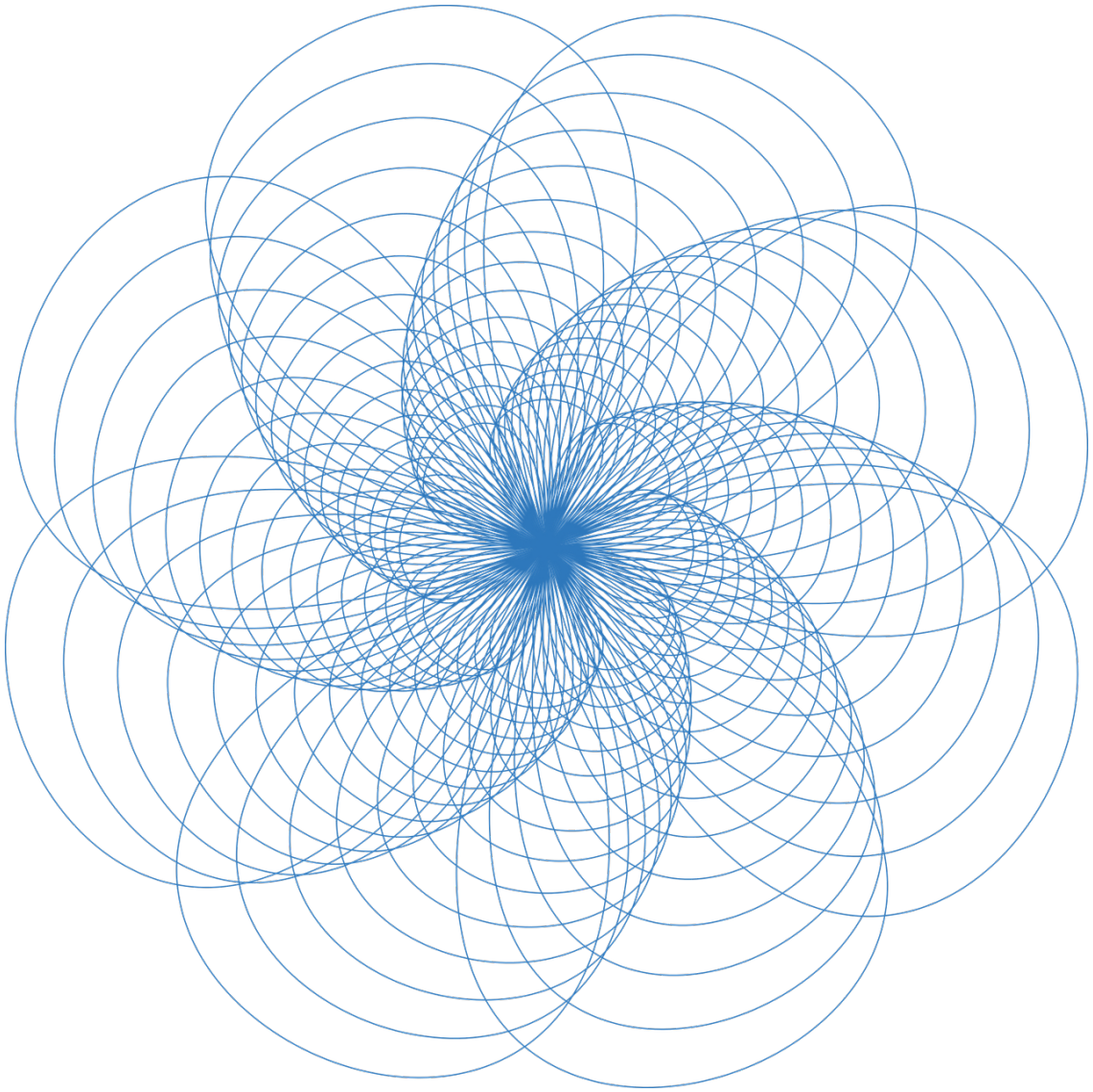
Mathématiques

Licence



Université Jean-François Champollion
Année 2022 - 2025

Table des matières

I — Raisonnements

Soit $\mathscr{P}$ une proposition et n un entier naturel.

Disjonctions & Conjonctions :

Si $\mathscr{P}$ est une disjonction de la forme $\mathscr{A} \vee \mathscr{B}$, il suffit alors de supposer **l'une des deux propriétés fausse** et de montrer que l'autre est vraie.

Si $\mathscr{P}$ est une conjonction de la forme $\mathscr{A} \wedge \mathscr{B}$, il faut simplement prouver $\mathscr{A}$ et $\mathscr{B}$.

Raisonnements par l'absurde :

Raisonner par l'absurde revient à utiliser le principe du **tiers exclu**, ie l'axiome qui affirme que la proposition ci-dessous est toujours vraie:

$$\mathscr{P} \vee \neg\mathscr{P}$$

Donc si on veut prouver $\mathscr{P}$, on peut alors simplement montrer que $\neg\mathscr{P} \implies \bot$ avec "$\bot$" comme notation d'une contradiction logique. Alors on peut conclure d'après l'axiome du tiers exclu que $\mathscr{P}$ est vraie.

Raisonnement par Analyse / Synthèse :

Le raisonnement par Analyse / Synthèse permet de déterminer **l'ensemble des solutions d'un problème**, il s'effectue en deux étapes, tout d'abord l'étape d'analyse suppose qu'une telle solution existe, alors on circonscrit son existence à des propriétés connues qu'elle vérifie nécessairement. Cette étape permet de "cerner" les solutions en question. Si les propriétés sont assez contraignantes, alors on peut même prouver **l'unicité**, ie l'ensemble des solutions se réduit à un singleton.

Puis lors de l'étape de synthèse, on considère un objet vérifiant les propriétés qu'on a utilisé lors de l'étape d'analyse, et on **vérifie** que cet objet est bien une solution au problème initial. C'est lors de cette étape qu'on prouve bien **l'existence** de solutions. Si aucun des objets circonscrits par l'analyse ne conviennent, le problème n'a alors pas de solutions.

Implications & Équivalences :

Si $\mathscr{P}$ est une implication de la forme $\mathscr{A} \implies \mathscr{B}$, on a les équivalences suivantes:

$$\mathscr{P} \iff \neg\mathscr{A} \vee \mathscr{B} \iff \neg\mathscr{B} \implies \neg\mathscr{A}$$

Aussi en raisonnant **par l'absurde**, il suffit alors de prouver:

$$\mathscr{A} \wedge \neg\mathscr{B} \implies \bot$$

Il est important de noter que l'implication **n'est pas une opération associative**, en effet, soit une propriété de la forme:

$$\mathscr{A}_1 \implies \mathscr{A}_2 \implies \mathscr{A}_3$$

Alors de manière générale, on a:

$$\mathscr{A}_1 \implies (\mathscr{A}_2 \implies \mathscr{A}_3) \not\iff (\mathscr{A}_1 \implies \mathscr{A}_2) \implies \mathscr{A}_3$$

Prouver une équivalence revient à prouver une **double implication** dans la majorité des cas.

Cas particulier : Si $\mathscr{P}$ est de la forme $A_1 \iff A_2 \iff \cdots \iff \mathscr{A}_{n-1} \iff \mathscr{A}_n$, il suffit alors de montrer:

$$\mathscr{A}_1 \implies \mathscr{A}_2 \implies \cdots \implies \mathscr{A}_{n-1} \implies \mathscr{A}_n \implies \mathscr{A}_1$$

Ainsi pour toute paire de $\mathscr{A}_i$, on a bien double implication entre les deux membres et donc la chaîne d'équivalence est démontrée.

