}(X) = (X-1).$$

Exercice 12. *Let u be an endomorphism of $\mathbb{C}^n$, whose eigenvalues are 0 and 1 ; we denote by K_0^i (resp. K_1^i) the kernel of u^i (resp. $(u - \mathrm{Id})^i$) and let d_0^i (resp. d_1^i) its dimension. We assume that the sequence (d_0^i) (resp. (d_1^i)) is equal to $(4, 7, 9, 10, 10, \cdots)$ (resp. $(3, 4, 5, 5, \cdots)$). Then determine the similarity invariants of u.*

Preuve : According to the reminders given in the previous exercise, the number of similarity invariants is equal to the maximum dimension of the proper subspaces, i.e. here 4 similarity invariants a_1, a_2, a_3, a_4. The minimal polynomial is written in the form $X^{\alpha_1}(X-1)^{\beta_1}$ with $\alpha_1 = r_0$ and $\beta_1 = r_1$ where we recall that r_0 (resp. r_1) is the index i such that $K_0^{i-1} \neq K_0^i = K_0^{i+k}$ (resp. $K_1^{i-1} \neq K_1^i = K_1^{i+k}$) for all $k \geq 0$, so here $a_1(X) = X^4(X-1)^3$. Similarly, we write the $a_i(X)$ in the form $a_i(X) = X^{\alpha_i}(X-1)^{\beta_i}$ for $2 \leq i \leq 4$ with $\alpha_i \geq \alpha_{i+1}$ (resp. $\beta_i \geq \beta_{i+1}$) and $\sum_{i=1}^4 \alpha_i = 10$ (resp. $\sum_{i=1}^4 \beta_i = 5$).

As in the previous exercise, we introduce $\delta_0^i = \dim K_0^i - \dim K_0^{i-1}$ (resp. $\delta_1^i = \dim K_1^i - \dim K_1^{i-1}$) ; we recall that δ_0^i (resp. δ_1^i) is the number of similarity invariants divisible by X^i (resp. $(X-1)^i$) (to see it, it suffices to reason on the Jordan form). As for the eigenvalue 0 : we have $\delta_0^4 = 1$ so that $\alpha_2 \leq 3$, furthermore $\delta_0^3 = 2$ imposes $\alpha_2 \geq 3$ or $\alpha_2 = 3$ and $\alpha_3 \leq 2$. Finally $\delta_0^2 = 3$ gives $\alpha_3 = 2$ and $\alpha_4 = 1$.

As for the eigenvalue 1, we find in the same way, $\beta_2 = \beta_3 = \beta_4 = 1$ which gives finally

$$a_1(X) = X^4(X-1)^3, \quad a_2(X) = X^3(X-1), a_3(X) = X^2(X-1), \quad a_4(X) = X(X-1).$$

Exercice 13. *Write in Jordan form and give the sequence of dimensions of the kernels* $\mathrm{Ker}(u-\lambda\mathrm{Id})^i$ *of the endomorphisms* u *whose similarity invariants are :*

(a) $P_1(X) = X$ *;*
(b) $P_1(X) = X(X-1)$ *;*
(c) $P_1(X) = X$ *and* $P_2(X) = X^2$ *;*
(d) $P_1(X) = X$ *and* $P_2(X) = X(X-1)$ *;*
(e) $P_1(X) = X^2(X-1)$, $P_2(X) = X^2(X-1)(X-2)$, $P_3(X) = X^3(X-1)^2(X-2)$ *and* $P_4(X) = X^4(X-1)^3(X-2)^4$ *;*

Preuve : We use the notations from the previous exercises. We recall that the dimension n of the vector space in question is the sum of the degrees of the similarity invariants.

(a) Here $n = 1$ and the endomorphism in question is the identity.

(b) We have $n = 2$ and a cyclic space with two distinct eigenvalues ; u is therefore diagonalizable and its matrix in a diagonalization basis is the diagonal matrix $\mathrm{diag}(0, 1)$.

(c) $n = 3$ and 0 is the only eigenvalue with $\delta_0^1 = 2$ and $\delta_0^2 = 1$ or $\dim K_0^1 = 2$ and $\dim K_0^2 = 3$ and the associated Jordan matrix is $\mathrm{diag}(0, J_2)$.

(d) $n = 3$ and $0, 1$ are the eigenvalues of u with $\delta_0^1 = 2$ (resp. $\delta_1^1 = 1$) and $\delta_1^i = \delta_0^i = 0$ for $i > 1$. We then obtain $\dim K_0^1 = 2$ and $\dim K_1^1 = 1$, the endomorphism is therefore diagonalizable.

(e) $n = 24$, the eigenvalues being $0, 1, 2$; the sequence δ_0^i (resp. δ_1^i, resp. δ_2^i) is $(4, 4, 2, 1, 0, \cdots)$ (resp. $(4, 2, 1, 0, \cdots)$, resp. $(3, 1, 1, 1, 0, \cdots)$) so that the sequence of dimensions of K_0^i (resp K_1^i, resp. K_2^i) is $(4, 8, 10, 11, \cdots)$ (reps. $(4, 6, 7, \cdots)$, resp. $(3, 4, 5, 6, \cdots)$). The Jordan form is the block diagonal matrix

$$\mathrm{diag}(J_2, J_2, J_3, J_4, I_1, I_1, I_2, I_3, 2I_2, 2I_4 + J_4).$$

Exercice 14. *In this problem, we propose to give an algorithm to compute the Dunford decomposition without computing the eigenvalues (which algorithmically can generally only be done approximately).*

Let $n \geq 1$ *and* $A \in \mathcal{M}_m(K)$ *be a square matrix with coefficients in the field* $K \subset \mathbb{C}$*. We denote by* χ_A *the characteristic polynomial of* A*. In* $\mathbb{C}$*,* $\chi_A(X)$ *is decomposed in the form* $\prod_i (X - \lambda_i)^{n_i}$ *with* $\sum_i n_i = m$*. We then introduce the polynomial* $P(X) = \prod_i (X - \lambda_i)$*.*

(a) Show that $P(X) = \lambda \frac{\chi_A(X)}{\chi_A(X) \wedge \chi_A'(X)}$*, where* $\chi_A(X) \wedge \chi_A'(X)$ *denotes the gcd of* χ_A *with its derived polynomial and* $\lambda \in K$*. Deduce then that* $P(X)$ *is a polynomial with coefficients in* K*.*

(b) Let U and N be matrices of $\mathcal{M}_n(K)$ respectively invertible and nilpotent, which commute with each other. Show that $U-N$ is invertible. Show then that $P'(A)$ is an invertible matrix of $\mathcal{M}_m(K)$ whose inverse commutes with A.

(c) We then consider the following sequence : $A_0 := A$ and $A_{n+1} = A_n - P(A_n).(P'(A_n))^{-1}$. We want to show by induction on n that the sequence is well-defined, i.e. that $P'(A_n)$ is an invertible matrix.

(i) Show that for any polynomial $Q \in K[X]$, there exists $\tilde{Q} \in K[X,Y]$ such that $Q(X+Y) = Q(X) + YQ'(X) + Y^2\tilde{Q}(X,Y)$.

(ii) Assuming the sequence A_n defined up to rank n, show that $P(A_n)$ can be written in the form $P(A)^{2^n}.B_n$ where B_n is a matrix of $\mathcal{M}_n(K)$ that is a polynomial in A.

(iii) Using a Taylor formula for the polynomial P', write $P'(A_{n+1})$ as the sum of an invertible matrix $P'(A_n)$ and a nilpotent matrix that commute between them.

(d) Show that for any polynomial $Q \in K[X]$, there exists $\tilde{Q} \in K[X,Y]$ such that $Q(X+Y) = Q(X)+YQ'(X)+Y^2\tilde{Q}(X,Y)$. Then show by induction on n that $P(A_n)$ can be written in the form $P(A)^{2^n}.B_n$ where B_n is a matrix of $\mathcal{M}_n(K)$.

(e) Deduce that the sequence A_n is stationary with limit D with D diagonalizable on $\mathbb{C}$ and $N := A - D$ nilpotent verifying $DN = ND$.

Preuve : (a) We note that the multiplicity of λ_i in P' is equal to $n_i - 1$ so that λ_i is a root of order 1 of $\frac{\mu_A(X)}{\mu_A(X) \wedge \mu'_A(X)}$ and that moreover these are the only roots from which the result comes. We will note in particular that knowledge of the λ_i is not necessary to calculate P which can be calculated via the Euclid algorithm.

(b) - The idea is to use the formal relation $(1-x)(1+x+x^2+\cdots+x^k) = 1 - x^{k+1}$ with $x = U^{-1}N$ and k such that $N^{k+1} = 0$ or $(1 - U^{-1}N)(1 + U^{-1}N + \cdots + (U^{-1}N)^k) = I_n$ because $(U^{-1}N)^{k+1} = U^{-k-1}N^{k+1}$ because U and N commute with each other ; or by multiplying on the left by U and on the right by U^{-1}, $(U - N)(U^{-1} + U^{-2}N + \cdots + U^{-k-1}N^k) = I_n$.

- The eigenvalues of A are not roots of P' and $P \wedge P' = 1$. We then consider a Bezout relation $RP' + SP = 1$ for P and P' which, when applied to A, gives $R(A)P'(A) = 1 - N$ with $N = S(A)P(A)$. However, according to the Cayley-Hamilton theorem, we have $P^r(A) = 0$ for $r \geq \max_i(n_i)$ so that N is nilpotent and therefore, by application of the above, $P'(A)$ is invertible whose inverse commutes with A as a polynomial in A.

(c) This is Newton's method applied to matrices, the goal being to construct a root of P, i.e. to find the diagonalizable part of A in its Dunford decomposition. Note that for $n = 0$, $P'(A_0)$ is invertible according to the previous question.

(i) For example, it suffices to verify this on the monomials X^m, i.e.

$$(X+Y)^m = X^m + mYX^{m-1} + Y^2 \sum_{k=2}^{m} \binom{k}{m} Y^{k-2} X^{m-k}.$$

(ii) It is clear from (a) that for all $0 \leq k \leq n$, A_k is a polynomial in A. We reason by induction on n. For $n = 0$, we have $P(A_0) = P(A)$. Let us therefore assume the result acquired up to rank k. According to (i), we write $P(A_{k+1}) = P(A_k + Y) = P(A_k) + YP'(A_k) + Y^2\tilde{Q}(A_k, Y)$ with Y such that $P(A_k) + YP'(A_k) = 0$. According to (a) $Y = P(A_k)Q(A_k)$ and therefore $P(A_{k+1})$ is of the form $P(A)^{2^{k+1}} B_{k+1}$ for a matrix B_{k+1} which as a polynomial in A_k commutes with A.

(iii) Taylor's formula gives $P'(A_{n+1}) - P'(A_n) = (A_{n+1} - A_n)Q(A_n)$ where $Q \in K[X]$. Now $A_{n+1} - A_n$ is of the form $P(A_n)\tilde{Q}(A_n)$ and is therefore nilpotent and commutes with A_n which is a polynomial in A. We then deduce that $P'(A_{n+1})$ is invertible according to (a).

(e) We recall that $P^r(A) = 0$ for $r = \max_i\{n_i\}$ so that the subsequence $(A_k)_{k\geq n}$ is constant as soon as $2^n \geq r$. The limit D is a polynomial in A such that $P(D) = 0$ so that D is diagonalizable because it has a split annihilating polynomial with simple roots (in $\mathbb{C}$). Moreover, for n such that $2^n \geq r$, we have $A - D = A_0 - A_n = \sum_{i=0}^{n-1}(A_i - A_{i+1})$ with $A_i - A_{i+1}$ nilpotent and which is a polynomial in A. Thus the $A_i - A_{i+1}$ commute in themselves so that their sum is nilpotent hence the result.

Exercice 15. *Let u be an endomorphism of $V \simeq K^n$ for which we denote χ_u and π_u respectively the characteristic and minimal polynomials :*

(1) χ_u is irreducible if and only if V has no stable subspace under u ;

(2) u is cyclic with π_u a power of an irreducible polynomial if and only if V is indecomposable under u ;

(3) propose an algorithm to test whether u is semi-simple.

Preuve : (1) If V has a stable subspace W by u, by completing a basis of W in a basis of V, the matrix of u is block diagonal and its characteristic polynomial is divisible by that of $u_{|W}$. Conversely, if χ_u is of the form PQ with P and Q coprime, the kernel lemma decomposes the space into a direct sum of $\operatorname{Ker} P(u)$ and $\operatorname{Ker} Q(u)$. If $\chi_u = P^r$ with P irreducible, we then have $E = \operatorname{Ker} P$ i.e. $\pi_u = P$. If we take any non-zero x, the vector space generated by $x, u(x), u^2(x), \cdots$ is therefore at most of dimension $\deg P$ (in fact we have equality), and by hypothesis is therefore equal to the entire space, i.e. $r = 1$.

(2) In the direct sense, using the structure of $K[X]$-module on V induced by u, we have $V \simeq K[X]/(\pi_u)$. If V were decomposable it would be as a $K[X]$-module isomorphic to a direct product $K[X]/P_1 \times K[X]/P_2$, which imposes $P_1 = P^r$ and $P_2 = P^s$ with P irreducible and $\pi_u = P^{r+s}$. Now the minimal polynomial of this direct product is clearly $P^{\max(r,s)}$, so $\min(r, s) = 0$.

Conversely, if V is indecomposable, then u is cyclic. Moreover, if its minimal polynomial were not a power of an irreducible polynomial, then the kernel lemma would contradict the indecomposability of V.

(3) u is semisimple if and only if π_u is without multiplicity, i.e. is prime with π'_u. We can test whether u is semisimple algorithmically : we calculate the characteristic polynomial χ_u and we test whether $\frac{\chi_u}{\chi_u \wedge \chi'_u}$ cancels u.

Exercice 16. *What are the complex endomorphisms u that have only a finite number of stable subspaces ?*

Preuve : We obviously place ourselves on an infinite field. It is already necessary that the proper subspaces be of dimension 1 otherwise, there would be an infinity of lines in one of these proper subspaces, which are obviously stable.

Since the proper subspaces are of dimension 1, we deduce that the endomorphism is cyclic (this is true on each characteristic subspace, we then apply the Chinese theorem). The space equipped with the structure of $K[X]$-module induced by u, is then isomorphic to $K[X]/(P(X))$ and the stable subspaces are in bijection with the divisors of P.

Exercice 17. *What are the endomorphisms u such that any stable subspace is of the form* $\operatorname{Ker} P(u)$ *or* $\operatorname{Im} P(u)$ *for P a polynomial.*

Preuve : We can already notice that the $\operatorname{Ker} P(u)$ and $\operatorname{Im} P(u)$ describe a finite set of subspaces : indeed if P is prime with the minimal polynomial of u then $\operatorname{Ker} P(u) = (0)$ and $\operatorname{Im} P(u) = E$. As in the previous question, the proper subspaces must be of dimension 1, and then the space being cyclic, the stable subspaces will be the $\operatorname{Ker} Q(X) = \operatorname{Im} \frac{P(X)}{Q(X)}$.

Exercice 18. *Given a flag and the associated parabolic, show that the stable subspaces by all elements of the parabolic subgroup are those of the flag.*

Preuve : Let E be a stable subspace and let i be such that $V_i \subset E \subsetneq V_{i+1}$. Suppose that $E \neq V_i$ and let $x \in (V_{i+1} \backslash V_i) \cap E$ and let $y \in V_{i+1} \backslash (V_i \cup E)$. There then exists $g \in P_W$ such that $g_{|V_i} = \mathrm{Id}$ and $g(x) = y$. Now since E is stable, we have $y \in E$ hence the contradiction.

Exercice 19. *Give the "list" of all stable subspaces of a vector isometry of* $\mathbb{R}^3$.

Preuve : The identity, the reversals and the reflections are diagonalizable ; we then know the "list" of their stable subspaces. In particular, we cannot distinguish reversals from reflections. As for a generic rotation, it is semi-simple and has only one proper line and therefore also one proper plane (as it is semi-simple, a proper plane has a supplementary stable which is therefore the unique proper line and which is also orthogonal to the stable plane).

Exercice 20. *(1) Let E be a K-vector space and F be a subspace. Show that the set of supplements of F in E is an affine space of direction* $\mathrm{Hom}_K(E/F, F)$.

(2) Let u be an endomorphism of E and let F be a subspace stable under u. Show that the set of supplements of F stable under u is, when it is nonempty, an affine space of direction the vector subspace of $\mathrm{Hom}_K(E/F, F)$ *of the s such that $s \circ \bar{u} - u_{|F} \circ s = 0$.*

(3) Let F be a subspace stable under M ; we assume that there exists a supplementary G stabke $E = F \oplus G$. Give a CNS so that G is the unique stable supplementary of F.

Preuve : (1) We identify the supplementaries of F in E with the sections s of the canonical surjection $\pi : E \longrightarrow E/F$. The affine space of supplementaries of F in E is then the set of solutions of the linear equation with right-hand side $\pi \circ s = \mathrm{Id}_{E/F}$. The direction of this affine space is therefore the set of s such that $\pi \circ s = 0$, i.e. the set of $s' : E/F \longrightarrow F$.