Raisonnements par récurrence :

Soit $\mathscr{P}$ une propriété dépendante de n qu'on veut démontrer sur $[\![\alpha\ ;\ +\infty]\!]$, soit k en entier fixé supérieur à α, démontrer $\mathscr{P}$ par récurrence simple revient à utiliser **l'axiome de récurrence** (issu de la construction de $\mathbb{N}$) ci-dessous:

$$\Big[\mathscr{P}_\alpha \wedge \big[\mathscr{P}_k \implies \mathscr{P}_{k+1}\big]\Big] \implies \forall n \in \mathbb{N}\ ;\ \mathscr{P}_n$$

Si la propriété à prouver est plus complexe, on peut avoir besoin de récurrences d'une autre type, en effet si $\mathscr{P}$ dépend **des deux rangs précédents**, et on utilise alors une récurrence à deux pas qui s'exprime:

$$\Big[\mathscr{P}_\alpha \ \wedge\ \mathscr{P}_{\alpha+1} \wedge \big[\mathscr{P}_{k-1} \wedge \mathscr{P}_k \implies \mathscr{P}_{k+1}\big]\Big] \implies \forall n \in \mathbb{N}\ ;\ \mathscr{P}_n$$

Enfin pour le cas limite, si $\mathscr{P}$ dépends **d'exactement tout les rangs précédents**, alors on peut utiliser une récurrence forte qui s'exprime:

$$\Big[\mathscr{P}_\alpha \wedge\big[\mathscr{P}_{\alpha+1} \ \wedge\ \ldots\ \mathscr{P}_{k-1} \ \wedge\ \mathscr{P}_k \implies \mathscr{P}_{k+1}\big]\Big] \implies \forall n \in \mathbb{N}\ ;\ \mathscr{P}_n$$

Un dernier type de récurrence appelé **récurrence limitée** permet simplement d'utiliser la récurrence sur un intervalle entier fini, et donc on initialise et on prouve l'hérédité avec la contrainte de cette intervalle.

Remarque sur la récurrence forte :

Une telle récurrence forte ne nécessitera qu'une **unique** initialisation pour compléter l'hérédité.

D'un point de vue heuristique, il peut arriver d'engager une récurrence forte sur un problème qui n'aurait nécessité qu'une récurrence à p pas.
Ce cas précis reviendra alors, lors de l'étape d'hérédité, à **ne pas utiliser l'ensemble de l'hypothèse de récurrence**, et alors il faudra modifier le nombre d'initialisation à réaliser et l'intervalle de notre hypothèse de récurrence.

Admettons que $\mathscr{P}_\alpha$ soit vraie, supposons qu'elle soit vraie sur $[\![\alpha\ ;\ k]\!]$. Alors, on doit montrer que la propriété est vraie au rang $k+1$.

Alors, selon **le plus petit rang** nécessaire à compléter l'hérédité, on a:

Si on a besoin de $\mathscr{P}_k$ alors **on se ramène à une récurrence simple.**
Si on a besoin de $\mathscr{P}_{k-1}$ alors **on se ramène à une récurrence double.**
Si on a besoin de $\mathscr{P}_{k-2}$ alors **on se ramène à une récurrence triple.**
.....

Et donc, les initialisations et l'intervalle de notre hypothèse de récurrence changeront en conséquence et on remarque alors que si le plus petit rang nécessaire est $\mathscr{P}_{k-p}$, alors on se ramène nécessairement à une **récurrence à p pas**, avec p initialisations et l'hypothèse de récurrence qui commence à $\alpha+p$.

Une récurrence forte n'est alors qu'une récurrence qui nécessite des hypothèses sur **tout les rangs précédents.**

Récurrences imbriquées :

Soit $\mathscr{P}_{n,m}$ une propriété qui dépend **de deux variables entières**, alors on pourrait prouver $\mathscr{P}_{n,0}$ par récurrence et alors cela constituerait l'initialisation d'une récurrence imbriquée qui supposerait par exemple $\mathscr{P}_{n,k}$ vraie pour prouver $\mathscr{P}_{n,k+1}$.

I — Ensembles

Soit E un ensemble de parties. On munit cet ensemble des opérations élémentaires d'union, d'intersection ainsi que de complémentation définies pour toutes familles d'ensembles $(A_i)_I$ par:

- **Union:** $x \in \bigcup_i A_i \iff \exists i \in I \ ; \ x \in A_i$
- **Intersection:** $x \in \bigcap_i A_i \iff \forall i \in I \ ; \ x \in A_i$
- **Complémentaire:** $x \in A^c \iff x \notin A$

Ces opérations sont compatibles entre elles au sens où elle sont distributives l'une sur l'autre et associatives (pour une suite d'un seul type d'opération, si il y a mélange d'intersections et d'unions, on n'a pas associativité).

Inclusion :

On définit une **relation d'ordre** sur l'ensemble des parties de E appelée inclusion, elle est **réflexive, transitive et antisymétrique**. Si l'inclusion est stricte, on parle de **sous-ensemble propre**. En particulier, les opérations élémentaires préservent l'inclusion, en effet si $F \subseteq G$, on a:

$$F \cap X \subseteq G \cap X$$
$$F \cup X \subseteq G \cup X$$

Néanmoins la complémentation inverse l'inclusion:

$$F^c \subseteq E^c$$

Complémentaire et différence :

On peut montrer deux propriétés fondamentales du complémentaire apellées **lois de De Morgan** qui nous donnent:

$$(E \cap F)^c = E^c \cup F^c$$
$$(E \cup F)^c = E^c \cap F^c$$

On peut aussi raffiner la notion de complémentaire en définissant **la différence ensembliste** pour tout partie $F, G \subseteq E$, on pose:

$$F \setminus G = F \cap G^c$$

Cas particulier : On peut aussi définir l'opération de **différence symétrique** notée Δ qui permet d'obtenir tout les éléments qui appartiennent exactement à un seul des deux ensembles:

$$E \Delta F = (E \cup F) \setminus (E \cap F)$$

Produit cartésien :

Soit n une entier naturel, le produit cartésien des ensembles E_1, E_2, ... , E_{n-1}, E_n est l'ensemble des n-uplets de la forme $(e_1, e_2, \ldots, e_{n-1}, e_n)$ avec $e_i \in E_i$ pour $i \in [\![1 \ ; \ n]\!]$. Il y a **unicité** de ces n-uplets.
Plus formellement, on note:

$$\prod_{i=1}^{n} E_i = \Big\{ (e_1, e_2, \ldots, e_{n-1}, e_n) \ ; \ e_1 \in E_1, e_2 \in E_2, \ldots, e_n \in E_n \Big\}$$

Le produit cartésien est distributif sur l'union et l'intersection ie si on note $\star$ une de ces deux opérations, on a:

$$A \times (B \star C) = (A \times B) \star (A \times C)$$

Cardinalité :

Supposons que E et F tout deux inclus dans X et ayant **un nombre fini d'éléments**. On a alors différentes propriétés:

$$\begin{aligned} |E \cup F| &= |E| + |F| - |E \cap F| \\ |E \times F| &= |E| \times |F| \\ |E^c| &= |X| - |E| \\ |\mathcal{P}_E| &= 2^{|E|} \end{aligned}$$

Partitions et recouvrements:

Soit $(P_i)_{i\in\mathbb{N}}$ une famille de parties **non vides et deux à deux disjointes** de E.
On dit que (P_i) est une **partition** de E si et seulement si:

$$\bigcup_{i\in\mathbb{N}} P_i = E$$

On remarque immédiatement deux partitions singulières:

- La famille contenant uniquement E qu'on appelle **partition grossière**.
- La famille contenant tout les singletons de E qui est la partition **la plus fine**.

On peut donc intuitivement parler de **finesse** d'une partition, en regard de la taille des parties de la famille.

On peut généraliser le concept de partitition à celui de **recouvrement**, alors E ne nécessite que d'être contenu par l'union des (P_i).

Algèbre de Boole :

On peut montrer que l'ensemble **ordonné** des parties de E muni de l'union, l'intersection, le complémentaires forment une **Algèbre de Boole**.

Cela signifie que la structure $(\mathcal{P}(E), \cup, \cap, X^c)$ vérifie les axiomes suivants:

- Les deux opérations binaires sont **associatives, commutatives et distributives l'une sur l'autre.**
- Les deux opérations binaires sont **idempotentes.**
- **L'élément neutre** pour l'union est l'ensemble vide, et pour l'intersection l'ensemble E.
- **L'élément absorbant** pour l'union est l'ensemble E, et pour l'intersection l'ensemble vide.

- Le complémentaire est **involutif**.
- L'intersection d'un élément et de son complémentaire est **vide**.
- L'union d'un élément et de son complémentaire est **l'ensemble tout entier**.
- Les **lois de De Morgan** sont vérifiées.