(2) Such a supplementary will then be stable if and only if $s \circ \bar{u} - u \circ s = 0$. The direction is then the intersection of the two vector spaces : $\pi \circ s = 0$ and $s \circ \bar{u} - u \circ s = 0$, that is, the subspace of $\mathrm{Hom}_K(E/F, F)$ of the s such that $s \circ \bar{u} - u_{|F} \circ s = 0$.

(3) According to the previous question, the direction of the affine space of stable supplements of F is that of the vector space of rectangular matrices $X \in \mathbb{M}_{p,n-p}(K)$ such that $AX - XB = 0$ with $p = \dim F$. We will show that a necessary and sufficient condition is that the characteristic polynomials of A and B are relatively prime.

Let μ_A and μ_B be the respective minimal polynomials of A and B and let Q be an irreducible factor of their gcd. According to reduction theory, we can decompose F (resp. G) into a stable subspace by A (resp. B) in the form $F_A \oplus F'_A$ (resp. $G_B \oplus G'_B$), such that F_A (resp. G_B) is cyclic relative to A (resp. B) with characteristic polynomial Q.

Let $X : G \longrightarrow F$ be defined as follows : it induces an isomorphism of G_B onto F_A and is zero on G'_B. We then have $AX = XB$.

Suppose that μ_A and μ_B are coprime, which is equivalent to the fact that their characteristic polynomials are. By linearity, we have $\mu_A(A)X = X\mu_A(B) = 0$ according to Cayley-Hamilton. Now, since μ_A is coprime with μ_B, a BÃ©zout relation $P_A\mu_A + P_B\mu_B = 1$ gives that $\mu_A(B)$ is invertible with inverse $P_A(B)$ so that X is zero.

Exercice 21. *Give the commutative subgroups of exponent r of $GL_n(\mathbb{C})$. Deduce that $GL_n(\mathbb{C}) \simeq GL_m(\mathbb{C})$ if and only if $n = m$.*

Preuve : Let G be a commutative subgroup of $GL_n(\mathbb{C})$ with exponent r. For all $g \in G$, we have $g^r = \mathrm{Id}$ i.e. $X^r - 1$ is a split annihilating polynomial with simple roots of g so that g is diagonalizable in a basis $(e_1, \cdots, e_n)$. For all $g' \in G$, g and g' commute and are diagonalizable, so we deduce that they

are simultaneously diagonalizable. Thus the matrix of any $g \in G$ in the basis $(e_1, \cdots, e_n)$, is diagonal of the form $\mathrm{diag}(\xi_1, \cdots, \xi_n)$ where ξ_i is an r-th root of unity. We therefore deduce that $G \simeq (\mathbb{Z}/r\mathbb{Z})^n$.

If the two groups $GL_n(\mathbb{C})$ and $GL_m(\mathbb{C})$ are isomorphic, their commutative subgroups of exponent r correspond to $(\mathbb{Z}/r\mathbb{Z})^n \simeq (\mathbb{Z}/r\mathbb{Z})^m$ and therefore $n = m$ by cardinality.

Exercice 22. *Show that A of $GL_n(\mathbb{C})$ is diagonalizable if and only if there exists k such that A^k is.*

Preuve : The direct meaning is obvious. In the other direction, we reason in each of the characteristic spaces so that we reduce to a unique eigenvalue λ. We write A in the form $\lambda(I_n + N)$ with N nilpotent. We then have $(I_n + N)^k = I_n + kN + \cdots$. Now we notice that $kN + \cdots$ is nilpotent and similar to N (use Jordan), hence the result.

Exercice 23. *Propose an effective test to know if a complex matrix is diagonalizable with distinct eigenvalues. Then treat the case of real matrices.*

Preuve : A matrix of $\mathbb{M}_n(K)$ is diagonalizable with distinct eigenvalues if and only if its minimal polynomial is of degree n and is split with simple roots. On $\mathbb{C}$, it is then sufficient to test whether the characteristic polynomial χ has simple roots. To do this, it is sufficient to check that a gcd of χ and χ' is equal to 1.

On $\mathbb{R}$, it is also necessary to test whether χ is split. To do this, we have the Sturm sequences which give us the number of real roots of χ.

Exercice 24. *Let A and B be two simultaneously diagonalizable matrices. Is there a polynomial P such that $B = P(A)$? Show that there exists a matrix C as well as polynomials P_A, P_B such that $A = P_A(C)$ and $B = P_B(C)$.*

Preuve : If we take $A = I_n$ then $P(A) = P(1)I_n$ so that if B is diagonalizable without being scalar, there cannot exist such a P.

The problem in the previous question came from multiple roots. Let P be the transition matrix from the canonical basis to the common diagonalization basis of A and B : $PAP^{-1} = \mathrm{diag}(a_1, \cdots, a_n)$ and $PBP^{-1} = \mathrm{diag}(b_1, \cdots, b_n)$. Let us then take C such that $PCP^{-1} = \mathrm{diag}(c_1, \cdots, c_n)$ with the c_i distinct two by two, for example $c_i = i$. We then choose P_A (resp. P_B) such that for all i, $P_A(c_i) = a_i$ (resp. $P_B(c_i) = b_i$) : this is possible by using for example the Lagrange interpolation polynomials. We therefore have $A = P_A(C)$ and $B = P_B(C)$.

Exercice 25. *Let a be a non-diagonalizable endomorphism, find the polynomials P such that $P(a)$ is diagonalizable (treat the case of $\mathbb{C}$ then of $\mathbb{R}$).*

Preuve : On $\mathbb{C}$: we denote by $\mu_a(X) = \prod_\lambda (X - \lambda)^{r_\lambda}$ the minimal polynomial of a. For any eigenvalue λ, we place ourselves on the associated

characteristic subspace $E_{(\lambda)}$. If $P(a)$ is diagonalizable then its restriction to $E_{(\lambda)}$ is $P(\lambda)\mathrm{Id}$. Now we have $P(a) - P(\lambda)\mathrm{Id} = (a - \lambda\mathrm{Id})Q_\lambda(a)$; for the latter to be zero it is necessary that $Q_\lambda(a)(E_{(\lambda)}) \subset E_\lambda$ where E_λ is the eigensubspace. Thus it is necessary that $(X-\lambda)^{r_\lambda - 1}$ divides $Q_\lambda(X)$ which is equivalent, for $r_\lambda > 1$ to

$$P'(\lambda) = P"(\lambda) = \cdots = P^{(r_\lambda - 1)}(\lambda) = 0$$

Conversely if this last condition is verified for any eigenvalue λ then $P(a)$ is diagonalizable.

On $\mathbb{R}$: we place ourselves in $\mathbb{C}$ so that the previous condition must be verified for any eigenvalue λ. In addition it is necessary to ensure that the eigenvalues $P(\lambda)$ are real. Conversely if these two conditions are verified then the matrix of a in the canonical basis is similar on $\mathbb{C}$ to a real diagonal matrix. It is well known that it is similar to it on $\mathbb{R}$.

Exercice 26. *Give the similarity invariants of a diagonalizable endomorphism.*

Preuve : We write the characteristic polynomial of a, $\chi(X) = \prod_\lambda (X-\lambda)^{r_\lambda}$. Since this is supposed to be diagonalizable, its minimal polynomial is $\mu(X) = \prod_\lambda (X - \lambda)$. We recall that the similarity invariants are polynomials

$$\mu_1 | \mu_2| \cdots \mu_r$$

with $\mu = \mu_r$ and $\chi = \prod_i \mu_i$. We thus deduce that all similarity invariants are multiplicity-free, that $r = \max_\lambda(r_\lambda)$ and that

$$\mu_i(X) = \prod_{\lambda \ / \ r_\lambda \geq r-i+1} (X - \lambda)$$

Exercice 27. *Is the matrix* $A = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}$ *diagonalizable ? Give its eigenvalues.*

Preuve : The matrix A is real symmetric and therefore diagonalizable. Moreover, its rank is clearly equal to 2 so that 0 is an eigenvalue of order $n-2$; it then remains to find two other eigenvalues λ_1 and λ_2. The trace gives us $\lambda_1 + \lambda_2 = 2$ while the trace of A^2 gives $\lambda_1^2 + \lambda_2^2 = 2n + n - 2$ which gives λ_1 and λ_2.

Exercice 28. *Are the following matrices squares in* $\mathbb{M}_2(\mathbb{R})$ *:* $\begin{pmatrix} -1 & 0 \\ 0 & -4 \end{pmatrix}$, $\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$.

Preuve : Let $A = \begin{pmatrix} -1 & 0 \\ 0 & -4 \end{pmatrix}$. Suppose that there exists B real such that $A = B^2$. The polynomial $(X^2+1)(X^2+4)$ is then a polynomial annihilating B so that its minimal polynomial, which is of degree 1 or 2, must be X^2+1 or X^2+4 (which are irreducible on $\mathbb{R}$). But then we would have $A + \mathrm{Id} = 0$ or $A + 4\mathrm{Id} = 0$ which is not the case.

Let $A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$. Suppose that there exists B such that $A = B^2$. The polynomial $(X^2+1)^2$ is then a annihilating polynomial of B so that its minimal polynomial of B is X^2+1 and therefore $A + \mathrm{Id} = 0$ which is not.

Exercice 29. *Show that if an open set of $\mathbb{M}_n(\mathbb{C})$ contains the diagonal matrices and is stable by similarity, then it is equal to $\mathbb{M}_n(\mathbb{C})$ as an entire set.*

Preuve : Let F be the complementary closed set ; if it were nonempty it would contain a matrix $M = S + N$, its Dunford decomposition, and would also contain its semisimple part S, which is in the adherence of the similarity class of M, hence the contradiction.

Exercice 30. *(1) On $\mathbb{R}$ or $\mathbb{C}$, show that*

(i) the interior of the set $\mathcal{D}$ of diagonalizable matrices is the set $\mathcal{D}^1$ of matrices with n distinct eigenvalues ;

(ii) the closure of $\mathcal{D}$ is the set $\mathcal{T}$ of trigonalizable matrices.

(2) On $\mathbb{C}$, show the connectedness of $\mathcal{D}$ and $\mathcal{D}^1$.

Preuve : (1) (i) Let A be a matrix with n distinct eigenvalues such that its characteristic polynomial is split with simple roots. From the continuity of the characteristic polynomial and that of the roots of a polynomial, we deduce that for any A' close to A, the characteristic polynomial of A' has, on $\mathbb{C}$, n distinct roots. If A and A' are real, then their roots are also real : for A it is true by hypothesis, for A', its roots are complex conjugate and close to those of χ_A, we conclude by noting that the latter are simple.

Thus $\mathcal{D}^1$ is open, moreover any diagonalizable matrix is a limit of matrices of $\mathcal{D}^1$: indeed for all $a_1, \cdots, a_n$ and for all $\epsilon > 0$, there exists $\epsilon_1, \cdots, \epsilon_n$ such that $0 < \epsilon_i < \epsilon$ and the $a_i + \epsilon_i$ are distinct two by two. Thus for $A = P\mathrm{diag}(a_1, \cdots, a_n)P^{-1}$, the ball with center A and radius ϵ contains $A' = P\mathrm{diag}(a_1 + \epsilon_1, \cdots, a_n + \epsilon_n) \in \mathcal{D}^1$.

We then deduce that the closure of $\mathcal{D}^1$ contains $\mathcal{D}$ and therefore that $\mathcal{D}$ is the interior of $\mathcal{D}$.

(ii) Let $(A_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{D}$ that converges to A. According to the continuity of the characteristic polynomial, χ_k converges to χ. The χ_k being split, we deduce as above that χ is split : its complex roots are limits of the roots of χ_k, if these are all real, their limits too. We then recall that A is trigonalizable.

Conversely, if A is trigonalizable $PAP^{-1} = T$, then $\epsilon_1, \cdots, \epsilon_n$ are small such that the $t_{i,i} + \epsilon_i$ are all distinct. The matrix $T + \mathrm{diag}(\epsilon_1, \cdots, \epsilon_n)$ is then diagonalizable because it has n distinct eigenvalues. Thus A is in the adherence of $\mathcal{D}$.

(2) For $\mathcal{D}$: we have a surjective application $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ on the set of diagonalizable matrices : we send $(P, (a_1, \cdots, a_n))$ on $P\mathrm{diag}(a_1, \cdots, a_n)P^{-1}$. The set $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ being connected, the same is true for the set of diagonalizable matrices.

We recall that $GL_n(\mathbb{C})$ is connected : let P_1, P_2 be two invertible matrices. We consider the polynomial $\det(P_1 z + (1-z)P_2)$. The complement of the (finite) set of zeros of this polynomial is connected ; we then consider a path that connects 0 to 1 in this complement, which provides a path from P_1 to P_2 in $GL_n(\mathbb{C})$.

For $\mathcal{D}^1$: we note that $\mathcal{D}^1$ is the complement of the zeros of the polynomial in n^2 variable defined as the discriminant of the characteristic polynomial.

Exercice 31. *Describe the closure of the orbit of a Jordan block*

$$J_n = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & 0 & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}$$

Preuve : We will show that this closure is the set of nilpotent matrices. Recall that according to the Jordan decomposition, any nilpotent matrix is similar to a block diagonal matrix, with on the diagonal Jordan blocks of distinct sizes $J(n_1, \cdots, n_r) := \mathrm{diag}(J_{n_1}, \cdots, J_{n_r})$ with $\sum_i n_i = n$. We then notice that J_n is similar to

$$\begin{pmatrix} J_{n_1} & \epsilon E_{n_1,n_2} & 0 & \cdots & 0 \\ 0 & J_{n_2} & \epsilon E_{n_2,n_3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & 0 & J_{n_{r-1}} & \epsilon E_{n_{r-1},n_r} \\ 0 & \cdots & \cdots & 0 & J_{n_r} \end{pmatrix}$$

where $E_{i,j}$ is the matrix of size $i \times j$ whose coefficients are all zero except that of the corner at bottom left which is worth 1, and where $\epsilon > 0$. By making ϵ tend towards zero we deduce that $J(n_1, \cdots, n_r)$ is in the adherence of the orbit of J_n.

Exercice 32. *(a) Show that if* $\mathrm{Ker}\, A^2 = \mathrm{Ker}\, A$*, then there exists a pseudo-inverse* X*, i.e. such that* $AX = XA$*,* $AXA = A$ *and* $XAX = X$*.*

(b) Under what condition on the similarity invariants of A *do we have* $\dim \mathrm{Ker}\, A^2 = 2 \dim \mathrm{Ker}\, A$ *?*

Preuve : (a) We decompose the space into a characteristic subspace for A. On the characteristic spaces associated with the non-zero eigenvalues, we set $X = A^{-1}$. On the characteristic space associated with the eigenvalue 0, the hypothesis implies that A is zero there, so we take any X.

(b) This corresponds to saying that the first two columns of the Young tableau associated with A are of the same length, which is equivalent to asking that there is no Jordan block of size 1.

Exercice 33. *Consider the sequence of dimensions of nested kernels. Describe the set of sequences obtained.*

Preuve : Let us denote for $k \geq 0$, $a_k := \dim \operatorname{Ker} a^k$. This is an increasing sequence bounded above by the dimension of the space n. We have $a_0 = 0$ and if $a_1 = 0$ then for all k, $a_k = 0$. More generally, let r be the first index such that $a_r = a_{r+1}$. Let then be $x \in \operatorname{Ker} a^{r+2}$ so that $a(x) \in \operatorname{Ker} a^{r+1} = \operatorname{Ker} a^r$ and therefore $a^{r+1}(x) = 0$ or $x \in \operatorname{Ker} a^{r+1}$ and therefore $a_{r+1} = a_r$ and by recurrence $a_r = a_{r+k}$ for all $k \geq 0$.

We introduce the sequence $d_k := a_k - a_{k-1}$ for all $k \geq 1$. Note that this sequence is decreasing : indeed a induces an injective endomorphism of $\operatorname{Ker} a^k / \operatorname{Ker} a^{k-1}$ into $\operatorname{Ker} a^{k-1} / \operatorname{Ker} a^{k-2}$.

Conversely, let $(a_k)_{k\geq 0}$ be an increasing sequence bounded above by n such that the sequence of differences d_k is decreasing. We then consider the nilpotent matrix A in Jordan form whose number of Jordan blocks of size r is equal to $d_r - d_{r+1}$. We then easily verify that $a_k = \dim \operatorname{Ker} A^k$.

Moreover, a graphical way to represent the similarity classes of nilpotent matrices is to introduce the Young diagram whose columns are the d_i for $i = 1, \cdots, r$. The Jordan blocks are then read on the lines.

Exercice 34. *Show that a sub-$\mathbb{R}$-vector space of the nilpotent cone is of dimension less than $\frac{n(n-1)}{2}$ and that this maximum is reached.*

Preuve : Consider the quadratic form q defined on the space of matrices that associates $\operatorname{tr} X^2$ to X. Obviously the isotropic cone is made up of isotropic vectors so that the vector space in question will be totally isotropic.

Moreover, if $X \neq 0$ is symmetric (resp. antisymmetric) then $q(X) > 0$ (resp. $q(X) < 0$) so that the signature of q is $(\frac{n(n+1)}{2}, \frac{n(n-1)}{2})$. Thus a totally isotropic subspace has a dimension less than or equal to $\frac{n(n-1)}{2}$.

The equality is clearly reached for strictly upper triangular matrices.