De manière analogue, en considérant $\{0, 1\}$ comme les valeurs de vérité d'une proposition, on a:

La structure $(\{0, 1\}, \vee, \wedge, \neg)$ est aussi une algèbre de Boole.

Cette structure est à la base de la logique formelle et vérifie les même axiomes que l'algèbre de l'ensemble des parties d'un ensemble.

I — Relations

Une **relation** entre des objets d'un ensemble est une propriété que vérifient ces objets **entre eux**.
Les relations sont des objets **fondamentaux** en mathématiques, elles sont entre autres des objets primitifs de la théorie des ensembles.

On appelle **arité** le nombre d'éléments mis en jeu par la relation.
Par exemple une relation d'arité 2 est appelée **relation binaire** et met en jeu deux éléments. On définit ainsi le cas général de relation **n-aire** qui met en jeu n éléments $x_1, x_2, \ldots, x_{n-1}, x_n$ et on note:

$$\mathscr{R}(x_1, x_2, \ldots, x_{n-1}, x_n)$$

Par abus de langage, on appelle **classe** un ensemble d'ensembles.
Formellement une classe n'est pas un ensemble mais un élément primitif de la théorie ZFC, mais ici on verra qu'on appelle classe des objets qui **sont** des ensembles.

Zoologie :

Il existe un grand nombre de relations très connues et élémentaires, par exemple:

- La relation d'appartenance à un ensemble
- La relation d'égalité
- La relation d'ordre
- La relation d'inclusion
- La relation de congruence
- La relation de parallélisme de deux droites du plan

On peut remarque que la relation d'appartenance à un ensemble est une relation binaire fondamentale, à la base de la théorie des ensembles.

Relations binaires :

Soit $x, y, z \in E$, une relation entre deux éléments peut vérifier plusieurs propriétés remarquables:

- **Réflexivité :** $\mathscr{R}(x, x)$
- **Irréflexivité :** $\not\mathscr{R}(x, x)$
- **Symétrie :** $\mathscr{R}(x, y) \implies \mathscr{R}(y, x)$
- **Antisymétrie :** $\mathscr{R}(x, y) \wedge \mathscr{R}(y, x) \implies x = y$

Elle peut aussi être **transitive**:

$$\mathscr{R}(x, y) \wedge \mathscr{R}(y, z) \implies \mathscr{R}(x, z)$$

On appelle aussi relation **totale** une relation telles si pour toute paire d'éléments, on a $\mathscr{R}(x, y) \vee \mathscr{R}(y, x)$.

Relations d'ordre :

Une **relation d'ordre** est une relation **réflexive, antisymétrique et transitive**. Elle induit un ordre sur l'ensemble E, qui peut potentiellement être **total**.

Des relations d'ordre très connues sont la relation $\leq$ sur les ensembles de nombres ou la relation $\subseteq$ sur l'ensemble des parties de E.

On appelle relation de **préordre** toute relation relation d'ordre qui n'est pas antisymétrique. Intuitivement, une relation de préordre est une relation d'ordre à "équivalence près" des éléments.

Relations d'équivalence :

Une **relation d'équivalence** est une relation **réflexive, symétrique et transitive**. Intuitivement, elle met en relation les éléments des ensembles qui sont "similaires".

Des relations d'équivalence très connues sont la relation $=$ et $\equiv$ sur les ensembles de nombres, ou encore la relation $\sim$ sur l'ensemble des fonctions.

Classes d'équivalence :

Soit $(E, \sim)$ un ensemble muni d'une relation d'équivalence.
Les **classes d'équivalence** de E par rapport à la relation $\sim$ sont alors les parties de E contenant des éléments en relation.

Soit $x \in E$, on définit alors la **classe d'équivalence** de x et on note $[x]$ l'ensemble:

$$[x] := \Big\{ \alpha \in E \ ; \ \alpha \sim x \Big\}$$

D'après les propriété de la relation, on a alors:

$$\boxed{x \sim y \iff [x] = [y]}$$

Et on appelle **représentant** de $[x]$ tout élément qui appartient à $[x]$.

Ensembles quotient :

L'ensemble des classes d'équivalence de E forme alors une **partition** de E, et on l'appelle alors **ensemble quotient**:

$$E/_\sim \ := \Big\{ [x] \in \mathcal{P}(E) \ ; \ x \in E \Big\}$$

On a alors une application $\pi : x \in E \longmapsto [x] \in E/_\sim$ apellée **surjection canonique** qui associe sa classe à tout élément.

Travailler avec l'ensemble quotient revient alors à **identifier** les éléments équivalents entre eux. C'est une opération fondamentale dans tout les domaines des mathématiques, en effet par exemple on pourra comprendre (à l'aide du chapitre suivant) que si on a une application $f : E \longrightarrow F$ et qu'on définit sur E la relation d'équivalence:

$$x \sim y \iff f(x) = f(y)$$

Alors on a identifié les éléments génants l'injectivité et on a donc obtenu une bijection entre $E/\sim$ et F, ie on a le modèle primitif du **premier théorème d'isomorphisme**:

$$\begin{array}{ccc} E & \xrightarrow{\quad f \quad} & F \\ {\scriptstyle \pi}\downarrow & & \uparrow{\scriptstyle \iota} \\ E/_\sim & \xrightarrow{\quad \widetilde{\phi} \quad} & \mathrm{Im}(f) \end{array}$$

I — Applications

On appelle **application** un cas particulier de relation entre deux ensembles, soit f, g deux applications telles que:

$$\begin{aligned} f : E &\longrightarrow F & \qquad g : G &\longrightarrow H \\ x &\longmapsto f(x) & \qquad x &\longmapsto g(x) \end{aligned}$$

Si $F \subseteq G$, alors on définit **la composée** $g \circ f$ par la fonction $h : x \in E \longmapsto g(f(x)) \in H$. On note alors E^F l'ensemble des **applications** de E vers F.

Cas des suites :

Une suite à valeurs dans E n'est alors qu'un cas particulier en la forme d'une fonction $u : \mathbb{N} \longrightarrow E$, ce sont des objets d'étude trés importants en analyse et notamment en topologie. Dans le cas des suites on peut définir la notion de **suite extraite**, car si u_n est une suite dans E et k_n est **suite d'entiers croissante**, alors on définit une suite extraite de u_n par:

$$u \circ k : \mathbb{N} \longrightarrow E$$

C'est simplement les termes de la suite u_n dont on ne choisit que les termes d'indices donnés par k_n.

Graphe :

On définit **le graphe** de f comme suit:

$$G_f := \Big\{(x, f(x)) \in E \times F \ ; \ x \in E\Big\}$$

Intuitivement, c'est l'ensemble des couples d'éléments, des points, qui caractérise uniquement la fonction.

Restrictions & Prolongements :

On note $f|_A$ la restriction de **l'ensemble de départ** de f à une partie A de E.
On note $f|^B$ la restriction de **l'ensemble d'arrivée** de f à une partie B de F.

Soit $x \in D_f$, on appelle **prolongement** de f, l'application g telle que $D_f \subset D_g$ et $g(x) = f(x)$

Image directe :

Une fonction induit canoniquement une autre fonction sur l'ensemble des parties, notée aussi f, qui à chaque partie associe la partie **image directe**, ie pour toute partie A de E on définit:

$$f(A) := \Big\{f(x) \ ; \ x \in A\Big\}$$

L'image directe est compatible avec **certaines opérations ensemblistes**, plus précisément:

- **Intersection:** $f(A \cap B) = f(A) \cap f(B)$
- **Union:** $f(A \cup B) \subset f(A) \cup f(B)$

Image Réciproque :

Toute fonction induit aussi une autre fonction sur l'ensemble des parties qui à chaque partie associe la partie **image réciproque**, ie pour toute partie A de E on définit:

$$f^{-1}(B) := \Big\{x \in A \ ; \ f(x) \in B\Big\}$$

L'image réciproque est compatible avec **toutes les opérations ensemblistes**, plus précisément:

- **Intersection:** $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$
- **Union:** $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$

Injections :

L'application f est dite **injective** si et seulement si:

$$\forall x_1, x_2 \in E^2 \ ; \ f(x_1) = f(x_2) \implies x_1 = x_2$$

Avoir une injection de $E \longrightarrow F$ permet **d'identifier une partie de F à E**. Réciproquement, la non-injectivité représente le fait que la fonction détruit de l'information [1], ou alors que l'espace d'arrivée est trop petit.

Si on considère une composée $g \circ f$ injective alors on peut montrer que f est nécessairement injective. L'injectivité est stable par composition.

Surjections :

L'application f est dite **surjective** si et seulement si:

$$\forall y \in F \ , \ \exists x \in E \ ; \ f(x) = y$$

Avoir une surjection de $E \longrightarrow F$ permet **d'identifier une partie de E à F**. Réciproquement, la non-surjectivité indique que la fonction transporte trop peu d'informations, ou que l'espace d'arrivée est trop grand.

Si on considère une composée $g \circ f$ injective alors on peut montrer que g est nécessairement surjective. La surjectivité est stable par composition.

Bijections :

L'application f est bijective si et seulement si elle est surjective et injective. Dans ce cas, **une application réciproque** g existe et elle vérifie:

$$\begin{cases} f \circ g &= Id_F \\ g \circ f &= Id_E \end{cases}$$

Réciproquement, si il existe une application g telle que f soit inversible à gauche et à droite par g, alors f est bijective. Aussi la bijectivité est stable par composition.

Equipotence & Cardinalités :

On peut étendre la notion de cardinalité d'un ensemble fini via la notion d'application, en particulier on dira que pour tout ensembles A, B, alors:

- Si il existe une injection $f : E \longrightarrow F$, alors on a nécessairement $|E| \leq |F|$
- Si il existe une surjection $f : E \longrightarrow F$, alors on a nécessairement $|E| \geq |F|$

Si il existe une bijection de E vers F, alors on dit que ces ensembles sont **équipotents**, et on a:

$$|E| = |F|$$

Cette définition du cardinal par les bijections permet de parler de cardinal d'un ensemble dans le cas **infini**. En particulier:

- Si il existe une bijection entre $\mathbb{N}$ et E, on dit que E est **dénombrable**[2] et on note $|E| = \aleph_0$
- Si il existe une bijection de $\mathbb{R}$ dans E, alors on dit que E est **indénombrable** et on note $|E| = \aleph_1$

Dans notre cadre théorique (ZFC), il n'existe aucun ensemble dont le cardinal se situerait entre $\aleph_0$ et $\aleph_1$, c'est **l'hypothèse du continu**.

[1] Dans le sens où si deux valeurs différentes ont la même image, on ne peut plus les distinguer à l'arrivée.
[2] En fait, une injection suffit car on considérera par la suite que les ensembles finis sont dénombrables.

I — Dénombrement

Soit E un ensemble, on dit que E est **fini** si il existe une bijection de $[\![1\ ;\ n]\!]$ sur E.

On considère maintenant que E est fini, dénombrer E consiste à déterminer sa cardinalité. Informellement il s'agit souvent de compter le nombres **d'issues possibles** d'une situation donnée, on dispose alors de trois grands modèles, les **listes**, les **arrangements** et les **combinaisons**.

Listes

On appelle **liste** à p éléments de E un p-uplet constitué d'éléments de E, c'est à dire **un élément du produit cartésien** E^p on remarque alors la propriété:

Dans une liste, l'ordre compte et les répétitions sont possibles

On peut alors montrer que le nombre d'applications d'un ensemble à p éléments dans un ensemble à n éléménts est p^n

Arrangements

On appelle **arrangement** tout liste à p éléments **distincts** de E, on remarque alors:

Dans un arrangement, l'ordre compte mais les répétitions sont impossibles

On note alors A_n^p le nombre d'arrangements de p éléménts d'un ensemble à n éléments et on a:

$$A_n^p = \frac{n!}{(n-p)!}$$

Et on peut alors montrer que le nombre d'applications **injectives** d'un ensemble à p éléments dans un ensemble à n éléménts est A_n^p.

Un arrangement de la forme A_n^n est apellée une **permutation** de E qui est simplement donnée par $n!$, c'est aussi le nombre de **bijections** de E dans E.

Combinaisons

On appelle **combinaison** de p éléments tout **partie** de E à p éléments, on remarque alors:

Dans une combinaison, l'ordre ne compte pas et les répétitions sont impossibles

On apelle alors **coefficient binomial** et on note $\binom{n}{p}$ le nombre de parties à p éléménts d'un ensemble à n éléments et on a:

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}$$

On peut remarquer que le nombre de parties à p éléménts de E est exactement le nombre d'arrangements à p éléménts de E auquel on retire toutes les permutations des p éléments choisis, ce qui revient exactement à **retirer la contrainte d'ordre**.

Propriétés du coefficient binomial

Le coefficient binomial possède plusieurs propriétés intéréssantes, on peut tout d'abord remarquer une **symétrie** évidente mais aussi:

Formule de Pascal: $\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}$ **Formule du capitaine:** $p\binom{n}{p} = n\binom{n-1}{p-1}$

La **formule de Pascal** se comprends si on considère un élément fixé de l'ensemble et qu'on dénombreux tout ceux qui le contiennent, et les autres, ie:

Le nombre de parties à p éléments est exactement la somme du nombre de parties qui ne contiennent pas un certain x et du nombre de partie qui contiennent ce x.

La **formule du capitaine** se comprends si on considère le choix d'une équipe sportive de p joueurs (dont un capitaine) parmi un groupe de n candidats:

Choisir une équipe de p joueurs puis un capitaine parmi les p joueurs revient à choisir un capitaine parmi les n candidats, puis les $p-1$ joueurs restants.

Enfin on a aussi:

$$\sum_{p=0}^{n} \binom{n}{p} = 2^n$$

Le cardinal de l'ensemble des parties d'un ensemble à n éléments est donc exactement la somme des parties qui ont respectivement $1, 2, \ldots, n$ éléments.

Généralisation

On peut remarquer que le coefficient binomial est le nombre de partitions en deux parties de E telles que le cardinal de la première soit p. Par exemple si on considère les partitions de $E := \{1, 2, 3\}$ en deux parties dont la première ait 1 élément, on remarque qu'il y a 3 telles partitions:

$$P = \big(\{1\}, \{2, 3\}\big) \text{ ou } \big(\{2\}, \{1, 2\}\big) \text{ ou } \big(\{3\}, \{1, 2\}\big)$$

On peut alors généraliser cette idée et définir le **coefficient multinomial** $\binom{n}{k_1,\ldots,k_p}$ qui sera le nombre de partitions en p parties telles que la p-ième partie soit de cardinal k_p avec la somme des k_p **qui soit égale au cardinal total**:

$$\boxed{\binom{n}{k_1, \ldots, k_p} = \frac{n!}{k_1! k_2! \ldots k_p!}}$$

Pour fixer les idées on remarque que si $p = 2$ on a bien notre coefficient binomial usuel[1][2]:

$$\binom{n}{k_1, k_2} = \binom{n}{k_1, n-k_1} = \binom{n}{k_1} = \frac{n!}{k_1!(n-k_1)} = \frac{n!}{k_1! k_2!}$$

On peut alors utiliser ce coefficient multinomial, pour compter le nombre d'anagramme d'un mot de n lettres avec m lettres distinctes répétées k_m fois, ou encore le nombre de façon de mettre n objets dans m boites qui peuvent en contenir k_m.
Par exemple, le nombre d'anagrammes de MISSISSIPI est donné par $\binom{11}{1,4,4,1} = \frac{11!}{4!4!} = 34650$

On peut même pour définir la **formule du multinôme de Newton** qui généralise celle du binôme:

$$(x_1 + x_2 + \ldots + x_p)^n = \sum_{k_1+k_2+\ldots+k_p=n} \binom{n}{k_1, k_2 \ldots, k_n} x_1^{k_1} x_2^{k_2} \ldots x_p^{k_p}$$

[1] La première égalité vient de la contrainte sur la somme des k_p.
[2] La seconde égalité se comprends par symétrie, compter le nombre de partitions en deux parties dont la première contient k_1 éléments revient à compter le nombre de parties à k_1 éléments et le reste sera nécessairement dans la seconde partie.