Exercice 35. *Let u be a nilpotent endomorphism ; for all $i \geq 0$, we denote K_0^i as the kernel of u^i and d_0^i as its dimension. We assume that the sequence (d_0^i) is equal to $(4, 7, 9, 10, 10, \cdots)$. Then determine the similarity invariants of u.*

Preuve : We denote by $V = K^n$ the vector space in question, which we equip with the structure of $A = K[X]$-module defined by the matrix to

be studied ; we denote by $a_r(X)|\cdots|a_1(X)$, its similarity invariants. The minimal polynomial is then $a_1(X)$ and the characteristic polynomial is the product of the similarity invariants. We denote by r the index i such that $K_0^{i-1} \neq K_0^i = K_0^j$ for all $j \geq i$. The integer r is the multiplicity of 0 in $a_1(X)$ while its dimension is the multiplicity of 0 in the product of the a_i. We note $\delta^i = \dim K_0^i - \dim K_0^{i-1}$; starting from the Jordan form it is easy to see that δ_0^i is equal to the number of a_k divisible by X^i. We thus note that the number r of similarity invariants is equal to the maximum of the dimensions of the eigensubspaces.

Thus the number of similarity invariants is equal to the dimension of the kernel, i.e. 4 similarity invariants a_1, a_2, a_3, a_4. The minimal polynomial is written in the form X^{α_1} with $\alpha_1 = r_0$ where r_0 is the index i such that $K_0^{i-1} \neq K_0^i = K_0^{i+k}$ for all $k \geq 0$, so here $a_1(X) = X^4$. Similarly, we write the $a_i(X)$ in the form $a_i(X) = X^{\alpha_i}$ for $2 \leq i \leq 4$ with $\alpha_i \geq \alpha_{i+1}$ and $\sum_{i=1}^4 \alpha_i = 10$.

We introduce as before $\delta_0^i = \dim K_0^i - \dim K_0^{i-1}$; δ_0^i is the number of similarity invariants divisible by X^i. From $\delta_0^4 = 1$ we deduce $\alpha_2 \leq 3$; furthermore $\delta_0^3 = 2$ imposes $\alpha_2 \geq 3$ or $\alpha_2 = 3$ and $\alpha_3 \leq 2$. Finally $\delta_0^2 = 3$ gives $\alpha_3 = 2$ and $\alpha_4 = 1$.

Exercice 36. *Under what conditions on the similarity invariants of nilpotent A does the equation $X^2 = A$ have solutions ?*

Preuve : We reason by analysis and synthesis. Let X be nilpotent and written in Jordan form : x_k is the number of Jordan blocks of size k. The Jordan decomposition of J_k^2 includes two blocks $J_{\lfloor \frac{k}{2} \rfloor}$ and $J_{\lceil \frac{k}{2} \rceil}$.

We then deduce that the Young tableau associated with X^2 satisfies one of the following equivalent conditions :

— it does not contain two consecutive columns of the same odd length ;

— if we group the rows two by two starting from the top (starting from the convention that we have a last row of zero length in the case where the kernel is of odd dimension), then the rows of the same pair differ by at most one cell.

The synthesis is then obvious.

Exercice 37. *Give, as a function of the similarity invariants, the dimension of the commutant of a nilpotent endomorphism.*

Preuve : This involves determining the number of degrees of freedom in the choice of an operator M that commutes with A. We reason in a Jordanization basis of A. We recall that we have

$$\operatorname{Ker} A \subsetneq \operatorname{Ker} A^2 \subsetneq \cdots \subsetneq \operatorname{Ker} A^r = \operatorname{Ker} A^{r+1}$$

We consider a basis $e_n, \cdots, e_{n-d_r+1}$ of $\operatorname{Ker} A^r - \operatorname{Ker} A^{r-1}$ of cardinality the length d_r of the last column of the Young tableau associated with A.

The image of this basis is totally free which gives $d_r n$ degrees of freedom ; in return the image of the u^k of these vectors are fixed. Let then r_1 be maximal such that $d_{r_1} \neq d_r$; we then obtain $d_{r_1} \dim \mathrm{Ker}\, A^{r_1}$ new degrees of freedom. We proceed in this way until exhausting all the space.

We then easily verify that we obtain a number of degrees of freedom equal to the sum of the squares of the lengths of the columns of the Young tableau.

Exercice 38. *(a) Show that the vector subspace generated by the nilpotent cone is the hyperplane of the trace-zero matrices.*

(b) Show that the vector subspace generated by the nilpotent matrices of rank 1 *is the same hyperplane.*

(c) Deduce that the vector subspace generated by the matrices of any similarity class of nilpotent matrices is the hyperplane of the trace-zero matrices.

Preuve : In all three cases, inclusion is immediate. We will show (b) directly. As usual this relies on a small calculation in dimension 2, namely : $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is similar to $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ by considering the new basis $e_1 + e_2$ and $e_1 - e_2$.

Let A then be a zero trace matrix ; by adding a linear combination of nilpotent matrices of rank 1, we reduce to A diagonal $\mathrm{diag}(a_1, \cdots, a_b)$ with $\sum_i a_i = 0$ which we write in the form

$$\mathrm{diag}(a_1, -a_1, 0, \cdots, 0) + \mathrm{diag}(0, a_2 + a_1, a_3, \cdots, a_n).$$

According to the previous calculation the first matrix is similar to a linear combination of nilpotent matrices of rank 1 ; the second also by induction hypothesis.

(c) The orbit of any similarity class contains in its adherence the similarity class of nilpotent matrices of rank 1. We then conclude from (b).

Exercice 39. *Show that any hyperplane H of $\mathbb{M}(n, \mathbb{C})$ contains at least $n^2 - n - 1$ linearly independent nilpotent matrices.*

Preuve : We reduce to the case where H has equation $tr(TX) = 0$ with triangular T and we consider the intersections of H with the subspaces of the upper or lower triangular nilpotent matrices.

Exercice 40. *Let $M = S + N$ be the Dunford decomposition of M into a semisimple plus nilpotent. Show that S is in the closure of the similarity class of M.*

Preuve : This simply follows from the fact that 0 is in the closure of the similarity class of N.

4 Bilinear Algebra

4.1 Sesquilinear Forms : Generalities

Recall that given an automorphism σ of the field $\mathbb{K}$, for example the complex conjugation of $\mathbb{C}$, *a semi-linear application* is an application θ such that for all $x, y \in E$ and $\lambda \in \mathbb{K}$ we have

$$\theta(x + \lambda y) = \theta(x) + \lambda^\sigma \theta(y)$$

where by convention we denote λ^σ for $\sigma(\lambda)$.

Definition 117. *We call* σ-sesquilinear form *any application* $\phi : E \times E \to \mathbb{K}$ *verifying the following conditions :*
— *for all* $x \in E$*, the application* $\phi_x : y \in E \mapsto \phi(x, y)$ *is linear ;*
— *for all* $y \in E$ *the application* $\phi_y : x \in E \mapsto \phi(x, y)$ *is* σ*-linear.*

Remarque: the notations ϕ_x and ϕ_y are not exemplary, we will be careful not to mix them.

Proposition 118. *Let* E *be a vector space with a basis* $(e_i)_{1 \leq i \leq n}$*. We denote by* $A_\phi = \big(\phi(e_i, e_j)\big) \in \mathbb{M}_n(\mathbb{K})$ *the matrix associated with the sesquilinear form* ϕ*, relative to the basis* $(e_i)_i$*. For all vectors* $x, y \in E$ *with column vector coordinates* X *and* Y*, we then have*

$$\phi(x, y) = {}^t X^\sigma A_\phi Y.$$

Remarque: if $P_{(e_i) \leftarrow (e'_i)}$ is the matrix of change of basis from $(e_i)_i$ to $(e'_i)_i$ then the matrix A'_ϕ relative to this new basis is such that

$$A'_\phi = {}^t P^\sigma_{(e_i) \leftarrow (e'_i)_i} A_\phi P_{(e_i) \leftarrow (e'_i)_i}.$$

In particular the determinant of A_ϕ, which we call the discriminant of ϕ, is defined only as an element of $\mathbb{K}/N(\mathbb{K})$ where $N(\mathbb{K}) = \{\lambda\lambda^\sigma,\ \lambda \in \mathbb{K}\}$.

Definition 119. *For* M *a subset of* E*, we denote*

$$M^\perp = \{y \in E,\ \phi(M, y) = 0\}, \qquad {}^\perp M = \{x \in E,\ \phi(x, M) = 0\}.$$

We say that $M^\perp$ *(resp.* ${}^\perp M$*) is the right orthogonal (resp. left orthogonal) of* M*.*

Remarque: $M^\perp$ and ${}^\perp M$ are clearly subspaces of E.

Lemme 120. *If* M *is a subspace of* E *then*

$$\dim M + \dim M^\perp = \dim E + \dim(M \cap {}^\perp E).$$

Preuve : Let $f : E \longrightarrow E^*$ be the semi-linear map which to x associates $y \mapsto \phi(x, y)$. By definition $M^\perp$ is orthogonal in the usual sense, cf. the proposition **??**, of $f(M)$ so that

$$\dim M^\perp = \dim E - \dim f(M) = \dim E - (\dim M - \dim(M \cap {}^\perp E))$$

since ${}^\perp E = \operatorname{Ker} f$.
Remarque: if we apply the formula to $M = E$, we obtain $\dim E^\perp = \dim {}^\perp E$.

Definition 121. *We say that ϕ is* nondegenerate *if $E^\perp = \{0\}$ (resp. ${}^\perp E = \{0\}$. The rank of A_ϕ is called* the rank *of ϕ, it is equal to the codimension of $E^\perp$ and ${}^\perp E$.*

Remarque: when ϕ is nondegenerate, the application $x \mapsto \phi_x$ induces a semi-linear *canonical* isomorphism from E to its dual E^*.

Lemme 122. *We have ${}^\perp(M^\perp) = M + {}^\perp E$.*

Remarque: In particular if ϕ is non-degenerate, we find the usual property of biduality ${}^\perp(M^\perp) = M$.
Preuve : We clearly have $M + {}^\perp E \subset {}^\perp(M^\perp)$ so that it suffices to show the equality of dimensions. Reasoning as in the previous lemma, we have

$$\dim N + \dim {}^\perp N = \dim E + \dim(N \cap E^\perp).$$

We apply the formula to $N = M^\perp$ so that by noting that $E^\perp \subset M^\perp$, we obtain

$$\dim {}^\perp(M^\perp) = \dim E - \dim M^\perp + \dim E^\perp = \dim M + \dim {}^\perp E - \dim(M \cap {}^\perp E)$$

because $\dim E^\perp = \dim {}^\perp E$ according to the above, and we recognize Grassman's formula which gives the dimension of $M + {}^\perp E$.

Definition 123. *A σ-sesquilinear form is called* reflexive *if for all $x, y \in E$, $\phi(x, y) = 0$ is equivalent to $\phi(y, x) = 0$.. It is called Hermitian (resp. antiHermitian) if $\phi(x, y) = \epsilon\Big(\phi(y, x)\Big)^\sigma$ with $\epsilon = 1$ (resp. $\epsilon = -1$).*

Remarque: for a Hermitian or antiHermitian form, σ is necessarily an involution ; in the antihermitian case in characteristic different from 2, we even have $\sigma = \mathrm{Id}$ and we simply say that ϕ is *anti-symmetric*.

Proposition 124. *We assume that ϕ is a non-degenerate hermitian or antihermitian form. For all $u \in \mathcal{L}(E)$, there exists a unique endomorphism $u^* \in \mathcal{L}(E)$, called* adjoint *of u, such that for all $x, y \in E$:*

$$\phi(u(x), y) = \phi(x, u^*(y)).$$

Furthermore we also have $\phi(x, u(y)) = \phi(u^*(x), y)$.

Remarque: the data of a non-degenerate σ-sesquilinear form, induces a canonical semi-linear isomorphism between E and E^* so that the usual adjoint of an endomorphism seen in $\mathcal{L}(E^*)$ is seen as an endomorphism of E. The equality $\phi(u(x), y) = \phi(x, u^*(y))$ of the proposition is then a simple translation of the isomorphism between $\mathcal{L}(E^*)$ and $\mathcal{L}(E)$ induced by ϕ.

Preuve : Consider for y fixed, the linear form $x \mapsto \phi(y, u(x))$; since ϕ is nondegenerate, there exists a vector depending on y that we denote by $u^*(y)$ such that for all x we have $\phi(u^*(y), x) = \phi(y, u(x))$ and therefore also, using the Hermitian nature of ϕ, $phi(u(x), y) = \phi(x, u^*(y))$. Finally we easily verify that $y \mapsto u^*(y)$ is linear.

We now assume that ϕ is a Hermitian or anti-Hermitian form, in which case the characteristic is also assumed to be different from 2.

Definitions 125.
- *A vector x of E is said to be* isotropic *if $\phi(x, x) = 0$.*
- *A subspace F of E is said to be* isotropic *if $F \cap F^\perp \neq \{0\}$.*
- *A subspace F of E is said to be* totally isotropic *and we write* SETI*, if $F \subset F^\perp$.*
- *A seti is said to be maximal and we write* SETIM*, if for any SETI G containing F then $G = F$.*

Remarque: since we are in finite dimension, any SETI is contained in a SETIM.
Remarque: if F is non-isotropic then $E = F \oplus F^\perp$.

Proposition 126. *All SETIMs have the same dimension called* index of ϕ.

Preuve : Let U and V be two SETIMs and introduce M and N which are respectively supplements of $U \cap V$ in U and V. We assume by absurdity $\dim U > \dim V$ so that $r := \dim M > s := \dim N$. For $f_1, \cdots, f_s$ a basis of N, consider the application $M \longrightarrow \mathbb{K}^s$ defined by

$$m \mapsto \big(\phi(f_1, m), \cdots, \phi(f_s, m)\big).$$

Its kernel is $M \cap N^\perp$ so that by the rank theorem

$$\dim(M \cap N^\perp) = \dim M - \dim \Im \geq r - s > 0.$$

Let us then consider a nonzero vector $x \in M \cap N^\perp \subset U$. Let us show that $x \in V^\perp$: so let $v \in V$ be decomposed by $v = u + n$. We then have $\phi(x, v) = \phi(x, u) + \phi(x, n)$ with $\phi(x, u) = 0$ because U is a SETI and $\phi(x, n) = 0$ because $x \in N^\perp$. Let us then consider $W = V + \mathbb{K}x$: for all $w = v + \lambda x$ we have

$$\phi(w, w) = \phi(v, v) + \lambda^2 \phi(x, x) + \bar{\lambda}\phi(v, x) + \lambda\phi(x, v),$$

where $\phi(v, v) = 0$ (resp. $\phi(x, x) = 0$) because V (resp. U) is a SETIM, and $\phi(v, x) = \phi(x, v) = 0$ because $x \in V^\perp$. Thus W is a SETI strictly containing V while the latter was assumed to be maximal, hence the contradiction.

Proposition 127. *If ϕ is non-degenerate, then there exists a so-called Witt decomposition of the space $E = F \oplus F' \oplus G$ with F, F' setims and G a non-isotropic subspace such that the matrix of ϕ in an adapted basis is of the form*

$$\begin{pmatrix} 0 & I_r & 0 \\ \epsilon I_r & 0 & 0 \\ 0 & 0 & B \end{pmatrix}.$$

Remarque: we will give a proof of this result later in the case where $\sigma = \mathrm{Id}$, i.e. for quadratic forms.

4.2 Remarkable endomorphisms

We now assume that E is equipped with a non-degenerate Hermitian form in characteristic different from 2.

Definition 128. *An automorphism u of E is called* unitary *if for all $x, y \in E$:*

$$\phi(u(x), u(y)) = \phi(x, y).$$

The set of unitary automorphisms is a subgroup of $GL(E)$ denoted U_ϕ and called the unitary group *of ϕ. The kernel of the determinant morphism with image $\mathbb{H} = \{\lambda \in \mathbb{K} : \ \lambda\lambda^\sigma = 1\}$, is denoted SU_ϕ and is called* the special unitary group *of ϕ.*

Remarque: u is unitary if and only if $u^{-1} = u^*$. Matrixically, the matrix U of u is unitary if and only if ${}^tU^\sigma A_\phi U = A_\phi$.

In the case where $A_\phi = I_n$, we find the usual condition ${}^tU^\sigma U = I_n$.

Definition 129. *A similarity of ratio $\lambda \in \mathbb{K}^\times$ is an automorphism such that for all $x, y \in E$:*

$$\phi\big(u(x), u(y)\big) = \lambda\phi(x, y).$$

The group of similarities is denoted by GU_ϕ.

Remarque: a classical definition of a similarity consists in asking :

$$\phi(x, y) = 0 \Rightarrow \phi\big(u(x), u(y)\big) = 0.$$

Definition 130. *An endomorphism u is said to be* self-adjoint *if it verifies $u = u^*$.*

Remarque: in the real (resp. complex) case we also say *symmetric* (resp. *Hermitian*).

Definition 131. *An endomorphism u of a Hermitian space is said to be* normal *if $u \circ u^* = u^* \circ u$.*

Theorem 132. *A normal endomorphism is semi-simple.*

Preuve : We will show more precisely that if F is stable by u normal then $F^\perp$ is also u-stable. Consider an orthonormal basis of F that we complete into an orthonormal basis of $F^\perp$. The matrix of u in this basis is of the form $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ and we calculate

$$\begin{aligned} MM^* &= \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}\begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix} = \begin{pmatrix} AA^*+BB^* & BC^* \\ CB^* & CC^* \end{pmatrix} \\ M^*M &= \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix}\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B+C^*C \end{pmatrix} \end{aligned}$$

which provides the equalities

$$\begin{cases} AA^*+BB^* = A^*A & BC^* = A^*B \\ CC^* = B^*B + C^*C. \end{cases}$$

Since $\mathrm{tr}(AA^*) = \mathrm{tr}(A^*A)$ we deduce that $\mathrm{tr}(BB^*) = 0$ and therefore $B = 0$ hence the result
Remarque: for normal f, we have $||f(x)|| = ||f^*(x)||$ for all x and therefore $\mathrm{Ker}\, f = \mathrm{Ker}\, f^*$. More generally if λ is an eigenvalue of f, using that $f - \lambda\mathrm{Id}$ is also normal, we deduce that the eigensubspace $\mathrm{Ker}(f - \lambda\mathrm{Id})$ is equal to the eigensubspace $\mathrm{Ker}(f^* - \bar{\lambda}\mathrm{Id})$. Their common orthogonal is therefore stabilized by both f and f^*. So in the case where the field is algebraically closed, we can use these arguments to prove that f and f^* are simultaneously diagonalizable.

4.3 Quadratic forms

We take the definitions of the previous paragraph in the case where $\sigma = \mathrm{Id}$, we then speak of symmetric bilinear form ϕ and $q(x) := \phi(x, x)$ is *a quadratic form.* In characteristic different from two, the polarization formula

$$\phi(x, y) = \frac{1}{4}\big(q(x + y) - q(x - y)\big)$$

allows us to identify symmetric bilinear forms and quadratic forms. We keep the vocabulary of the previous paragraph with the notions of isotropy, of non-degenerate form allowing us to canonically identify the space E with its dual E^*. The matrix translation of $\phi(x, y)$ is ${}^tXAY$ and the change of basis is expressed as $A' = {}^tPAP$.

Definition 133. *Given two symmetric bilinear forms (V, ϕ) and (V', ϕ'), the orthogonal sum $(V, \phi) \perp (V', \phi')$ is the symmetric bilinear form $\phi \perp \phi'$ defined on $V \oplus V'$ by the formula*

$$(\phi \perp \phi')(x + x', y + y') = \phi(x, y) + \phi'(x', y'), \qquad \forall x, y \in V,\ \forall x', y' \in V'.$$

Remarque: if V is equipped with a symmetric bilinear form ϕ and two subspaces U, W such that $V = U \oplus W$ and $\phi(u, w) = 0$ for all $u \in U$ and $w \in W$, then

$$(V, \phi) \simeq (U, \phi_{|U}) \perp (W, \phi_{|W}).$$

In particular we have $(V, \phi) \simeq (V^\perp, 0) \perp (V/V^\perp, \phi')$ with ϕ' non degenerate.

Lemme 134. *Let (V, ϕ) be a symmetric bilinear form and W be a subspace of V such that $b_{|W}$ is non-degenerate. We then have*

$$(V, \phi) = (W, \phi_{|W}) \perp (W^\perp, b_{|W^\perp}).$$

Preuve : Let us start by showing that $V = W \oplus W^\perp$. Since ϕ_W is non-degenerate, we have $W \cap W^\perp = \{0\}$. Let then $v \in V$ and $f_v \in W^*$ be defined by

$$f_v : w \in W \mapsto \phi(w, v).$$

Since $\phi_{|W}$ is non-degenerate, the application $w \in W \mapsto f_w \in W^*$ is an isomorphism and there therefore exists $w \in W$ such that $f_v = f_w$ and therefore $v - w \in W^\perp$ hence the assertion. The isomorphism of the statement then follows from the previous remark.

Notation 3. *For $a_1, \cdots, a_n \in K^\times$, we denote by $\langle a_1, \cdots, a_n \rangle$ the symmetric bilinear form on K^n defined by*

$$(x, y) \in K^n \mapsto \sum_{i=1}^{n} a_i x_i y_i \in K.$$

Remarque: the matrix of $\langle a_1, \cdots, a_n \rangle$ in the canonical basis is the diagonal matrix $\mathrm{diag}(a_1, \cdots, a_n)$. Since the a_i are non-zero, this matrix is invertible and $\langle a_1, \cdots, a_n \rangle$ is non-degenerate.

Theorem 135. *Any symmetric bilinear form is equivalent to a form $\langle a_1, \cdots, a_r \rangle$ for some $a_i \in K^\times$ and where r is the rank of the form.*

Preuve : We reason by induction on the dimension of the space; the case of a line being obvious. The bilinear form ϕ being non-zero, from the polarization formula, we deduce the existence of a vector v such that $q(v) \neq 0$ so that the restriction of ϕ to $W = K.v$ is non-degenerate. According to the previous lemma, we have $(V, \phi) = (W, \phi_{|W}) \perp (W^\perp, \phi_{W^\perp})$ and we conclude by induction.
Remarque: the statement means that there exists a basis $(e_1, \cdots, e_n)$ such that $q(e_i) = a_i$ where for $i > r$ we have set $a_i = 0$. Let $l_1, \cdots, l_n$ be the associated dual basis, we then have

$$q(x) = a_1 l_1(x)^2 + \cdots + a_r l_r(x)^2.$$

The Gauss decomposition algorithm allows us to directly find the independent linear forms $l_1, \cdots, l_r$. Note further that changing e_i to λe_i modifies a_i by $\lambda^2 a_i$, so that the a_i are a priori well-defined only in $K^\times/K^{\times,2}$. However, they are not necessarily invariants in the sense that $\langle a_1, \cdots, a_n\rangle$ can be isomorphic to $\langle a'_1, \cdots, a'_n\rangle$ even, modulo permutations, the $a_i a'_i \notin K^{\times,2}$. For

— $K = \mathbb{C}$, since any non-zero element is a square, $\phi \simeq \langle 1, \cdots, 1\rangle$ where the number of 1 is equal to the rank of ϕ, which is indeed an invariant.

— For $K = \mathbb{R}$, we have $\phi \simeq \langle 1, \cdots, 1, -1, \cdots, -1\rangle$. We denote by r (resp. s) the number 1 (resp. of -1); the pair (r, s) is called the signature of ϕ. To show that these are invariants, it suffices for example to note that r (resp. s) is the maximum dimension of a subspace on which the restriction of ϕ is positive definite (resp. negative).

— Over a finite field $K = \mathbb{F}_q$, in dimension n we have two classes of non-degenerate symmetric bilinear forms, namely $\langle 1, \cdots, 1\rangle$ and $\langle 1, \cdots, 1, \alpha\rangle$ where α is not a square of $\mathbb{F}_q^\times$. To demonstrate this we reason by induction on n, the case $n = 1$ being obvious. Let then be $n \geq 2$; for all $\alpha, \beta \in \mathbb{F}_q^\times$, the sets $\{\alpha x^2 : \ x \in \mathbb{F}_q\}$ and $\{1 - \beta y^2 : \ y \in \mathbb{F}_q\}$ are of cardinality $\frac{q+1}{2}$ and cannot be disjoint. Thus the equation $\alpha x^2 + \beta y^2 = 1$ admits a solution, i.e. there exists v such that $q(v) = 1$. We then apply the induction hypothesis to $(Kv)^\perp$.

Definition 136. *Given a vector v such that $q(v) \neq 0$, the reflection τ_v with respect to $(K.v)^\perp$ is defined by*

$$\tau_v(x) = x - 2\frac{\phi(x, v)}{q(v)}v.$$

For$x = x_v + x' \in (Kv) \oplus (Kv)^\perp$, we have $\tau_v(x) = -x_v + x'$ and τ_v is an isometry with respect to q, i.e. $\phi(x, y) = \phi(\tau_v(x), \tau_v(y))$.

Lemme 137. *Let ϕ be a symmetric bilinear form on a K-vector space V and let $x, y \in V$ be such that $q(x) = q(y) \neq 0$. Then there exists an isometry τ of V for ϕ such that $\tau(x) = y$.*

Preuve : Let $u = \frac{x+y}{2}$ and $v = \frac{x-y}{2}$ with thus $x = u + v$ and $y = u - v$. Since $q(x) = q(y)$, we calculate $\phi(u, v) = 0$ and $0 \neq q(x) = q(u) + q(v)$ so that, even if we exchange the roles of x and y, we can assume that $q(v) \neq 0$. The reflection τ_v of the previous definition then verifies $\tau_v(x) = y$ hence the result.

Theorem 138. ***(de Witt)*** *Let (V_1, ϕ_1), (V_2, ϕ_2) and (V, ϕ) be three spaces equipped with a symmetric bilinear form with non-degenerate ϕ. Then*

$$(V_1, \phi_1) \perp (V, \phi) \simeq (V_2, \phi_2) \perp (V, \phi) \Leftrightarrow (V_1, \phi_1) \simeq (V_2, \phi_2).$$

Preuve : Since ϕ is non-degenerate, it is diagonalizable and it is therefore sufficient to treat the case where $(V,\phi) = (Kv, \langle a\rangle)$. For $i = 1, 2$, we denote $v_i = (0, v) \in V_i \perp V$ and $v_2' = f(v_1)$ where $f : (V_1, \phi_1) \perp (Kv, \langle a\rangle)) \simeq (V_2, \phi_2) \perp (Kv, \langle a\rangle))$. We then apply the previous lemma to v_2 and v_2' with $\tau(v_2') = v_2$ so that $\tau \circ f$ sends v_1 to v_2 and thus induces an isomorphism between their orthogonals i.e. between (V_1, ϕ_1) and (V_2, ϕ_2) hence the result.

Proposition 139. *Let V be a K-vector space of dimension 2 equipped with a nondegenerate ϕ symmetric bilinear form. The following properties are equivalent :*

(i) ϕ is isotropic ;

(ii) $\det\phi = -1 \in K^\times/K^{\times,2}$;

(iii) $\phi \simeq \langle 1, -1\rangle$;

(iv) there exists a basis in which the matrix of ϕ is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Preuve : We will follow the chain of implications (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Rightarrow$ (iv) $\Rightarrow$ (iii). The only ones that are not obvious are (ii) $\Rightarrow$ (i) $\Rightarrow$ (iv).

- Let us show (ii) $\Rightarrow$ (i) : we have $\phi\langle\alpha, \beta\rangle$ with $\alpha, \beta \in K^\times$ and $\alpha\beta = -1 \in K^\times/K^{\times,2}$. We therefore have $\phi \simeq \langle\alpha, -\alpha\rangle$ which is indeed isotropic since $q(e_1 + e_2) = 0$.

- Let us show (i) $\Rightarrow$ (iv) : let $x \neq 0$ be such that $q(x) = 0$. Since ϕ is non-degenerate, there exists y such that $\phi(x, y) \neq 0$ and $\phi(x, x + \lambda y) = \lambda\phi(x, y)$. We can therefore choose y such that $\phi(x, y) = 1$ with therefore necessarily y non-collinear to x. We denote $\alpha = q(y)$ and we set $z = y - \frac{\alpha}{2}x$ so that $q(z) = 0$ and $\phi(x, z) = 1$. In the basis (x, z), the matrix of ϕ is that of (iv).

Definition 140. *A form that satisfies one of the equivalent properties of the previous proposition is called* hyperbolic plane *: we denote it by H. A basis such that the matrix of ϕ is as in (iv) is called hyperbolic. Finally, we call* hyperbolic form *an orthogonal sum of hyperbolic planes.*

Lemme 141. *Let ϕ be a nondegenerate symmetric bilinear form on a K-vector space V.*

The following assertions are equivalent :

(i) the form ϕ is hyperbolic ;

(ii) there exists a subspace W of V with $2\dim W = \dim V$ with $\phi_{|W} = 0$;

(iii) there exists a subspace W of V such that $W^\perp W$.

Preuve : - (i) $\Rightarrow$ (ii) : let $(e_1, \cdots, e_{2n})$ be the basis of V in which $\phi \simeq \langle 1, -1, \cdots, 1, -1\rangle$. The vector subspace W generated by $e_{2i-1} + e_{2i}$ for $i = 1, \cdots, n$ then verifies (ii).

- Implication (ii) $\Rightarrow$ (iii) let us verify (iii) $\Rightarrow$ (i). Let $(e_1, \cdots, e_n)$ be a basis of W. For all $i = 2, \cdots, n$, the subspace $(Ke_1)^\perp$ is of dimension $2n - 1$ and contains W : according to the rank theorem there then exists y orthogonal to $e_2, \cdots, e_n$ and not belonging to W. According to the previous proposition,

the plane H generated by e_1 and y is then hyperbolic and $W \cap H^\perp$ verifies (iii) relatively to $H^\perp$. We then conclude by induction on the dimension.
Remarque: in particular $(V,\phi) \perp (V,-\phi)$ is a hyperbolic form since the subspace $W = \{(v,v) : v \in V\}$ verifies point (ii) of the previous lemma.

Lemme 142. *Let ϕ be a hyperbolic form and ψ a non-degenerate form. Then $\phi \otimes \psi$ is hyperbolic.*

Preuve : By hypothesis $\phi = \langle 1,-1\rangle \perp \cdots \perp \langle 1,-1\rangle$ so that

$$\phi \otimes \psi \simeq (\psi \perp -\psi) \perp \cdots \perp (\psi \perp -\psi)$$

and the result follows from the previous remark.

Proposition 143. ***(Witt decomposition)*** *Let ϕ be an isotropic nondegenerate form. There then exists an anisotropic nondegenerate form (V_a,ϕ_a) and a hyperbolic form (V_h,ϕ_h) such that*

$$(V,\phi) \simeq (V_h,\phi_h) \perp (V_a,\phi_a).$$

Such a decomposition is also unique up to isomorphism.

Remarque: in other words, to study a nondegenerate symmetric bilinear form, we reduce to anisotropic forms.
Preuve : If ϕ is anisotropic, there is nothing to do. Otherwise let $x \neq 0$ such that $q(x) = 0$ then, cf. the proof of the previous proposition, y such that $\phi(x,y) = 1$. Since (x,y) is necessarily free, they generate a hyperbolic plane W and the restriction ϕ' of ϕ to $W^\perp$ is non-degenerate which allows to iterate the construction until an anisotropic form is obtained.

The fact that the decomposition is unique up to isomorphism follows from Witt's theorem 138 and from the fact that a hyperbolic form is of the form $H \perp \cdots \perp H$.

Definition 144. *The number of factors H in*

$$(V_h,\phi_h) \simeq H \perp \cdots \perp H$$

is called the Witt index *of ϕ.*

Notation 4. *Let $\phi \simeq \langle a_1,\cdots,a_n\rangle$ and $\phi' \simeq \langle a'_1,\cdots,a'_m\rangle$ be nondegenerate symmetric bilinear forms. We denote $\phi \otimes \phi'$ the non-degenerate symmetric bilinear form isomorphic to*

$$\langle a_1a'_1,\cdots,a_1a'_m,a_2a'_1,\cdots,a_na'_m\rangle.$$

Remarque: the above notation is an artifice to avoid having to canonically define the tensor product of two vector spaces.

Definition 145. *We denote $\mathcal{M}_K$ the set of isomorphism classes of non-degenerate symmetric bilinear forms on K and we equip it with the equivalence relation*

$$\phi \sim \phi' \Leftrightarrow \phi_a \simeq \phi'_a.$$

We denote $W(K)$ the quotient set.

Remarque: we easily verify that $W(K)$ equipped with the orthogonal direct sum $\perp$ and the tensor product $\otimes$ is a commutative ring.
Examples :

— in $\mathbb{C}$, -1 is a square and therefore $\langle 1, \cdots, 1\rangle \simeq \langle 1, -1, 1 \cdots\rangle$ which is a hyperbolic form if and only if the dimension of the space is even. Thus $W(\mathbb{Z}) \simeq \mathbb{Z}/2\mathbb{Z}$ where the arrow is given by the dimension of the space modulo 2.

— The same reasoning in the case $K = \mathbb{R}$, shows that $W(\mathbb{R}) \simeq \mathbb{Z}$ where the arrow is given by $r - s$ where (r, s) is the signature.

— For $K = \mathbb{F}_p$ where $p \equiv 1 \mod 4$ as -1 is a square, we reduce ourselves as before to the following forms

$$\langle 1\rangle,\ \langle \alpha\rangle,\ \langle 1, \alpha\rangle,\ H$$

where $\alpha \in \mathbb{F}_p^\times - \mathbb{F}_p^{\times,2}$. These four forms are distinct in $W(\mathbb{F}_p)$ and we easily verify that they are all of order 2

$$\langle 1\rangle \perp \langle 1\rangle = H \sim 0, \quad \langle \alpha\rangle \perp \langle \alpha\rangle \simeq 0, \quad \langle 1, \alpha\rangle \perp \langle 1, \alpha\rangle = \langle 1, 1\rangle \perp \langle \alpha, \alpha\rangle \sim 0.$$

Thus $W(\mathbb{F}_p) \simeq (\mathbb{Z}/2\mathbb{Z})^2$.

— Let us finish with the $K = \mathbb{F}_p$ with $p \equiv 3 \mod 4$, we then obtain the forms

$$\langle 1\rangle,\ \langle -1\rangle,\ \langle 1, 1\rangle,\ H = \langle 1 - 1\rangle$$

which are distinct in $W(\mathbb{F}_p)$. Now $\langle 1\rangle \perp \langle 1\rangle = \langle 1, 1\rangle \not\sim H$ and therefore $W(\mathbb{F}_p) \simeq \mathbb{Z}/4\mathbb{Z}$.