I — Arithmétique élémentaire

Dans ce chapitre on énonce quelques définitions et propriétés arithmétiques simples dans $\mathbb{Z}$, qui seront généralisées plus tard dans le chapitre d'algèbre au cas général. Dans cet ensemble on peut définir une relation d'ordre de divisibilité définie par:

$$a|b \iff \exists k \in \mathbb{Z}\ ;\ b = ak$$

On dira alors que b est un multiple de a et que a divise b. On notera $\mathcal{D}(n)$ l'ensemble des diviseurs de n et $n\mathbb{Z}$ l'ensemble de ses multiples. Une première propriété trés utile de cette relation et que si a divise b, c alors pour tout $n, m \in \mathbb{Z}$ on a:

$$a|nb + mc$$

Division Euclidienne :

Soit $a, b \in \mathbb{Z} \times \mathbb{Z}^*$, on peut montrer qu'il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ avec $r < |b|$ tel que:

$$a = bq + r$$

On appelle alors cette décomposition **la division euclidienne** de a par b. La preuve se fait par l'exhibition de l'algorithme bien connu.

Plus grand diviseur commun :

Soit $a, b \in \mathbb{Z}$ non simultanément nuls, alors le pgcd est l'entier $a \wedge b$ qui vérifie:

$$a \wedge b := \max\left\{n \in \mathbb{N}\ ;\ n|a \text{ et } n|b\right\}$$

Alors on l'appelle **plus grand diviseur commun** de a et de b et on le note $a \wedge b$. On peut alors monter plusieurs propriétés de cette quantité:

- **Maximalité:** Si d est un diviseur commun de a, b alors $d|a \wedge b$.
- **Réduction:** On peut réduire le pgcd par division euclidienne, ie $a \wedge b = b \wedge r$.

On peut alors définir la notion de deux entier n, m **premiers entre eux** par le fait que $n \wedge m = 1$.

Plus petit commun multiple:

Soit $a, b \in \mathbb{Z}$ non simultanément nuls, alors le ppcm est l'entier $a \vee b$ qui vérifie:

$$a \vee b := \min\left\{n \in \mathbb{N}\ ;\ a|n \text{ et } b|n\right\}$$

Alors on l'appelle **plus petit commun multiple** de a et de b et on le note $a \vee b$. On peut alors monter plusieurs propriétés de cette quantité:

- **Minimalité:** Si d est un multiple commun de a, b alors c'est un multiple de $a \vee b$.

Identité de Bézout :

Soit $a, b \in \mathbb{Z}^2$, on peut montrer par une extension de l'algorithme d'Euclide appelé **algorithme d'Euclide étendu**[1] qu'il existe deux entiers $u, v \in \mathbb{Z}^2$ tels que:

$$au + bv = a \wedge b$$

Il existe donc une combinaison linéaire (à coefficients entiers) de a, b qui donne leur PGCD.

[1] En effet remontrer l'algorithme par subsitution permet d'écrire le dernier reste non nul, le pgcd, comme combinaison linéaire des deux nombres de départ.

Lemme de Gauss :

Soit 3 entiers $a, b, c \in \mathbb{Z}$, alors grâce à l'identité de Bézout, on peut montrer le **lemme de Gauss**:

$$\begin{cases} a \mid bc \\ a \wedge b = 1 \end{cases} \implies a \mid c$$

Nombres premiers :

On appelle nombres premiers tout nombre différent de 1 qui n'admet aucun diviseur. Il existe une infinité de nombres premiers et on a le **théorème de décomposition**:

$$\forall n \in \mathbb{Z}\ ;\ n = \prod_{p \in \mathcal{P}} p^{v_p(n)}$$

Où $v_p(n) = \max\left\{k \in \mathbb{N}\ ;\ p^k | n\right\}$ est apellée **valuation p-adique** de n.

Indicatrice d'Euler :

En algèbre, il sera utile de connaître **le nombre d'entiers inférieurs à n et premiers avec n**, pour ceci on définit **la fonction indicatrice d'Euler** par:

$$\begin{aligned} \varphi : \mathbb{N} &\longrightarrow \mathbb{N} \\ n &\longmapsto n \prod_{p \mid n} \left(1 - \frac{1}{p}\right) \end{aligned}$$

Le produit se faisant sur tout les diviseurs premiers distincts de n. L'utilité de cette fonction vient de la propriété suivante que justement $\phi(n)$ est exactement le nombre d'entiers inférieurs à n et premiers avec n.

Exemple: $\varphi(30) = \varphi(2 \times 3 \times 5) = 30 \times \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right) = 30 \times \frac{1}{2} \times \frac{2}{3} \times \frac{4}{5} = 8$

II — Introduction

On appelle **structure algébrique** un ensemble muni d'une (ou plusieurs) opérations appelées **lois**, c'est l'étude de telles structures mathématiques, des relations entre celles-ci (que nous appeleront morphismes), et de leurs propriétés que nous appeleront **algèbre générale**.

Soit E un ensemble non-vide, on appelle **loi de composition interne** une opération binaire sur les éléments de E (qu'on notera temporairement $\star$) telle que:

$$\forall a, b \in M \ ; \ a \star b \in E$$

Soit K un ensemble non-vide, on appelle **loi de composition externe** une opération binaire entre un élément de K et un élément de E (qu'on notera temporairement $\cdot$) telle que:

$$\forall \lambda, a \in K \times E \ ; \ \lambda \cdot a \in E$$

Soit $a \in E$, alors on peut aussi rencontrer dans les structures usuelles des éléments remarquables qui peuvent exister ou non:

- On dira que $e \in E$ est un **élément neutre** pour la loi si $\forall a \in E \ ; \ a \star e = e \star a = a$
- On dira que $a^{-1} \in E$ est **l'inverse** de a pour la loi si $a^{-1} \star a = a \star a^{-1} = e$

Ces éléments, si ils existent, sont alors **uniques**.

Monoides :

Soit M un ensemble qu'on munit d'une **loi de composition interne**, alors le couple $(M, \star)$ est appelé un **magma**, c'est la structure algébrique primitive la plus faible, en effet la seule contrainte étant que la loi soit interne.

On peut alors enrichir la structure de magma par les deux contraintes supplémentaires suivantes:

- La loi est **associative**.
- Il existe **un élément neutre** pour la loi.

Cette structure plus riche, qu'on appelle **monoide** nous permet alors d'identifier des exemples remarquables:

- Les entiers naturels munis de l'addition forment un monoide.
- L'ensemble des chaines de caractères muni de la concaténation forme un monoide.

Les éléments neutres respectifs de ces exemples sont $0_{\mathbb{N}}$ et la chaine de caractère vide.

Sous-structures :

Une fois une structure algébrique définie sur E, on peut alors s'intéresser aux parties de E qui conservent cette structure, on les appellera alors **sous-structures** de E.

En particulier, on dira que F est une **sous-structure** de E (et on notera $F < E$) si elle vérifie:

- La partie F est stable par les lois.
- Les éléments neutres[1] appartient à F
- Les inverses[2] des éléments de F appartient à F

On montre alors facilement que **l'intersection** de deux sous-structures est aussi une sous-structure mais que l'union de deux sous-structures n'est en général pas une sous-structure.

[1] **Si la structure impose leur existence**
[2] **Si la structure impose leur existence**

Sous-structure engendrée :

On se donne une partie A de E, on peut alors définir la **sous-structure engendrée** par A. Si on considère $(V_i)_{i\in I}$ la famille des sous-structures de E qui contiennent A, alors on pose:

$$\langle A\rangle = \bigcap_{i\in I} V_i$$

C'est alors clair que c'est la plus petite sous-structure (pour l'inclusion) qui contienne A et on peut alors la caractériser par la propriété suivante:

C'est l'ensemble des combinaisons finies obtenues par applications des lois sur des éléments de A.

Morphismes :

Soit $(E,\star)$ et $(F,\cdot)$ deux ensembles munis de la meme structure[1] et $\varphi : M \to N$, alors φ est appelé **morphisme**, si il vérifie:

$$\forall x,y \in E\ ;\ \varphi(x\star y) = \varphi(x)\cdot\varphi(y)$$

Les morphismes préservent dans une certaine mesure la structure opératoire.