4.4 The real case

In this case we necessarily have $\sigma = \mathrm{Id}$ and we therefore return to the quadratic forms of the previous paragraph with the notions of positivity.

Definition 146. *A symmetric bilinear form is said to be :*

— *positive (resp. negative) if for all $x \in E$, we have $\phi(x, x) \geq 0$ (resp. $\phi(x, x) \leq 0$);*

— *positive definite (resp. negative definite) if it is positive (resp. negative) and that $\phi(x, x) = 0$ if and only if x is the zero vector.*

Theorem 147. ***(Sylvester's law of inertia)***
Let ϕ be a symmetric bilinear form.

(i) There then exists a decomposition

$$E = E^{\perp} \oplus E^{+} \oplus E^{-}$$

such that the restriction of ϕ to E^+ (resp. E^-) is positive definite (resp. negative definite). Such a decomposition is not unique but the dimensions r of E^+ and s of E^- are the same for any such decomposition and are respectively equal to the maximum of the dimensions of the subspaces F of E such that the restriction of ϕ is positive definite (resp. negative). We say that the pair (r, s) is the signature *of ϕ.*

(ii) The rank of ϕ is equal to $s + r$ and its index is equal to $(n - \mathrm{rg}\phi) + \min\{s, r\}$.

Preuve : (i) The existence of the decomposition was seen above in a general framework, it then remains to verify the characterization of r and s. Let us therefore consider a space F on which q is positive definite and suppose by the absurd that its dimension is $> r$. According to the rank theorem, it then intersects non-trivially $E^{\perp} \oplus E^{-}$ space on which is $q(x) \leq 0$ hence the contradiction. Thus r is indeed equal to the maximum dimension of a space on which q is positive definite.

(ii) The rank of ϕ is clearly equal to $r + s$. Recall that its index is the dimension of a SETIM. Let $(e_1^+, \cdots, e_r^+)$ (resp. $(e_1^-, \cdots, e_s^-)$) be a basis of E^+ (resp. E^-) as well as $(e_1^0, \cdots, e_t^0)$ be a basis of $E^{\perp}$. Let us then consider

$$F = \mathrm{Vect}(e_1^+ + e_1^-, \cdots, e_{\min(r,s)}^+ + e_{\min(r,s)}^-, e_1^0, \cdots, e_t^0).$$

We easily verify that F is totally isotropic of dimension $(n - \mathrm{rg}\phi) + \min(r, s)$.

Let us then suppose by absurdity that there exists a SETI F of dimension $> (n - \mathrm{rg}\phi) + \min(r, s)$. According to the rank theorem it then intersects non-trivially E^+ if $r > s$ and E^- otherwise while for any non-zero x of E^+ (resp. E^-) we have $q(x) > 0$ (resp. $q(x) < 0$), hence the contradiction.
Application : any sub-$\mathbb{R}$-vector space of the nilpotent cone, i.e. of the set of nilpotent matrices of $\mathbb{M}_n(\mathbb{R})$, has a dimension less than $\frac{n(n-1)}{2}$. Indeed, we consider the quadratic form q defined on the space of matrices that to X associates $\mathrm{tr}X^2$. Obviously, the isotropic cone is made up of isotropic vectors so that the vector space in question will be totally isotropic. Moreover, if $X \neq 0$ is symmetric (resp. antisymmetric) then $q(X) > 0$ (resp. $q(X) < 0$) so that the signature of q is $(\frac{n(n+1)}{2}, \frac{n(n-1)}{2})$. Thus a totally isotropic subspace is of dimension less than or equal to $\frac{n(n-1)}{2}$. The equality is clearly reached for strictly upper triangular matrices.

Definition 148. *Let F be a non-isotropic subspace of E ; the unique unitary involution u such that $F = \mathrm{Im}(u + \mathrm{Id})$ is called* the orthogonal symmetry with respect to the non-isotropic subspace F. *If F is a hyperplane, we say that u is* a reflection *; if F is of codimension 2 we say that u is* a reversal *or* a reversal.

Remarque: if v is a normed vector orthogonal to a non-isotropic hyperplane H, then the reflection with respect to H is the application $x \mapsto x - 2\phi(x,v)v$.

Theorem 149. ***(Cartan-DieudonnÃ©)***
Let q be a non-degenerate quadratic form on E.

(i) Any element u of $O(q)$ can be written as a product of $p := \dim \Im(u - \mathrm{Id})$ reflections; moreover any writing of u as a composite of reflections requires at least p different reflections.

(ii) In dimension ≥ 3, any element of $SO(q)$ is can be written as a product of q reversals with $q \leq \dim E$.

Remarque: In dimension 2, we have a group isomorphism $\mathbb{R}/2\pi\mathbb{Z} \to SO(2)$ which to θ associates the matrix

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

We then notice that it is not possible to exhibit a particular generating subfamily.

Preuve : (i) First note that if r_H denotes the orthogonal reflection relative to the hyperplane H, then $u = r_{H_1} \circ \cdots r_{H_r}$ verifies $H_1 \cap \cdots \cap H_r \subset \mathrm{Ker}(u - \mathrm{Id})$ so that $\dim \Im(u - \mathrm{Id}) = n - \dim \mathrm{Ker}(u - \mathrm{Id}) \leq r$ and therefore any writing of u as a composite of reflections requires at least $\dim \Im(u - \mathrm{Id})$ factors.

Let us then show by induction on $p = \dim \Im(u - \mathrm{Id})$, that there exists a writing $u = r_1 \circ \cdots r_q$ with $q \leq p$. For $p = 0$ it is clear since then $u = \mathrm{Id}$. Let us assume the result acquired up to $p - 1$. Let us then consider $x \in \mathrm{Ker}(u - \mathrm{Id})^\perp$ such that $u(x) \neq x$ and let r_0 be the orthogonal reflection relative to the hyperplane H orthogonal to $u(x) - x$. Since $u \in O(q)$, we have $u(x) \in \mathrm{Ker}(u - \mathrm{Id})^\perp$ and therefore $\mathrm{Ker}(u - \mathrm{Id}) \subset H$ and therefore $\mathrm{Ker}(u - \mathrm{Id}) + \mathrm{Vect}(x) \subset \mathrm{Ker}(r_0 \circ u - \mathrm{Id})$. Thus $\dim \Im(r_0 \circ u - \mathrm{Id}) \leq p - 1$ and according to the induction hypothesis, there exists a writing $r_0 \circ u = r_1 \circ \cdots \circ r_q$ with $q \leq p - 1$ and the result is deduced by composing on the left by r_0.

(ii) It suffices to note that for r an orthogonal reflection relative to a hyperplane H, then $-r$ is a reversal. Thus for $u = (r_1 \circ r_2) \circ \cdots \circ (r_{2k-1} \circ r_{2k})$, where we will note that the number of factors is necessarily even for $u \in SO(q)$, by writing $r_{2i-1} \circ r_{2i} = (-r_{2i-1}) \circ (-r_{2i})$, we make u appear as a composite of reversals.

In dimension 3 any matrix of $SO(3)$ is similar to a matrix of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

We also have a group isomorphism between $SO(3)$ and the group of quaternions of norm 1 quotiented by $\{\pm 1\}$. This description is particularly useful when it comes to composing rotations of space.

Proposition 150. *An endomorphism u is normal if and only if in an orthonormal basis it admits a block diagonal matrix of size* 1 *or* 2 *of the form* $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Preuve : Since on $\mathbb{R}$ an irreducible polynomial is of degree at most 2, it follows from (3) and using that the submodules of $\mathbb{R}[X]/(P)$ correspond to the divisors of P, that u admits either a line or a stable plane. If we have a stable line and since according to 132 a normal endomorphism is semi-simple, we conclude by induction on the dimension. If we have a stable design and since the restriction of u to a stable subspace is still normal, for $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the equality $MM^* = M^*M$ gives

$$\begin{cases} b^2 = c^2 \\ ac + bd = ab + dc \end{cases}$$

and therefore

— let $c = b$ in which case the matrix is symmetric with a split characteristic polynomial, i.e. u admits a stable line ;

— let $c = -d$ and $a = d$ and we find the matrix of the statement.

We conclude again by induction on the dimension using the semi-simplicity of u.

Conversely, it suffices to check that the matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ is normal, which has already been done above.

Corollory 151. *An endomorphism u of E is*

— *symmetric if and only if it is diagonalizable in an orthonormal basis.*

— *anti-symmetric if and only if, in an orthonormal basis, its matrix is diagonal by zero blocks or of the form* $\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$.

— *orthogonal if and only if, in an orthonormal basis, its matrix is block-diagonal where the blocks are either* ± 1 *or rotation matrices* $\begin{pmatrix} cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Remarque: we can use this reduction theorem to show, for example, that the special orthogonal group is arc-connected, by relating every positive orthogonal matrix O to the identity matrix. To do this, we transform any rotation matrix of O and angle θ, into a rotation matrix of angle $t\theta$ for $t \in [0, 1]$, and by grouping the blocks of -1 two by two to identify them with a rotation matrix of angle π, which we transform into a matrix of angle $t\pi$.

Proposition 152. *Any compact subgroup G of $GL_n(\mathbb{R})$ is contained in a conjugate of the orthogonal group.*

Preuve : It is therefore a question of showing that G stabilizes a positive definite quadratic form. Indeed if A is the matrix of such a form, it is diagonalizable in orthonormal basis ${}^tPAP = D$ with ${}^tP = P^{-1}$ and $D = \text{diag}(\lambda_1, \cdots, \lambda_n)$ where the λ_i are strictly positive. We can then write $D = D_1^2$ where D_1 is a diagonal matrix with strictly positive eigenvalues so that $A = B^2$ with $B = {}^tPD_1P$. Thus for all $M \in G$, we have ${}^tBB = {}^tMB^2M$ or

$$ {}^t(BMB^{-1})(BMB^{-1}) = 1$$

i.e. BMB^{-1} is orthogonal.

Let us then consider the action of G on the set $\mathfrak{S}_n^{++}$ of positive definite quadratic forms according to the formula ${}^tBB \mapsto {}^tM{}^tBBM = {}^t(BM)BM$. The idea is to use a fixed point theorem which requires restricting ourselves to a compact convex. To have a compact we simply need to consider $\{{}^tBB : \ B \in G\}$ and for the convexity to take the convex envelope of this compact which therefore remains compact according to CarathÃ©odory's theorem. We thus have the action of a compact group on a compact convex and the lemma **??** below ensures the existence of a fixed point. It only remains to verify that the fixed point corresponds to a positive definite form, which follows from the following lemma.

Lemme 153. *The set $\mathfrak{S}_n^{++}$ of positive definite quadratic forms is convex.*

Preuve : The idea is to make simultaneous congruence. Let A and B be two symmetric positive definite matrices : A then defines a scalar product ϕ_A. The matrix $U = A^{-1}B$ defines a symmetric endomorphism relatively to ϕ_A since $AU = {}^tUA$ and therefore there exists an orthonormal basis for ϕ_A diagonalizing U, i.e. ${}^tPAP = I_n$ and $P^{-1}UP = D$ and therefore $(P^{-1}A^{-1}{}^tP^{-1}){}^tPBP = D$ or ${}^tPBP = D$. Thus for $0 \leq t \leq 1$, we have

$$ tA + (1-t)B = {}^tP^{-1}(tI_n + (1-t)D)P^{-1}$$

which is positive definite.

We now conclude with the fixed point lemma needed to finish proving the previous proposition.

Lemme 154. *Let $G \subset GL(E)$ be a compact group acting on a convex compact K of a Euclidean vector space E. There then exists a point of K fixed by all the elements of G.*

Preuve : We introduce the application

$$ x \in E \mapsto N_G(x) = \sup_{g \in G} ||g(x)||,$$

where the sup is in fact a maximum, i.e. is reached, because G is compact. We easily verify that N_G is a norm that is also strictly convex since

$$ N_G(x+y) = ||g_0(x) + g_0(y)|| \leq ||g_0(x) + g_0(y)|| \leq N_G(x) + N_g(y)$$

and to have equality it is necessary in particular that the first inequality is an equality and therefore that there exists $\lambda > 0$ such that $g_0(x) = \lambda g_0(y)$ and since g_0 is linear and invertible, $x = \lambda y$, hence the assertion.

Thus there exists a unique $x_0 \in K$ minimizing N_G and as trivially for all x we have $N_G(g(x)) = N_G(x)$, from the uniqueness of the minimum we have $g(x_0) = x_0$, hence the result.

4.5 The Hermitian case

In this paragraph for $\mathbb{K} = \mathbb{C}$, in order not to repeat the framework of quadratic forms, we consider the case where σ is the complex conjugation. For $A \in GL_n(\mathbb{C})$, we denote A^* for ${}^t\overline{A}$. Note in particular that any Hermitian form ϕ verifies $\phi(x,x) \in \mathbb{R}$. We then say that it is *positive* (resp. negative) if $\phi(x,x) \geq 0$ (resp. ≤ 0) for all $x \in E$ and we say that it is also *defined* if $\phi(x,x) = 0 \Rightarrow x = 0$.
Remarque: if E is equipped with a positive definite Hermitian form we say that E is *a Hermitian space.*

Proposition 155. *If ϕ is positive then*

$$\phi(x,y)\overline{\phi(x,y)} \leq q(x)q(y).$$

Theorem 156. *As in the real case,*

- *there exists a decomposition $E = E^{\perp} \oplus E^+ \oplus E^-$ such that the restriction of ϕ to E^+ (resp. E^-) is positive definite (resp. negative). Moreover the dimension s of E^+ and t of E^- are independent of this decomposition and the pair (s,t) is called* the signature *of ϕ.*
- *There exists a basis $(e_i)_{1\leq i\leq n}$ such that*

$$\phi\Big(\sum_{i=1}^{n}\lambda_i e_i, \sum_{j=1}^{n}\mu_j e_j\Big) = \sum_{i=1}^{s}\lambda_i\overline{\mu_i} - \sum_{i=s+1}^{s+t}\lambda_i\overline{\mu_i}.$$

- *The rank of ϕ is $s+t$ and its index $n-(s+t)+\min\{s,t\}$.*

Let us finish these algebraic reminders with a fundamental example of a Hermitian matrix.

Definition 157. *Let $(x_1,\cdots,x_m)$ be a family of vectors of a Hermitian space E.* The Gram matrix *defined by*

$$Gram(x_1,\cdots,x_m) = \Big(\phi(x_i,x_j)\Big)_{1\leq i,j\leq m}$$

is a Hermitian matrix whose determinant is denoted by $G(x_1,\cdots,x_m)$.

Proposition 158. *For all $(x_1,\cdots,x_m)$, the Gram determinant $G(x_1,\cdots,x_m)$ belongs to the positive real numbers; it is nonzero if and only if the family $(x_1,\cdots,x_m)$ is free.*

Corollory 159. *Let $x \in E$ and F be a subspace of E ; if $(x_1, \cdots, x_m)$ is a basis of F then the distance $d(x, F)$ from x to F is given by*

$$d(x, F)^2 = \frac{G(x, x_1, \cdots, x_m)}{G(x_1, \cdots, x_m)}.$$

Proposition 160. *An endomorphism is normal if and only if it is diagonalizable in an orthonormal basis.*

Preuve : If u is normal, then it is semi-simple according to 132 : for x an eigenvector of u, we have $E = \langle x \rangle \oplus \langle x \rangle^\perp$ and we conclude by induction on the dimension.

The converse is obvious and follows from the fact that a diagonal matrix commutes with its diagonal adjoint.

Corollory 161. *An endomorphism of a Hermitian space is*

- *Hermitian (resp. anti-Hermitian) if and only if, in an orthonormal basis, its matrix is real diagonal (resp. pure imaginary).*
- *unitary if and only if, in an orthonormal basis, its matrix is diagonal with coefficients of modulus* 1.

4.6 Congruence classes

Definition 162. *Two square matrices A and B of $\mathbb{M}_n(\mathbb{C})$ are said to be congruent, if there exists an invertible matrix P such that $B = P^*AP$.*

The congruence relation is clearly an equivalence relation. This relation is really only used for Hermitian matrices and corresponds to the effect of a change of basis on the matrix associated with a Hermitian product, cf. the proposition **??**. In this context, we will see that the classes are parameterized by the signature, cf. the theorem 147 in the real case and 156 in the Hermitian case. On the other hand unlike the case of similarity classes, we have a statement of simultaneous congruence as follows.

Proposition 163. *Let two Hermitian matrices A and B with B positive definite. There then exists an invertible matrix P verifying*

$$P^*AP = D \qquad \text{and} \qquad P^*BP = I_n,$$

where D is a diagonal matrix.

Remarque: we will note that the matrix P is not unitary so that A and B are not simultaneously diagonalizable !

Preuve : Let Q denote a matrix whose column vectors form an orthonormal basis for the Hermitian product defined by B, i.e. $(X, Y) \mapsto X^*BY$. We thus have $Q^*BQ = I_n$. Since Q^*AQ is visibly Hermitian, we diagonalize it into an orthonormal basis for the canonical Hermitian product, i.e. there exists a unitary matrix R such that $R^*(Q^*AQ)R = D$. We then have $R^*(Q^*BQ)R = R^*R = I_n$, so that the matrix $P = QR$ is suitable.

4.7 Unitary similarity classes

We are interested here in unitary similarity classes. Let us first note that if A and B are in the same unitary similarity class then

$$\sum_{i,j=1}^{n} |b_{ij}|^2 = \sum_{i,j=1}^{n} |a_{i,j}|^2$$

Indeed this follows from the equality $\sum_{i,j=1}^{n} |a_{i,j}|^2 = \mathrm{tr}(A^*A)$. More generally given a word $M(s,t) = s^{m_1}t^{n_1}s^{m_2}t^{n_2}\cdots s^{m_k}t^{n_k}$ in two variables s,t with $m_1, n_1, \cdots, m_k, n_k \geq 0$: the degree of $M(s,t)$ is by definition equal to $m_1 + n_1 + \cdots + m_k + n_k$. For $A \in \mathbb{M}_n(\mathbb{C})$, we set

$$M(A, A^*) = A^{m_1}(A^*)^{n_1}\cdots A^{m_k}(A^*)^{n_k}$$

We then note that for all $M(s,t)$, $\mathrm{tr}M(A, A^*)$ is constant on the unitary similarity class.

Theorem 164. ***(Specht)*** *Two matrices A, B are unitarily similar if and only if $\mathrm{tr}M(A, A^*) = \mathrm{tr}M(B, B^*)$ for all words $M(s,t)$.*

The obvious drawback of Specht's theorem is that an infinite number of conditions must be verified. The following Pearcy theorem allows us to reduce to a finite number.

Theorem 165. *(Pearcy [**?**]) Two matrices A, B are unitarily similar if and only if $\mathrm{tr}M(A, A^*) = \mathrm{tr}M(B, B^*)$ for any word $M(s,t)$ of degree at most $2n^2$.*

Remarque: a quick calculation shows that there are at most 4^{n^2} distinct words of degree at most $2n^2$. For $n = 2$, it is easy to show that it is in fact sufficient to consider the words s, s^2 and ts. For $n = 3$, we can show that it is sufficient to consider the 9 words, $s, s^2, ts, ts^2, t^2s^2, tsts, ts^2ts, ts^2t^2s$ instead of all the words of degree at most 18.

4.8 Eigenvalues of Hermitian matrices

For a general complex matrix, the eigenvalues are the roots of the characteristic polynomial. If the matrix is Hermitian, the eigenvalues are also the solutions of several optimization problems : this paragraph is devoted to the presentation of some of these, the interested reader can consult [**?**] §4.2. In the following A denotes a complex Hermitian matrix whose real eigenvalues are ranked in ascending order :

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Theorem 166. *(Courant-Fisher) For all $1 \leq k \leq n$, we have*

$$\lambda_k = \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{x^* A x}{x^* x} = \max_{F \in \mathcal{E}_{n-k+1}} \min_{0 \neq x \in F} \frac{x^* A x}{x^* x}$$

where $\mathcal{E}_k$ denotes the set of k-dimensional subspaces of $\mathbb{C}^n$.

Remarque: the case of λ_1 and λ_n is known as the Rayleight-Ritz theorem.

Preuve : Let us denote for all $1 \leq k \leq n$, x_k a unitary eigenvector for the eigenvalue λ_k. Then let $F \in \mathcal{E}_k$ be such that

$$F \cap \mathrm{vect}(x_k, x_{k+1}, \cdots, x_n)$$

is of dimension greater than 1. For $x = \sum_{i=k}^n \alpha_i x_i \neq 0$ a unit vector of this intersection, we have

$$x^* A x = \sum_{i=k}^n \lambda_i |\alpha_i|^2 \geq \lambda_k \sum_{i=k}^n |\alpha_i|^2 = \lambda_k.$$

We then deduce the inequality

$$\lambda_k \geq \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{x^* A X}{x^* x}.$$

Moreover for $F = \mathrm{vect}(x_1, \cdots, x_k)$, we have

$$\lambda_k = \max_{0 \neq x \in F} \frac{x^* A X}{x^* x},$$

hence the equality. The other case is treated in a strictly identical manner.

Let us cite some applications of the previous theorem, cf. [**?**] §4.3.

Theorem 167. *(Weyl) Let A, B be Hermitian matrices with eigenvalues $\lambda_i(A), \lambda_i(B)$ ranked in ascending order. By similarly classifying the eigenvalues of $A + B$, we obtain :*

(a) $\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A + B) \leq \lambda_k(A) + \lambda_n(B)$, for all $k = 1, \cdots, n$;

(b) if B has rank at most r :

- *$\lambda_k(A + B) \leq \lambda_{k+r}(A) \leq \lambda_{k+2r}(A + B)$ for $k = 1, \cdots, n - 2r$;*
- *$\lambda_k(A) \leq \lambda_{k+r}(A + B) \leq \lambda_{k+2r}(A)$, for $k = 1, \cdots, n - 2r$;*

(c) $\lambda_{j+k-n}(A + B) \leq \lambda_j(A) + \lambda_k(B)$, for all $1 \leq j, k \leq n$ such that $j + k \geq n + 1$;

(d) $\lambda_j(A) + \lambda_k(B) \leq \lambda_{j+k-1}(A + B)$, for all $1 \leq j, k \leq n$ such s that $j + k \leq n + 1$.

Preuve : (a) We have $\frac{x^*(A+B)x}{x^*x} = \frac{x^*Ax}{x^*x} + \frac{x^*Bx}{x^*x}$; the result then follows simply from the framework $\lambda_1(B) \leq \frac{x^*Bx}{x^*x} \leq \lambda_n(B)$ and the theorem 166.

(b) We write B in the form $\alpha_1 u_1 u_1^* + \cdots + \alpha_r u_r u_r^*$ where the vectors u_i of $\mathbb{C}^n$ are not necessarily independent. We then have

$$\begin{aligned}\lambda_{k+2r}(A+B) &= \min_{F\in\mathcal{E}_{k+2r}} \max_{0\neq x\in F} \frac{x^*(A+B)x}{x^*x} \\ &\geq \min_{F\in\mathcal{E}_{k+2r}} \max_{0\neq x\in F\cap \mathrm{vect}(u_1,\cdots,u_r)^\perp} \frac{x^*(A+B)x}{x^*x}.\end{aligned}$$

Noting $\mathcal{E}'_{k+r}$ the set of subspaces of $\mathrm{vect}(u_1,\cdots,u_r)^\perp$ of dimension $k+r$, the last term above is equal to

$$\min_{F\in\mathcal{E}'_{k+r}} \max_{0\neq x\in F} \frac{x^*Ax}{x^*x} \geq \min_{F\in\mathcal{E}_{k+r}} \max_{0\neq x\in F} \frac{x^*Ax}{x^*x} = \lambda_{k+r}(A).$$

According to the same scheme, we have

$$\begin{aligned}\lambda_k(A+B) &= \max_{F\in\mathcal{E}_{n-k+1}} \min_{0\neq x\in F} \frac{x^*(A+B)x}{x^*x} \\ &\leq \max_{F\in\mathcal{E}_{n-k+1}} \min_{0\neq x\in F\cap \mathrm{vect}(u_1,\cdots,u_r)^\perp} \frac{x^*(A+B)x}{x^*x} \\ &= \max_{F\in n-k+1-r} \min_{0\neq x\in F} \frac{x^*Ax}{x^*x} \\ &\leq \max_{F\in\mathcal{E}_{n-k+1-r}} \min_{0\neq x\in F} \frac{x^*Ax}{x^*x} \\ &= \lambda_{k+r}(A)\end{aligned}$$

These two families of inequalities then provide those of the statement.

(c) We diagonalize $A = UD_1U^*$ and $B = VD_2V^*$ and we denote u_i (resp. v_i) the column vectors of the unitary matrix U (resp. V). For a pair (j,k) verifying the conditions of the statement, we note for β large enough such that for all $j+1 \leq i \leq n$ and for all $k+1 \leq i' \leq n$, $\lambda_i(A) - \beta < \lambda_j(A)$ and $\lambda_{i'}(B) - \beta < \lambda_k(B)$:[1]

$$A_j = \beta(u_n u_n^* + \cdots + u_{j+1}u_{j+1}^*) \qquad B_k = \beta(v_n^* + \cdots + v_{k+1}v_{k+1}^*).$$

Noting that $(A - A_j)u_i = (\lambda_i - \beta)u_i$ for $i = j+1,\cdots,n$, and is equal to $\lambda_i u_i$ for $i = 1,\cdots,j$, we note that $\lambda_n(A - A_j) = \lambda_j(A)$. Similarly we have $\lambda_n(B - B_k) = \lambda_k(B)$.

Moreover, since A_j (resp. B_k) is of rank $n-j$ (resp. $n-k$), $A_j + B_k$ is of rank at most $2n - j - k$ and therefore according to (b)

$$\begin{aligned}\lambda(A - A_j + B - B_k) &= \lambda_n(A + B - (A_j + B_k)) \\ &\geq \lambda_{n-(2n-j-k)}(A+B) \\ &= \lambda_{j+k-n}(A+B)\end{aligned}$$

According to (a) for $k = n$, we also have

$$\lambda_n(A - A_j + B - B_k) \leq \lambda_n(A - A_j) + \lambda_n(B - B_k)$$

1. We will note that in [?] theorem 4.3.6 (c), the author forgets that $\lambda_j(A)$ can be strictly negative, so that $\lambda_n(A-B)$ would be zero and not equal to $\lambda_{n-r}(A)$.

so that

$$\begin{aligned}\lambda_k(A)+\lambda_k(B) &= \lambda(A-A_j)+\lambda_n(B-B_k) \geq \lambda_n(A-A_j+B-B_k)\\ &= \lambda_n((A+B)-(A_j+B_k)) \geq \lambda_{j+k-n}(A+B)\end{aligned}$$

(d) The result follows directly from (c) considering $-A$ and $-B$ and noting that $\lambda_i(-A) = -\lambda_{n-i+1}(A)$.
Remarque: in (b) for the case $r = 1$, the eigenvalues are interleaved, i.e.

$$\lambda_k(A+B) \leq \lambda_k(A) \leq \lambda_{k+2}(A+B) \qquad \lambda_k(A) \leq \lambda_{k+1}(A+B) \leq \lambda_{k+2}(A)$$

The same phenomenon occurs in the following situation.

Theorem 168. *Let $A \in \mathbb{M}_{n+1}(\mathbb{C})$ be a hermitian eigenvalue with $\lambda_1 \leq \cdots \leq \lambda_{n+1}$. Then the eigenvalues $\lambda'_1 \leq \cdots \leq \lambda'_n$ of a principal extracted matrix[2] of A verifies the following inequalities :*

$$\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \lambda'_2 \leq \cdots \leq \lambda'_{n-1} \leq \lambda_n \leq \lambda'_n \leq \lambda_{n+1}$$

Preuve : For the sake of simplicity, let us assume that $i = n+1$, we then have

$$\begin{aligned}\lambda_{k+1} &= \min_{F\in\mathcal{E}_{k+1}} \max_{0\neq x\in F} \frac{x^*Ax}{x^*x}\\ &\geq \min_{F\in\mathcal{E}_{k+1}} \max_{0\neq x\in F\cap \mathrm{vect}(e_{n+1})^\perp} \frac{x^*Ax}{x^*x}\\ &= \min_{F'\in\mathcal{E}'_k} \max_{0\neq x'\in F'} \frac{(x')^*A'x'}{(x')^*x'}\\ &= \lambda'_k\end{aligned}$$

where we wrote for $x \in \mathbb{C}^{n+1}$, $x' \in \mathbb{C}^n$ is the vector obtained by removing the last coordinate : we adopted similar notations for F' and $\mathcal{E}'_k$ denotes the k-dimensional subspaces of $\mathbb{C}^n$. The upper bound $\lambda_k \leq \lambda'_k$ is proved in the same way by considering $\max_{F\in\mathcal{E}_{n-k+1}} \min_{0\neq x\in F}$.
Remarque: obviously the demonstration simply adapts to the case where we consider a principal extracted matrix A_r of A, where we have deleted r-rows and the corresponding r-columns. The result is then :

$$\lambda_k(A) \leq \lambda_k(A_r) \leq \lambda_{k+n-r}(A).$$

The previous theorem also admits a reciprocal :

Theorem 169. *Let for an integer $n \geq 1$, be real numbers such as :*

$$\lambda_1 \leq \lambda'_1 \leq \ lambda_2 \leq \lambda'_2 \leq \cdots \leq \lambda'_{n-1} \leq \lambda_n \leq \lambda'_n \leq \lambda_{n+1}$$

Let $A' = \mathrm{diag}(\lambda'_1, \cdots, \lambda'_n)$, then there exists a real a and a vector $y \in \mathbb{R}^n$ such that $\{\lambda_1, \cdots, \lambda_{n+1}\}$ is the set of eigenvalues of the symmetric matrix : $A = \begin{pmatrix} A' & y \\ {}^t y & a \end{pmatrix}$.

2. i.e. for an index $1 \leq i \leq n+1$, we remove from A its i-th column and its i-th row

Preuve : Computing the trace gives $a = \sum_{i=1}^{n+1} \lambda_i - \sum_{i=1}^{n} \lambda'_i$. For all t distinct from λ_i, we have the following equality :

$$\begin{pmatrix} I & 0 \\ {}^t((tI-A')^{-1}y) & 1 \end{pmatrix} \begin{pmatrix} tI-A' & -y \\ -{}^t y & t-a \end{pmatrix} \begin{pmatrix} I & (tI-A')^{-1}y \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} tI-A' & 0 \\ 0 & (t-a) - {}^t y(tI-A')^{-1}y \end{pmatrix}$$

Taking the determinants, we then obtain the following equality between the characteristic polynomials :

$$\chi_A(t) = \chi_{A'}(t)\Big[(t-a) - \sum_{i=1}^{n} y_i^2 \frac{1}{t-\lambda_i}\Big]$$

with therefore $\chi_{A'}(t) = \prod_{i=1}^{n}(t-\lambda_i)$. It is then a question of proving the existence of n real y_i such that $\chi_A(\lambda_k) = 0$ for all $k = 1, \cdots, n+1$. We then consider the Euclidean division $\xi_A(t) := \prod_{i=1}^{n+1}(t-\lambda_i) = Q(t)\chi_{A'}(t) + R(t)$ with therefore $Q(t) = t-a$ and for all $1 \leq k \leq n$, $R(\lambda'_k) = \xi_A(\lambda'_k)$, which determines only the polynomial R of degree lower or than n using for example the Lagrange interpolating polynomials. Let us assume for simplification that all λ'_i are distinct, we then have

$$R(t) = \sum_{i=1}^{n} \xi_A(\lambda'_i) \frac{\chi_{A'}(t)}{\chi'_{A'}(t)(t-\lambda'_i)}$$

so that

$$\frac{\xi_A(t)}{\chi_{A'}(t)} = (t-a) - \sum_{i=1}^{n} \frac{-\xi_A(\lambda'_i)}{\chi'_{A'}(t)} \frac{1}{t-\lambda'_i}$$

It is then sufficient to show that for all $i = 1, \cdots, n$, $\xi_A(\lambda'_i)\chi'_{A'}(\lambda'_i) \leq 0$ so that by setting $y_i^2 = -\frac{\xi_A(\lambda'_i)}{\chi'_{A'}(\lambda'_i)}$, we will have $\chi_A(t) = \xi_A(t)$.

This involves using the hypothesis of interlacing of eigenvalues : we thus note that

$$\xi_A(\lambda'_i) = (-1)^{n-i+1} \prod_{j=1}^{n+1} (\lambda'_i - \lambda_j) \qquad \chi'_{A'}(\lambda'_i) = (-1)^{n-i} \prod_{\substack{j=1 \\ j\neq i}}^{n} (\lambda'_i - \lambda'_j)$$

are effectively of opposite signs.
Remarque: in the case where some of the λ'_i are equal, for example $\lambda'_1 = \lambda'_2 = \cdots \lambda'_k < \lambda'_{k+1} \leq \cdots$, we then notice that $(t-\lambda'_1)^{k-1}$ divides $\xi_A(t)$ and we resume the previous reasoning by dividing $\xi_A(t)$ and $\chi_{A'}(t)$ by $(t-\lambda'_1)^{k-1}$.

Corollory 170. *Let $A \in \mathbb{M}_n(\mathbb{C})$ be Hermitian ; for $1 \leq r \leq n$, U denotes a matrix of $\mathbb{M}_{n,r}(\mathbb{C})$ such that its column vectors form an orthonormal family, i.e. $U^*U = I \in \mathbb{M}_r(\mathbb{C})$. We then have the following properties :*

(i) for all $k = 1, 2, \cdots, r$, $\lambda_k(A) \leq \lambda_k(U^*AU) \leq \lambda_{k+n-r}(A)$;
(ii) $\lambda_1(A) + \cdots + \lambda_r(A) = \min_{U^*U=I \in \mathbb{M}_r(\mathbb{C})} \mathrm{tr}(U^*AU)$ *and*

$$\lambda_{n-r+1}(A) + \cdots + \lambda_n(A) = \max_{U^*U=I \in \mathbb{M}_r(\mathbb{C})} \mathrm{tr}(U^*AU).$$

Remarque: (i) is known as the PoincarÃ© separation theorem and is used in quantum mechanics where we have access to the calculations of $u_i^*Au_j$ for an orthonormal family $(u_i)_{1 \leq i \leq r}$.