En termes de vocabulaire, on définit alors:

- **Les endormorphismes** commme les morphismes de M dans lui-meme.
- **Les isomorphismes** commme les morphismes bijectifs.
- **Les automorphismes** commme les morphismes bijectifs de M dans lui-meme.

La recherche d'isomorphismes est un thême principal en algèbre des structures, en effet, trouver un isomorphisme entre une structure simple et une structure complexe permet de mieux comprendre cette dernière par l'intermédiaire du morphisme.

Propriétés des morphismes :

Pour une structure donnée, on peut montrer que la composée de morphismes et l'inverse d'un morphisme bijectif est un morphisme. En outre on peut caractériser la structure des images directes et réciproques par un morphisme:

L'image et la préimage d'une sous-structure par un morphisme est une sous-structure.

Par ailleurs si $F = \langle f_1,\dots,f_n\rangle$ est un sous groupe de E et que ϕ est un morphisme, on a que:

$$\phi(H) = \langle \phi(h_1),\dots,\phi(h_n)\rangle$$

En d'autres termes, **l'image des générateurs engendre l'image**.

Structures Quotients :

On considère maintenant un ensemble quotient $E/_\sim$ tel que E soit muni d'une structure, on cherche alors une condition sur la relation d'équivalence pour que **la structure soit conservée au passage au quotient**. On peut alors montrer que c'est le cas si et seulement si $\sim$ est **compatible** avec les lois, ie que pour toute loi $\star$, on ait:

$$x_1 \sim x_2 \text{ et } y_1 \sim y_2 \implies x_1\star y_1 \sim x_2 \star y_2$$

Alors la **surjection canonique** est un morphisme $\pi : E \to E/\sim$ qui a chaque élément associe sa classe d'équivalence pour la relation.

[1] Si les structures présentent plusieurs lois, alors les morphismes doivent vérifier la compatibilité pour **toutes les lois**. Aussi dans le cas particulier de structures qui requièrent l'existence d'un élément neutre, l'image de l'élement neutre de la structure de départ doit être celui de celle d'arrivée.

II — Groupes

Soit G un ensemble **non-vide** muni d'une loi de composition interne associative[1] telle que:

- Il existe **un élément neutre** pour la loi.
- Tout élément de G admet **un inverse** pour la loi.

Alors le couple $(G, \star)$ est appellé **groupe**. De plus si le groupe est **commutatif**, on dira alors que c'est un groupe **abélien**.

On appellera **ordre du groupe** le cardinal (potentiellement infini) de l'ensemble sous-jacent, noté $|G|$.

Exemples :

On peut alors considérer plusieurs groupes remarquables:

- Les **entiers relatifs** muni de l'addition usuelle.
- Les **isométries du plan** muni de la composition, on l'apelle le **groupe dihédral**.
- Les **matrices inversibles** muni de la multiplication, on l'apelle le **groupe linéaire**.
- Les **bijections** sur un ensemble muni de la composition, on l'apelle le **groupe symétrique**.

Morphismes de groupes :

Soit G, H deux groupes et $\varphi : G \to H$, l'existence d'un élément neutre nous permet de définir alors **le noyau d'un morphisme** par:

$$\mathrm{Ker}(\varphi) := \Big\{ x \in G \ ; \ \varphi(x) = e_H \Big\}$$

On montre facilement que c'est un sous-groupe (normal) et on peut alors montrer qu'un morphisme est **injectif si et seulement si son noyau est réduit à l'élément neutre.**

Sous-groupes :

Les sous-structures dans le cas des groupes sont naturellement les sous-groupes. On peut alors caractérisé **le sous-groupe engendré** par H (défini au premier chapitre) par:

$$\langle H \rangle := \Big\{ h_1^{k_1} h_2^{k_2} \dots h_n^{k_n} \ ; \ n \in \mathbb{N} \ , \ h_i \in H \ , \ k_i \in \mathbb{Z} \Big\}$$

On peut alors considérer le sous-groupe engendré par un élément $h \in H$, en effet on a:

$$\langle h \rangle := \Big\{ h^k \ ; \ k \in \mathbb{Z} \Big\}$$

On peut alors définir **l'ordre d'un élément** comme étant l'ordre du sous-groupe engendré associé (potentiellement infini).

Ce sous-groupe permet de définir des groupes remarquables, en effet si un groupe est engendré par un unique élément, il est appelé **groupe cyclique** dont nous parleront plus loin dans ce chapitre.

[1] Dans la suite, la loi de composition des groupes sera notée multiplicativement sauf exceptions.

Classes :

On considère maintenant un sous-groupe $H \leq G$, alors on peut définir deux relations d'équivalences sur G par:

$$\begin{cases} g_1 \sim g_2 \iff \exists h \in H \ ; \ g_1 = g_2 h \\ g_1 \sim g_2 \iff \exists h \in H \ ; \ g_1 = h g_2 \end{cases}$$

On appelle alors **classe à gauche** (resp. classe à droite) les classes d'équivalences pour ces deux relations et on note alors gH (resp. Hg) la classe d'un élément g pour cette relation. On note alors G/H l'ensemble quotient associé aux classes à gauche.

Théorème de Lagrange :

Ces classes induisent donc une partition de G en classes **de même cardinal**, en effet:

$$|gH| = \left|\{gh \ ; \ h \in H\}\right| = |H|$$

En outre on a une bijection qui associe à chaque élément de g sa classe et l'élément de H lui correspondant:

$$\begin{aligned} f : G &\longrightarrow (G/H, H) \\ g &\longmapsto (aH, h) \end{aligned}$$

Ceci nous permet donc de montrer le **théorème de Lagrange** qui nous donne que pour tout groupe fini G, on a:

$$|G| = |G/H||H|$$

Et comme corollaire immédiat la propriété suivante:

Le cardinal d'un sous-groupe divise le cardinal du groupe.

Sous-groupes normaux :

On cherche alors à caractériser les sous-groupes tels que la relation d'équivalence définie ci-dessous soit **compatible** avec les opération de groupe, en d'autres termes on cherche à définir un groupe quotient pour cette relation. On peut alors montrer que les sous-groupes vérifiant cette compatibilité vérifient:

$$\forall g \in G \ ; \ gH = Hg$$

En d'autres termes les classes à droite et à gauche coincident. C'est alors immédiat que **tout sous-groupe d'un groupe abélien est normal**. Par ailleurs on peut caractériser les sous-groupes normaux d'une autre façon (détaillée au chapitre sur les actions de groupe) comme les sous-groupes qui vérifient:

$$\forall h \in H \ , \ \forall g \in G \ ; \ ghg^{-1} \in H$$

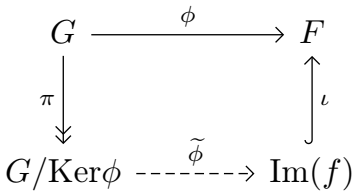
La propriété fondamentale de ces groupes, qui utilise le premier résultat du chapitre suivant est que les sous groupes normaux de G sont exactement les **noyaux** de morphismes de domaine G.

II — Théorèmes d'isomorphismes

Une des motivations de la notion de groupe quotient est entre autres de pouvoir trouver des **isomorphismes** entre des groupes connus, dans ce chapitre, on énonce les trois grands théorèmes utilisables pour atteindre cet objectif.

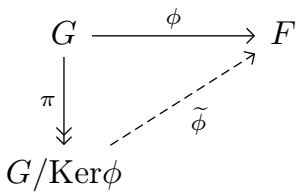
Premier théorème d'isomorphisme :

Soit $\phi : G \longrightarrow F$ un morphisme, on rapelle que tout les noyaux sont normaux et on peut alors montrer qu'il existe un unique isomorphisme $\widetilde{\phi} : G/\mathrm{Ker}\phi \longrightarrow F$ tel que le diagramme soit commutatif [1]:



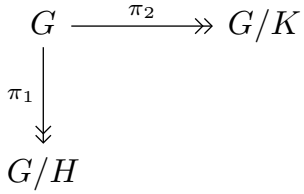
En effet, le passage au quotient rend le morphisme injectif, donc surjectif sur son image, et le diagramme commute, ie on a $\phi = \iota \circ \widetilde{\phi} \circ \pi$.