Preuve : (i) if $r < n$, we complete the column vectors of U in an orthonormal basis ; the matrix U' is then unitary and $(U')^*AU'$ has the same eigenvalues as A and U^*AU is a principal extracted matrix, the result then follows from 168, or rather from the remark that follows.

(ii) the upper bounds follow directly from (i), the equalities are then obtained if the columns of U correspond to the eigenvectors of the r smallest eigenvalues.

We have the same kind of result for singular values.

Theorem 171. *Let* $A \in \mathbb{M}_{m,n}(\mathbb{C})$ *and* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q$ *be its singular values for* $q = \min\{m, n\}$. *For* $1 \leq k \leq q$, *we have*

$$\sigma_k = \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{||Ax||_2}{||x||_2} = \min_{F \in \mathcal{E}_{n-k+1}} \max_{0 \neq x \in F} \frac{||Ax||_2}{||x||_2}$$

Remarque: following the same scheme as before, we can obtain similar results on the entanglement of singular values. For example for $A \in \mathbb{M}_{m,n}(\mathbb{C})$ with $m \geq n$, for $1 \leq i \leq n$, $A^{(i)}$ denotes the matrix extracted from A by removing its i-th column and row. We denote by σ_i (resp. σ_i') the singular values of A (resp. A'), so that we have

$$\sigma_1 \geq \sigma_1' \geq \sigma_2 \geq \sigma_2' \geq \cdots \geq \sigma_{n-1}' \geq \sigma_n \geq 0$$

We then deduce cf. [**?**] 3.1.3, that if A_r denotes a submatrix of A obtained by removing r rows and/or columns then $\sigma_k(A) \geq \sigma_k(A_r) \geq \sigma_{k+r}(A)$.

Corollory 172. *Let* $A \in \mathbb{M}_n(\mathbb{C})$ *have singular values* $\sigma_1 \geq \cdots \geq \sigma_n$ *and let* $H(A) = \frac{1}{2}(A + A^*)$ *be its Hermitian part with eigenvalues* $\lambda_1 \geq \cdots \lambda_n$. *For all* $k = 1, \cdots, n$, *we have* $\lambda_k \leq \sigma_k$.

Preuve : For $x \in \mathbb{C}^n$ unitary, we have $x^*H(A)x = \mathrm{Re}\ (x^*Ax) \leq |x^*Ax| \leq ||x||_2.||Ax||_2 = ||Ax||_2$. So we have

$$\lambda_k = \min_{F \in \mathcal{E}_k} \max_{\substack{0 \neq x \in F \\ ||x||_2 = 1}} x^*H(A)x \leq \min_{F \in \mathcal{E}_k} \max_{\substack{0 \neq x \in F \\ ||x||_2 = 1}} ||Ax||_2 = \sigma_k$$

Corollory 173. *(cf. [**?**] 3.3.2) Let* $A \in \mathbb{M}_n(\mathbb{C})$ *have singular values* $\sigma_1 \geq \cdots \geq \sigma_n$ *and eigenvalues* $\{\lambda_1, \cdots, \lambda_n\}$ *ordered so that* $|\lambda_1| \geq \cdots \geq |\lambda_n|$. *We then have*

$$|\lambda_1 \cdots \lambda_k| \leq \sigma_1 \cdots \sigma_k \quad \forall k = 1, \cdots, n$$

with equality for $k = n$.

Preuve : Let U be unitary such that $U^*AU = T$ is upper triangular and where the diagonal of T is $(\lambda_1, \cdots, \lambda_n)$. Let $U_k \in \mathbb{M}_{n,k}(\mathbb{C})$ be the matrix extracted from U consisting of its first k columns such that $U^*AU = \begin{pmatrix} U_k^*AU_k & * \\ * & * \end{pmatrix}$. The matrix $T_k = U_k^*AU_k$ is therefore upper triangular with diagonal equal to $\lambda_1, \cdots, \lambda_k$ such that $|\det T_k| = |\lambda_1 \cdots \lambda_k|$ is equal to the product $\sigma_1(T_k) \cdots \sigma_k(T_k)$ of the singular values of T_k. The result then follows from the remark that follows the theorem 171, i.e. $\sigma_1(T_k) \cdots \sigma_k(T_k) \leq \sigma_1 \cdots \sigma_k$.

4.9 Exercises

Exercice 1. *Show that two unitarily similar real matrices are orthogonally similar.*

Preuve : We have $A = UBU^{-1}$ and ${}^tA = U{}^tBU^{-1}$. We recall that two real matrices that are similar on $\mathbb{C}$ are similar on $\mathbb{C}$. There therefore exists $P \in GL_n(\mathbb{R})$ such that $A = PBP^{-1}$ and ${}^tA = P{}^tBP^{-1}$. Let $P = OS$ be the polar decomposition of P and $A = OSB^{-1}B^{-1}o^{-1}$. The result then follows from the fact that B and S commute : indeed we have $PBP^{-1} = A = {}^t({}^tA) = {}^tP^{-1}B{}^tP$. Thus B commutes with ${}^tPP$ and therefore also with S which is a polynomial in ${}^tPP$.
Remarque: As an application, we can deduce the reduction of isometries, from the diagonalization of unitary endomorphisms.

Exercice 2. *Show that the only real symmetric Bourdaud matrices (i.e. the eigenvalues are read on the diagonal) are the diagonal matrices.*

Preuve : Let λ_1 be the largest of the eigenvalues, we have $\lambda_1 = \sum_{||x||=1}(A(x), x)$ and where the sup is reached, we are in the presence of an eigenvector associated with λ_1 : indeed the differential vanishes at x_0 on the tangent space to the sphere at x_0 (related extrema), i.e. $(A(x_0), y) = 0$ for all y such that $(x_0, y) = 0$ so that $A(x_0)$ is collinear with x_0.

We then note that there exists k such that $\lambda_1 = a_{k,k} = (A(e_k), e_k)$ and we reason by induction on the orthogonal to e_k.

We could also have used the quadratic form $A \mapsto \mathrm{tr}({}^tAA) = \sum_{i,j} \lambda_{i,j}^2$ which is clearly invariant under the action by conjugation of the orthogonal group so that $\mathrm{tr}({}^tAA) = \sum_i \lambda_{i,i}^2$, hence the result.

Exercice 3. *Let G be a subgroup of $O(n)$ then G is finite if and only if G has finite exponent if and only if the set of traces of the elements of G is finite.*

Preuve : If G is finite it is clearly of finite exponent. If G is of finite exponent then any element of G is similar to a block diagonal matrix with on the diagonal I_r, $-I_s$ and matrices of size 2, of rotations whose angles are then of the form $\frac{2k\pi}{n}$, which gives therefore a finite set of possible traces.

If the set of traces is finite, consider vect$G$ the vector subspace of $\mathcal{L}_n(\mathbb{R})$ generated by the elements of G whose basis is fixed $g_1, \cdots, g_r$. We also consider the scalar product $(A, B) := \mathrm{tr}({}^tAB)$. We denote $a_i(g)$ as the component of g on g_i, i.e. $g = \sum_i a_i(g)g_i$. We compose on the right with g_j^{-1} and take the trace. Since $g_i^{-1} = {}^tg_j$, we have $\mathrm{tr}(g_i g_j^{-1}) = (g_i, g_j)$ and therefore

$$\mathrm{tr}(gg_j^{-1}) = \sum_i a_i(g)(g_i, g_j)$$

and since the g_i are linearly independent, the matrix M whose elements are the (g_i, g_j) is invertible therefore

$$a_i(g) = \sum_j (M^{-1})_g \mathrm{tr}(gg_j^{-1})$$

so that we obtain a finite number of $a_i(g)$ and therefore of g.

Exercice 4. *A rotation of $SO(3)$ will be denoted by $r = (k, \theta)$ where k is the unit vector of the axis of the rotation and θ its angle.*

Let $r = (OA, 2\alpha)$ and $s = (OB, 2\beta)$ be two rotations such that $\frac{\alpha}{\pi}$ and $\frac{\beta}{\pi}$ are irrational. Show that if we except a countable infinity of values for the measure c of the angle between the axes OA and OB, the group generated by r and s is dense in $SO(3)$.

Preuve : If P_3 is the plane OAB and $P_2 = (OA, -\alpha)(P_3)$ then r is written as the product of the reflections with respect to the planes P_2 and P_3. Similarly s is the product of P_1 and P_2 where $P_1 = (OB, \beta)(P_3)$.

In order to approximate a rotation $(k, 2\theta)$, we approximate its axis then its angle. To approximate $\mathbb{R}.k$, we approximate the planes it determines with OA, and OB. By the Jacobi-Kronecker theorem, they are respectively approximated by $P_2' = (OA, -p\alpha)(P_3)$ and $P_1' = (OB, q\beta)(P_3)$ if p and q are adequate integers. Thus $\mathbb{R}k' = P_1' \cap P_2'$ approaches $\mathbb{R}k$.

Since $r^p = (OA, 2p\alpha)) = (P_3)(P_2')$ and $s^q = (OB, 2q\beta) = (P_1')(P_3)$, we have $s^q r^p = (P_1')(P_2')$ whose angle measure $2\gamma'$ is given by the fundamental formula of spherical trigonometry

$$\cos\gamma' = \sin(p\alpha)\sin(q\beta)\cos c - \cos(p\alpha)\cos(q\beta)$$

We are looking for an irrational $\frac{\gamma'}{\pi}$; the previous formula shows that if p and q describe the integers and if $\frac{\gamma'}{\pi}$ describes the rationals, $\cos c$ takes only a countable infinity of values. We then choose c so that $\cos c$ does not belong to this set of values. It then follows that $\frac{\gamma'}{\pi}$ is irrational for all p, q. The Jacobi-Kronecker theorem then shows that we can choose n so that $2n\gamma'$ approaches 2θ so that $(s^q r^p)^n$ approaches $(k, 2\theta)$.

Exercice 5. *(Tauvel p.417) Show that u is diagonalizable with real spectrum if and only if u is the product of two Hermitian endomorphisms, at least one of them being positive definite*

Preuve : Suppose that u is diagonalizable with real spectrum : let $(e_1, \cdots, e_n)$ be an orthonormal basis and let $(x_1, \cdots, x_n)$ be a basis formed by eigenvectors of u with real eigenvalues $\lambda_1, \cdots, \lambda_n$. We define f and g by $f(x_i) = e_i$ and $g(e_i) = \lambda_i e_i$ so that $u = f{-}1 \circ g \circ f$. Let $f = qr$ be the polar decomposition of f with q unitary and r Hermitian positive definite. We thus obtain by setting $l = r \circ l \circ r^{-1} = q^* \circ g \circ q$ which is Hermitian, then $u = r^{-1} \circ l \circ r = (r^{-1} \circ l \circ r^{-1}) \circ r^2$ with therefore $r^{-1} \circ l \circ r^{-1}$ and r^2 Hermitian.

Conversely if $u = v \circ w$ with w (resp. v) Hermitian (resp. Hermitian positive definite). Let us note $l = v^{1/2}$ which is Hermitian positive definite, we have $u = l \circ (l \circ w \circ l) \circ l^{-1}$. Since $l \circ w \circ l$ is Hermitian, it is diagonalizable with real spectrum and so is u which is similar to it.

Exercice 6. *(Tauvel p.418) For $n \geq 2$ and A non-zero Hermitian of size n, show that A is positive definite or negative if and only if for any Hermitian matrix B, AB is diagonalizable.*

Preuve : The direct meaning follows from the previous exercise. Let us therefore assume that for any Hermitian matrix B, AB is diagonalizable. Let P be unitary such that $A = PDP^*$ with D real diagonal. Since $AB = P(DP^*BP)P^{-1}$, we see that AB is diagonalizable if and only if DP^*BP is. Moreover, since P^*BP is Hermitian, we can assume $A = D$.

We easily reduce to dimension 2 with $A = \mathrm{diag}(x^2, -y^2)$ with x, y real. Let then $B = \begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix}$ and $B' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. If $y \neq 0$, we have $AB \neq 0$ and $(AB)^2 = 0$. If $y = 0$ then $AB' \neq 0$ and $(AB')^2 = 0$. Thus AB (resp. AB') is non-zero nilpotent and therefore not diagonalizable.

Exercice 7. *(M-T p.219 or M p.92) Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of a complex matrix $A = (a_{i,j})$. Show that A is normal if and only if $\sum_{i,j} |a_{i,j}|^2 = \sum_i |\lambda_i|^2$.*

Preuve : Recall that $\mathrm{tr}(AA^*) = \sum_{i,j} |a_{i,j}|^2$ is $U(n)$-invariant so that if A is normal it is then unitarily similar to the diagonal matrix of λ_i hence the direct sense. Conversely any complex matrix is unitarily similar to a triangular matrix T with diagonal formed by the λ_i. The equality then implies that the terms that are not on the diagonal are zero, i.e. that T is diagonal.

Exercice 8. *Show that A is normal if and only if $\mathrm{tr}(AA^*)^2 = \mathrm{tr}A^2A^{*2}$.*

Preuve : According to the previous exercise, the Hermitian matrix $H = AA^* - A^*A$ is zero if and only if $\mathrm{tr}H^2 = 0$. Thus A is normal if and only if the trace of $(AA^*)^2 - A(A^*)^2A - A^*A^2A^* + (A^*A)^2$ is zero. Or let us recall that $\mathrm{tr}AB = \mathrm{tr}BA$ so that $\mathrm{tr}(A^*A)^2 = \mathrm{tr}(AA^*)^2$ and $\mathrm{tr}A^*A^2A^* = \mathrm{tr}A(A^*)^2A = \mathrm{tr}A^2A^{*2}$, hence the result.

Exercice 9. *Show that two matrices are unitarily equivalent if and only if they have the same singular values*

Preuve : The direct meaning follows from $PAP^* = B$, $PA^*P^* = B^*$ or $PAA^*P^* = BB^*$. Conversely, the polar decomposition gives $A = H_A U_A$, $B = H_B U_B$ with H_A, H_B Hermitian positive definite, and U_A, U_B unitary. We therefore have $AA^* = H_A H_A^*$ and $BB^* = H_B H_B^*$. The matrices H_A, H_B are diagonalizable so that if AA^* and BB^* have the same eigenvalues then the eigenvalues of H_A are equal to those of H_B up to the sign, i.e. $H_A = P_A D P_A^*$ and $H_B = P_B D D' P_B^*$ with $D' = \mathrm{diag}(\epsilon_1, \cdots, \epsilon_n)$ where the ϵ_i are equal to ± 1 so that D' is unitary. We therefore have $D = P_B^* U_B^* B P_B D'$ and $A = U_A P_A P_B^* U_B^* B P_B D' P_A^*$ hence the result.

Exercice 10. *Give the center Z of $O(q)$ (resp. Z^+ of $O^+(q)$) and show that $O(q)$ is a semi-direct product of $O^+(q)$ by $\mathbb{Z}/2\mathbb{Z}$; under what condition can this semi-direct product be taken direct ?*

Preuve : It is clear that $\{Id, -Id\} \subset Z$; conversely, let $z \in Z$ and τ_D be a reflection of line D. We have $z\tau_D z^{-1} = \tau_D = \tau_{z(D)}$ so that z leaves all the lines of the space stable ; it is therefore a homothety (classical result) and therefore $z = \pm Id$.

Concerning Z^+ let us note that $-Id$ belongs to $O^+(q)$ if and only if n is even. For $n \geq 3$ let τ_P be a reversal of plane P ; we have $z\tau_P z^{-1} = \tau_P = \tau_{z(P)}$ so that z leaves all planes of the space stable. Since any line is the intersection of two planes, we deduce that z leaves all lines of the space stable, i.e. $Z^+ = \{Id\}$ for n odd and otherwise $Z^+ = Z$ for n even. For $n = 2$, it is well known that O^+ is commutative.

It is clear that the exact sequence $1 \to O^+(q) \longrightarrow O(q) \longrightarrow \mathbb{Z}/2\mathbb{Z} \to 0$ is split, a lift being given for example by any reflection. To obtain a direct product, it is necessary to find an element of order 2 which is not in O^+ and which commutes to all the elements of O^+ ; the only possibility is then $-Id$ in odd dimension.

□

Exercice 11. *Let $u \in O(q)$ and $F_u = \{x \in E \ / \ u(x) = x\}$ and we denote $p_u = n - \dim F_u$. Show by induction on p_u, that u is the product of at most p_u reflections. Then show that u is the product of at least p_u reflections.*

Preuve : We reason by induction on p_u, the case $p_u = 0$ corresponding to $u = Id$. Let us therefore suppose $p_u > 0$ and let $x \in F_u^\perp$ be non-zero and let $y = u(x) \neq x$ because $x \notin F_u$; we have $y \in F_u^\perp$ because F_u being stable by u, $F_u^\perp$ is also stable. Moreover, since x and y have the same norm, we deduce that $(x - y, x + y) = 0$ (isosceles triangle). We then consider the reflection τ defined by $x - y$ so that $\tau(x - y) = y - x$ and $\tau(x + y) = x + y$ is therefore $\tau(y) = x$ with $\tau_{|F_u} = Id$. Thus we have $F_u \subset F_{\tau \circ u}$ the latter containing x so that $p_{\tau \circ u} < p_u$ and we conclude by recurrence.

Furthermore if u is the product of r reflections then F_u is clearly of dimension greater than or equal to $n-r$ (the intersection of r hyperplanes) is therefore $p_u \leq r$.

□

Exercice 12. *Show that for $n \geq 3$, any element of $O^+(q)$ is a product of at most n inversions.*

Preuve : The case $n = 3$ is obvious by noting that if τ is a reflection, then $-\tau$ is a inversion so that the product of two reflections (and hence any product of an even number) is a product of two inversions $\tau_1 \circ \tau_2 = (-\tau_1) \circ (-\tau_2)$.

For $n \geq 3$, let τ_1 and τ_2 be reflections with respect to the hyperplanes H_1 and H_2 and $u = \tau_1 \circ \tau_2$. Then let $V \subset H_1 \cap H_2$ be a subspace of dimension $n-3$: $u_{|V} = Id$ and $V^\perp$ is stable under u. According to the case $n = 3$, we have $u_{V^\perp} = \sigma_1 \circ \sigma_2$ where σ_1, σ_2 are reversals of $V^\perp$. We obtain the result by extending the σ_i by the identity on V.

□

Exercice 13. *Let u_1 and u_2 be two orthogonal symmetries of the same nature (i.e. such that $\dim \mathrm{Ker}(u_1 - Id) = \dim \mathrm{Ker}(u_2 - Id)$). Show that u_1 and u_2 are conjugate by $O^+(q)$. Deduce then that $D(O(q)) = D(O^+(q)) = O^+(q)$.*

Preuve : We decompose the space $E = E_1 \oplus E_1^\perp = E_2 \oplus E_2^\perp$ where $E_i = \mathrm{Ker}(u_i - Id)$. We then choose orthonormal bases (e_i^1) and (e_i^2) of E adapted to these decompositions. Let then u such that $u(e_i^1) = e_i^2$; u is an isometry and if we change ϵ_1 to $-\epsilon_1$, we can assume that u is positive. We then immediately verify that $u \circ u_1 \circ u^{-1} = u_2$.

The inclusion $D(O(q)) \subset O^+(q)$ is obvious ; conversely, let τ_1 and τ_2 be two reflections and let u be such that $u \circ \tau_1 \circ u^{-1} = \tau_2$ so that $\tau_1 \circ \tau_2 = [\tau_1, u]$. Since any element of $O^+(q)$ is the product of an even number of reflections, we obtain the reciprocal inclusion.

Similarly, to show that $O^+(q) \subset D(O^+(q))$ for $n \geq 3$, it suffices to show that any inversion is a commutator. Let V be a subspace of dimension 3 and (e_1, e_2, e_3) an orthonormal basis. Let $\sigma_1, \sigma_2, \sigma_3$ be the reversals defined by $(\sigma_i)_{|V^\perp} = Id$ and $\sigma_i(e_i) = e_i$) and therefore $\sigma_i(e_j) = -e_j$ for $i \neq j$. We then have $\sigma_3 = \sigma_1 \circ \sigma_2$. Furthermore there exists $u \in O^+(q)$ such that $\sigma_2 = u \circ \sigma_1 \circ u^{-1}$ and therefore $\sigma_3 = [\sigma_1, u]$.

□

Exercice 14. *Show that for all $u \in O(q)$, there exists an orthogonal decomposition*

$$E = \mathrm{Ker}(u - Id) \oplus \mathrm{Ker}(u + Id) \oplus P_1 \oplus \cdots \oplus P_r$$

where the P_i are stable planes by u, such that the restriction of u there is a rotation.

Preuve : We proceed by induction on the dimension, the cases $n = 1$ and $n = 2$ being well known. If u admits a real eigenvalue (necessarily ± 1), it is over (in particular if n is odd). Otherwise let $\lambda \in \mathbb{C}$ be an eigenvalue of the complexified of $u_{\mathbb{C}}$, so that $\bar{\lambda}$ is also an eigenvalue. Let then $x \in E \otimes_{\mathbb{R}} \mathbb{C}$ be an eigenvector of the complexified relative to λ and let $\bar{x}$ be its conjugate which is then eigen for $\bar{\lambda}$ relative to $u_{\mathbb{C}}$. The complex plane $P = \mathbb{C}x + \mathbb{C}\bar{x}$ is then invariant by $u_{\mathbb{C}}$. We then note that the vectors $\frac{x+\bar{x}}{2}$ and $\frac{x-\bar{x}}{2i}$ are real and form a basis of P so that the real plane they generate is stable under u. □

Exercice 15. - *We want to prove the simplicity of $O^+(3,\mathbb{R})$. Let N be a distinguished subgroup not reduced to the identity ; explain why it is sufficient to show that N contains a reversal.*

- Let $u \in N$ be a rotation of axis D and let P be the plane orthogonal to D at the origin so that the restriction of u to P is a rotation of angle θ that we assume $0 < \theta < \pi$. Let then x and $y = u(x)$ be points of the unit sphere of E ; let d be the distance between x and y. Show that for all $0 \leq d' \leq d$, there exist x_1, x_2 points of the unit sphere at a distance d' from each other and such that $x_2 = u(x_1)$.

- Deduce from the above that given y_1, y_2 points of the unit sphere at a distance of d' with $0 \leq d' \leq d$, there exist $u' \in N$ such that $u'(y_1) = y_2$. Considering the rotation of axis z and angle π/m for m large enough, construct a reversal of N and conclude.

Preuve : - Since the inversions generate $O^+(3,\mathbb{R})$ and are conjugate under $O^+(3,\mathbb{R})$, it suffices to show that N contains one.

- A classical calculation gives $d^2 = 2(1 - \cos\theta)$. Let a be one of the points of $D \cap S^2$; the result follows from the observation that u sends the meridian containing a and x, onto the one containing a and y and that when x_1 varies from x to a, the distance $||x_1 - u(x_1)||$ varies continuously from d to 0. Precisely, we consider $x + \lambda a$ with a squared norm equal to $1 + \lambda^2$ so that $x_1 = \frac{x+\lambda a}{\sqrt{1+\lambda^2}} \in S^2$. We then have $||u(x_1) - x_1|| = \frac{d}{\sqrt{1+\lambda^2}}$ so that it suffices to take $\lambda = \frac{\sqrt{d^2-m^2}}{m}$.

- Let x_3 (resp. y_3) be a vector of norm 1 orthogonal to the plane generated by x_1 and x_2 (resp. y_1 and y_2) and let u be such that $s(x_i) = y_i$ for $i = 1, 2, 3$. It is clear that s preserves the scalar product and therefore $u \in O(3,\mathbb{R})$; even if we change y_3 to $-y_3$, we can assume that s is positive. We set $u' := s \circ u \circ s^{-1} \in N$ and $u'(y_1) = y_2$. Then let r_n be the rotation of angle π/n and of axis a. Since $\mathbb{R}$ is Archimedean, the ratio π/n tends to 0 when n tends to $+\infty$ and therefore for n large enough $||x - r_n(x)|| \leq d$. We then set $x_0 = x$ and $x_{i+1} = r_n(x_i)$ with therefore $x_n = -x$. Since we have $||x_{i+1} - x_i|| \leq d$ there then exists $u_i \in N$ such that $u(x_i) = x_{i+1}$ so that $v = u_n \circ \cdots \circ u_1 \in N$ and $v(x) = -x$ and v is therefore a reversal, hence the result.

□

Exercice 16. *Let H denote the field of quaternions and let G be those of norm 1 : $G = \{a + bi + cj + dk \ / \ a^2 + b^2 + c^2 + d^2 = 1\}$. We then consider the action of G on H by interior automorphisms. By restricting this action to the set P of pure quaternions, show that we then obtain an isomorphism $G/\{\pm 1\} \simeq O(3, \mathbb{R})^+$. Is the associated exact sequence split ?*

Preuve : We have $P \simeq \mathbb{R}^3$ and we easily check that the conjugation action of G is $\mathbb{R}$-linear and preserves the norm so that it defines a group morphism $G \longrightarrow O(3, \mathbb{R})$. We also note that $G \simeq S^3$ is connected and that the previous morphism is continuous so that the image of $G \to O(3, \mathbb{R}) \to \{\pm 1\}$ is connected and therefore equal to $\{1\}$. We therefore obtain a group morphism $\phi : G \longrightarrow O^+(3, \mathbb{R})$. Let us show surjectivity : let $p \in P \cap G$, we have $\phi_p(p) = p$ which proves that ϕ_p fixes p (and is non-trivial), so it is a rotation of axis p. In addition we have $p^2 = -1$ or ϕ_p of order 2 ; it is therefore a reversal. We therefore obtain all the reversals, but these generate $O^+(3, \mathbb{R})$, hence surjectivity. For the kernel, we have $\phi_g(p) = p$ for all $p \in P$ if and only if g commutes to all elements of P and therefore to all elements of H, so $g \in \mathbb{R} \cap G = \{\pm 1\}$.

If the exact sequence

$$1 \to \{\pm 1\} \longrightarrow G \longrightarrow^{\phi} O^+(3, \mathbb{R}) \to 1$$

were split, we would have a subgroup H of G such that $\phi_{|H}$ is an isomorphism of H on $O^+(3, \mathbb{R})$. But then for $g \in G$, we would have g or $-g$ which would belong to H. Taking $o \in P \cap G$, we have $p^2 = (-p)^2 = -1$ or therefore $-1 \in H$, contradiction.

□

Exercice 17. *We consider the action of $G \times G$ on H defined by $(q_1, q_2).q := q_1 q \bar{q}_2$. Show that we thus define an isomorphism $G \times G/\{(1, 1), (-1, -1)\} \simeq O(4, \mathbb{R})^+$ and deduce that $PO(4, \mathbb{R})^+ \simeq O(3, \mathbb{R})^+ \times O(3, \mathbb{R})^+$.*

Preuve : The application ϕ_{q_1,q_2} is clearly $\mathbb{R}$-linear and preserves the norm. By continuity, we conclude as before that its image is contained in the positive isometries, i.e.

$$\phi : G \times G \longrightarrow O^+(4, \mathbb{R})$$

Let $(q_1, q_2) \in \operatorname{Ker} \phi$, i.e. $q_1 q \bar{q}_2 = q$ for all $q \in H$. For $q = 1$, we find $q_1 = q_2$ so that then q_1 is central and therefore $\operatorname{Ker} \phi = \{(1, 1), (-1, -1)\}$.

For surjectivity, let $u \in O^+(4, \mathbb{R})$, if we have $u(1) = 1$, as $P = 1^{\perp}$, we have $u(P) = P$ with $u_{|P} \in O^+(3, \mathbb{R})$ and from what there exists $q \in G$ such that $\phi_{q,q} = u$. If we have $u(1) = g$, we have then $\phi_{\bar{g},1} \circ u(1) = 1$ and we conclude thanks to the previous case. Finally we obtain

$$G \times G/\{(1, 1), (-1, -1)\} \simeq O(4, \mathbb{R})^+$$

By passing to the projective group, we look for the pairs (q_1, q_2) such that $\phi_{q_1,q_2} = -Id$, i.e. $q_1 q \bar{q_2} = -q$ for all $q \in H$. By making $q = 1$, we obtain $q_1 = -q_2$, then we see that q_1 is central, i.e.

$$G \times G/V \simeq PO(4, \mathbb{R})^+$$

where $V = \{(1,1), (1,-1), (-1,1), (-1,-1)\}$. Furthermore the canonical projection $G \to G/\{\pm 1\}$ induces an isomorphism

$$(G \times G)/V \simeq G/\{\pm 1\} \times G/\{\pm 1\}$$

and therefore according to the above

$$PO(4, \mathbb{R})^+ \simeq O(3, \mathbb{R})^+ \times O(3, \mathbb{R})^+.$$

□

Exercice 18. *(Tauvel p.417) Show that u is diagonalizable with real spectrum if and only if u is the product of two Hermitian endomorphisms, at least one of them being positive definite*

Preuve : Suppose that u is diagonalizable with real spectrum : let $(e_1, \cdots, e_n)$ be an orthonormal basis and let $(x_1, \cdots, x_n)$ be a basis formed by eigenvectors of u with real eigenvalues $\lambda_1, \cdots, \lambda_n$. We define f and g by $f(x_i) = e_i$ and $g(e_i) = \lambda_i e_i$ so that $u = f - 1 \circ g \circ f$. Let $f = qr$ be the polar decomposition of f with q unitary and r Hermitian positive definite. We thus obtain by setting $l = r \circ l \circ r^{-1} = q^* \circ g \circ q$ which is Hermitian, then $u = r^{-1} \circ l \circ r = (r^{-1} \circ l \circ r^{-1}) \circ r^2$ with therefore $r^{-1} \circ l \circ r^{-1}$ and r^2 Hermitian.

Conversely if $u = v \circ w$ with w (resp. v) Hermitian (resp. Hermitian positive definite). Let us note $l = v^{1/2}$ which is Hermitian positive definite, we have $u = l \circ (l \circ w \circ l) \circ l^{-1}$. Since $l \circ w \circ l$ is Hermitian, it is diagonalizable with real spectrum and so is u which is similar to it.

Exercice 19. *(Tauvel p.418) For $n \geq 2$ and A non-zero Hermitian of size n, show that A is positive definite or negative if and only if for any Hermitian matrix B, AB is diagonalizable.*

Preuve : The direct meaning follows from the previous exercise. Let us therefore assume that for any Hermitian matrix B, AB is diagonalizable. Let P be unitary such that $A = PDP^*$ with D real diagonal. Since $AB = P(DP^*BP)P^{-1}$, we see that AB is diagonalizable if and only if DP^*BP is. Moreover, since P^*BP is Hermitian, we can assume $A = D$.

We easily reduce to dimension 2 with $A = \mathrm{diag}(x^2, -y^2)$ with x, y real. Let then $B = \begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix}$ and $B' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. If $y \neq 0$, we have $AB \neq 0$ and $(AB)^2 = 0$. If $y = 0$ then $AB' \neq 0$ and $(AB')^2 = 0$. Thus AB (resp. AB') is non-zero nilpotent and therefore not diagonalizable.

Exercice 20. *A is normal if and only if A^* is a polynomial in A*

Preuve : The converse being obvious, let us show the direct sense. Let $(e_1, \cdots, e_n)$ be an orthonormal basis of eigenvectors for A : $Ae_i = \lambda_i e_i$. Similarly, we have $A^* e_i = \overline{\lambda_i} e_i$. We then construct a Lagrange interpolating polynomial P such that $P(\lambda_i) = \overline{\lambda_i}$ so that $A^* = P(A)$.

Exercice 21. *If A is normal, show that the orbit of A under the action of the unitary group is homeomorphic to $U(n)/U(k_1) \times \cdots \times U(k_r)$.*

Preuve : It suffices to note that the centralizer of A is isomorphic to $U(k_1) \times \cdots \times U(k_r)$ where the k_i are the dimensions of the eigensubspaces.

Exercice 22. *Show that two matrices are unitarily equivalent if and only if they have the same singular values*

Preuve : The direct sense follows from $PAP^* = B$, $PA^*P^* = B^*$ or $PAA^*P^* = BB^*$. Conversely, the polar decomposition gives $A = H_A U_A$, $B = H_B U_B$ with H_A, H_B positive-definite Hermitian and U_A, U_B unitary. We therefore have $AA^* = H_A H_A^*$ and $BB^* = H_B H_B^*$. The matrices H_A, H_B are diagonalizable so that if AA^* and BB^* have the same eigenvalues then the eigenvalues of H_A are equal to those of H_B up to the sign, i.e. $H_A = P_A D P_A^*$ and $H_B = P_B D D' P_B^*$ with $D' = \mathrm{diag}(\epsilon_1, \cdots, \epsilon_n)$ where the ϵ_i are equal to ± 1 so that D' is unitary. We therefore have $D = P_B^* U_B^* B P_B D'$ and $A = U_A P_A P_B^* U_B^* B P_B D' P_A^*$ hence the result.

Exercice 23. *(M-T p.187) The Householder matrices are exactly the matrices of $U(n)$ that are Hermitian of signature $(n-1, 1)$.*

Preuve : Recall that a Householder matrix associated with the column vector $v \in \mathbb{C}^n - \{0\}$ is $H(v) = I - 2\frac{vv^*}{v^*v}$. The matrix vv^* is Hermitian of rank 1 ; its eigenvalues are 0 at order $n-1$ and $\mathrm{tr}(vv^*) = v^*v$. Thus the eigenvalues of $H(v)$ which is clearly Hermitian and unitary are 1 at order $n-1$ and -1 at order 1.

Conversely, if H is a unitary Hermitian matrix, its eigenvalues are real with modulus 1 ; given the hypothesis on the signature and the fact that H is diagonalizable in an orthonormal basis $H = UDU^*$ with $D = \mathrm{diag}(-1, 1, \cdots, 1)$, i.e. $H = UH(e_1)U^{-1} = H(U(e_1))$.

5 List of possible projects

1. Infinite dimension with notably Hilbert, Banach and FrÃ©chet spaces
2. Uses of matrix conditioning
3. Fadaev's algorithm for computing the minimal polynomial : case of other similarity invariants.
4. Perron-Frobenius theorem, Markov chains and Google pagerank
5. Matrix exponential
6. $PSL_2(\mathbb{Z})$: generators and relations ; free groups and Banach Tarski paradox
7. Character theory
8. Compact subgroups of $GL_n(\mathbb{R})$
9. Simplicity of PSL_n and SO_3.
10. Endoscopy
11. Bistochastic matrices
12. Hadamard matrices