De manière plus générale, on a la **propriété universelle du quotient** pour $H \trianglelefteq G$ tel que $H \subseteq \ker(\phi)$, alors on a l'existence d'un morphisme $\widetilde{\phi}$ tel que le diagramme suivant commute:

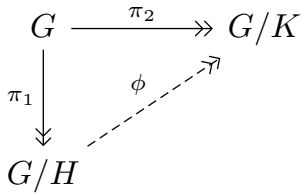


Deuxième théorème d'isomorphisme :

On considère ici deux sous groupe normaux H, K de G tel que $H \subseteq K$, alors on a les deux projections suivantes:

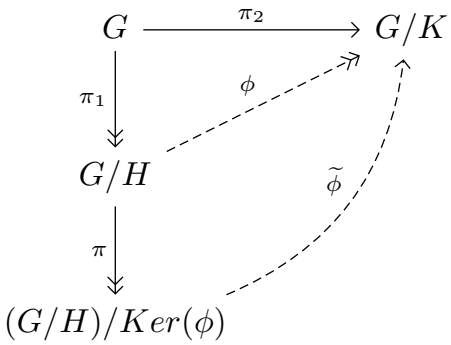


On peut alors utiliser la propriété universelle du quotient pour compléter le diagramme par un morphisme ϕ (par ailleurs surjectif):



[1] Un **diagramme commutatif** est une collection d'objets et de morphismes tels tout les chemins (de composition) partant d'un objet vers un autre donnent le meme résultat (ie sont le meme morphisme).

Enfin, on peut appliquer le premier théorème d'isomorphisme à ϕ pour obtenir le diagramme suivant:



On peut alors montrer que $\mathrm{Ker}(\phi) = K/H$ et donc qu'on a l'isomorphisme suivant:

$$(G/H)/(K/H) \cong G/K$$

Troisième théorème d'isomorphisme :

Caractère universel :

Le parti pris a été fait de mettre cette section dans le chapitre sur les groupes, mais ceci est trompeur, les trois théorèmes ci-dessus sont en fait vrais dans un cadre bien plus général, et pour des objets bien plus généraux apellés **algèbres universelles**, en particulier tout structure sur laquelle on peut définir une notion de quotient compatibles avec les opérations vérifie alors des analogues de ces théorèmes. En particulier:

- On peut quotienter un ensemble par la relation d'équivalence "avoir la même image" et obtenir alors de tels théorèmes.
- On peut quotienter un anneau par un **idéal** et obtenir alors de tels théorèmes.
- On peut quotienter un espace vectoriel (ou même un module) par un sous-espace et obtenir alors de tels théorèmes.

Applications :

II — Actions de groupe

Soit G un groupe et X un ensemble quelconque, dans ce chapitre on définit un notion fondamentale en théorie des groupes, la notion **d'action d'un groupe sur un ensemble.** En effet on appelera **action** du groupe G sur X une application de la forme:

$$\begin{aligned} G \times X &\longrightarrow X \\ (g, x) &\longmapsto g \cdot x \end{aligned}$$

En outre une action doit vérifier deux autres propriétés:

- **Le neutre n'agit pas:** $\forall x \in X \ ; \ e \cdot x = x$
- **Associativité mixte:** $\forall g_1, g_2, x \in G \times G \times X \ ; \ (g_1 g_2) \cdot x = g_1(g_2 \cdot x)$

On dira alors que G **agît** sur X et on notera alors $G \curvearrowright X$.

Morphisme structurel:

On se donne une action $G \curvearrowright X$, alors il peut être utile de considérer la currifiée[1] de cette action, ie:

$$\begin{aligned} \phi : G &\longrightarrow (X \longrightarrow X) \\ g &\longmapsto (x \longmapsto g \cdot x) \end{aligned}$$

On peut alors montrer que cette fonction prends son image dans l'ensemble des bijections sur X (dont on montrera que c'est un groupe au chapitre sur le groupe symétrique) et que c'est un **morphisme de groupe**. L'action de G induit donc un morphisme de groupe, appelé **morphisme structurel** de la forme:

$$\phi : G \longmapsto \mathfrak{S}(X)$$

En outre cette correspondante est bijective, il est donc équivalent de considérer une action d'un groupe sur un ensemble ou un morphisme structurel.

Action induite sur l'ensemble des parties :

Si G agît sur X alors G agît alors naturellement sur $\mathcal{P}(X)$ par l'action:

$$(g, P) \mapsto g \cdot P := \{g \cdot x \ ; \ x \in P\}$$

Action induite sur les sous structures:

On se pose alors deux questions naturelles:

- Une action de G sur X induit-elle nécessairement une action de G sur $Y \subseteq X$?
- Une action de G sur X induit-elle nécessairement une action de $H \leq G$ sur X ?

On peut alors montrer que la première question admet une réponse positive si et seulement si Y est **stable par l'action**.

Pour la seconde question, elle admet toujours une réponse positive et on a même le résultat général suivant gràce au morphisme structurel, on considère deux groupes G, H reliés par un morphisme ϕ, et une action de H sur X de morphisme structurel ψ, alors on a le diagramme:

$$G \xrightarrow{\ \phi\ } H \xrightarrow{\ \psi\ } X$$

Et donc $\phi \circ \psi$ définit bien un morphisme structurel de G sur $\mathfrak{S}(X)$ et donc une action. Le cas particulier des sous-groupes se déduit en considérant ϕ le morphisme d'inclusion d'un sous-groupe dans le groupe total.

[1] On rapelle que $\mathcal{F}(E \times F, G) \cong \mathcal{F}(E, \mathcal{F}(F, G))$ en tant qu'ensembles.

Orbites :

Considérons un point $a \in X$, alors on définit **l'orbite** de a sous l'action du groupe G par:

$$\mathrm{Orb}_G(a) := \{g \cdot a \ ; \ g \in G\}$$

Intuitivement, ce sont tout les points atteints par l'action de G sur le point initial a. Une propriété fondamentale des orbites est la suivante, si on considère la relation suivante:

$$x \sim y \iff y \in \mathrm{Orb}_G(x)$$

Alors c'est une **relation d'équivalence**, et on a donc toujours une **partition** de X associée à l'action de G, c'est la partition en orbites.

Stabilisateurs :

Considérons un point $a \in X$, alors on définit **le stabilisateur** de a sous l'action du groupe G par:

$$\mathrm{Stab}_G(a) := \big\{g \in G \ ; \ g(a) = a\big\}$$

Intuitivement, ce sont tout les éléments du groupe qui laissent a invariant. Une propriété fondamentale des stabilisateurs est que c'est un **sous-groupe** du groupe G. En outre si on considère le morphisme structurel ϕ de l'action, on a:

$$\mathrm{Ker}(\phi) = \bigcap_{x \in X} \mathrm{Stab}_G(x)$$

Généralisations aux parties :

On peut alors noter qu'il est aussi possible de définir les orbites et stabilisateurs de **parties**, en considérant les orbites et stabilisateurs pour l'action induite sur les parties définie plus haut.

Vocabulaire :

On peut alors nommer les actions de groupes qui vérifient certaines propriétés relatives aux ensembles définis plus haut, on appelle alors:

- Action **transitive** une action qui n'admet qu'une seule orbite.
- Action **libre** une action dont tout les stablisateurs sont triviaux.
- Action **fidèle** une action dont le noyau du morphisme structurel est trivial[1].

On dira aussi qu'une action transitive et libre est **simplement transitive**, et on peut caractériser cette action par le fait que pour tout paire d'éléments $x, y \in E$, il existe un **unique** élement de g qui relie x à y.

Action par automorphismes intérieurs:

On peut alors aussi étudier l'action du groupe G sur **lui-même**, on obtient alors un nouveau moyen d'étude du groupe G, en particulier, on a deux actions remarquables:

- **L'action par translation:** $\forall g, h \in G, g \cdot h = gh$
- **L'action par conugaison:** $\forall g, h \in G, g \cdot h = ghg^{-1}$

Ceci permet une reformulation plus élégante du concept de sous-groupe normal, en effet un sous-groupe est normal si et seulement si il est **stable par l'action de conjugaison**.

[1] On a alors d'aprés la caractérisation du noyau ci-dessus que toute action **libre** est **fidèle**.

Centralisateur:

Le stabilisateur d'un élément g pour la relation de conjugaison est alors appelé **centralisateur** et noté $Z(g)$, et c'est l'ensemble des éléments qui commutent avec g.

On peut définir le centralisateur d'une partie, noté $Z(H)$ qui est l'intersection de tout les centralisateurs de ses éléments, ie l'ensemble des éléments du groupe qui commutent avec tout les éléments de H, ie on a:

$$Z(H) := \{g \in G \ ; \ \forall h \in P \ , \ gh = hg\}$$

En particulier pour tout groupe G, on appelle **centre** du groupe et on note $Z(G)$, l'ensemble des éléments qui commutent avec tout les autres éléments.

Normalisateur:

En affaiblissant la définition ci dessus, on peut définir le **normalisateur** d'une partie H, noté $N(H)$, et c'est le stabilisateur de l'action par la conjugaison sur les **parties**, ie:

$$N(H) := \{g \in G \ ; \ gH = Hg\}$$

Relation orbites stabilisateurs:

On peut alors montrer que si on fixe $x \in X$, alors il existe une bijection entre $G/\text{Stab}(x) \longrightarrow \text{Orb}(x)$ et en particulier, on a alors la relation fondamentale suivante dite **relation orbites-stabilisateurs**:

$$|G| = \big|\text{Orb}(x)\big| \, \big|\text{Stab}(x)\big|$$

Et donc en particulier, le cardinal d'une orbite (ou d'un stabilisateur) **divise l'ordre de** G.

Formules des classes:

On peut alors utiliser le fait que X se partitionne en orbites pour obtenir une expression du cardinal de X apellée **formule des classes** où n désigne le nombre d'orbites:

$$|X| = \sum_{i=1}^{n} \big|\text{Orb}(x_i)\big| = \sum_{i=1}^{n} \frac{|G|}{\big|\text{Stab}(x_i)\big|}$$

Un des intérêts de cette formule est par exemple qu'elle permet de connaître le nombre d'orbites d'une action ou de montrer l'existence de points fixes (ie de points dont l'orbite est de cardinal 1), en effet on considère les diviseurs de l'ordre du groupe $(d_1, \ldots, d_k)$ et $(a_1, \ldots, a_k)$ le nombre d'orbites de cette taille, on obtient alors une equation de la forme suivante, qui peut souvent s'étudier facilement dans les cas simples avec peu de diviseurs:

$$|X| = \sum a_k d_k$$

Formules de Burnside:

Une autre formule important liée aux actions de groupe est **la formule de Burnside** qui permet de dénombrer les orbites de l'action, et en particulier, on peut alors **compter des éléments modulo une action de groupe**, on a la formule suivante:

$$n = \frac{1}{|G|} \sum_{g \in G} |\text{Fix(g)}|$$

Où $\text{Fix(g)} := \{x \in X \ ; \ g \cdot x = x\}$ est l'ensemble des points fixés par g. Cette formule est fondamentale en combinatoire, par exemple imaginons que nous souhaitions compter le nombre de colliers **différents** de 5 perles à deux couleurs. Alors ici "différents" signifie que un des colliers dénombré est égal à un autre aprés une rotation ou une reflexion, on considèrera ce collier comme le même que le premier.

L'idée principale est donc bien de compter des éléments modulo l'action sur l'ensemble, ie en identifiant deux éléments dans la même orbite.

Compter ces colliers revient donc à compter les orbites de l'action par symétries d'un groupe sur l'ensemble de tout les colliers possibles. Et la formule de Burnside nous permet donc d'effectuer ce calcul.

II — Groupes Symétriques

On appelle **groupe symétrique** et on note $\mathfrak{S}_n$ le groupe des **permutations** de l'ensemble $[\![1\ ;\ n]\!]$ muni de la composition des applications.

On remarque alors aisément que l'ordre de $\mathfrak{S}_n$ est $n!$.

Soit $\sigma \in \mathfrak{S}_n$ une permutation de $[\![1\ ;\ n]\!]$, alors c'est une fonction bijective sur cet ensemble. En particulier, sachant que l'ensemble est fini, c'est une fonction définie par cas qu'on note alors par commodité horizontalement dans un tableau:

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

Support :

On appelle alors **support** d'une permutation le complémentaire des points fixes de σ, ie on a:

$$\mathrm{Supp}(\sigma) := \{i \in \mathbb{N}\ ;\ \sigma(i) \neq i\}$$

Une des propriétés fondamentale qu'on peut déduire de cette définition est que **deux permutations à supports disjoints commutent**.

Cycles :

On appelle **k-cycle** une permutation σ telle qu'il existe $k \geq 2$ et k éléments deux à deux distincts $a_1, \dots, a_k$ tels que:

$$\begin{cases} \forall i \in [\![1\ ;\ k-1]\!]\ ;\ \sigma(a_i) = \sigma(a_{i+1}) \\ \forall i \notin [\![1\ ;\ k]\!]\ ;\ \sigma(a_i) = \sigma(a_i) \\ \sigma(a_k) = \sigma(a_1) \end{cases}$$

Un k-cycle laisse fixe tout les éléments sauf pour une certaine famille (a_i) pour laquelle chaque élément est envoyé sur le suivant.

On peut alors noter un tel cycle par la notation suivante qui décrit tout les éléments affectés par la permutation:

$$\sigma = (a_1, \dots, a_n)$$

Le cas particulier des 2-cycles est intéressant, en effet un 2-cycle **échange deux valeurs** de $[\![1\ ;\ n]\!]$, ils sont d'une importance particulière et on les appelle **transpositions**.

Exemple: La permutation suivante est un 3-cycle:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} = (1\ 2\ 3)$$

Si σ est un k-cycle et que a n'est pas un point fixe, alors on en déduit[1] que le support de σ est donné par:

$$\mathrm{Supp}(\sigma) = \{a, \sigma(a), \sigma^2(a), \dots, \sigma^{k-1}(a)\}$$

Ordre :

On peut alors démontrer une propriété fondamentale de l'ordre des cycles:

Un k-cycle est d'ordre k.

En effet si on considère le sous-groupe engendré par un tel cycle, on remarque que pour tout élément $a \in [\![1\ ;\ n]\!]$ $\sigma^k(a) = a$, donc $\sigma^k = \mathrm{Id}$.

[1] En effet par exemple $\sigma(a)$ est la valeur suivante dans le cycle, et le cycle parcourt tout les points non-fixes par construction

Théorèmes de décomposition :

Une des problématiques principales à propos des groupes symétriques est la question de la **décomposition d'une permutation** en cycles. On peut en effet montrer que **toute permutation se décompose en produit de cycles à support disjoints**.

Pour ceci, on utilise le fait que toute permutation induit une **partition en orbites** de $[\![1\ ;\ n]\!]$, ces orbites correspondront alors au supports des cycles dans la décomposition.

Par la suite, on peut alors constater directement que pour tout k-cycle $\sigma = (a_1\ \ \dots\ \ a_k)$, on a:

$$\sigma = (a_1\ \ a_2)(a_2\ \ a_3)\dots(a_{k-1}\ \ a_k)$$

Enfin, on conclura de ces deux propositions que **toute permutation se décompose en produit de transpositions**, ou en d'autres termes si on note $\mathfrak{T}_n$ l'ensemble des transpositions:

$$\langle \mathfrak{T}_n \rangle = \mathfrak{S}_n$$

Conjugaison et permuations :

On considère alors l'action de $\mathfrak{S}_n$ sur lui-même par conjugaison, on peut alors montrer que pour toute permutation σ, on a:

$$\sigma(a_1, \dots, a_n)\sigma^{-1} = (\sigma(a_1), \dots, \sigma(a_n))$$

En particulier, on a alors que deux cycles sont conjugués si et seulement si ils ont la même longueur, et si on définit le **type d'une permutation** par le n-uplet **non ordonné** $[l_1, \dots, l_k]$ des longueurs des cycles dans sa décomposition en cycles, on a alors une caractérisation des classes de conjugaisons:

Deux permutations sont conjuguées si et seulement si elles ont même type.

Signature :

A REFAIRE.

II — Groupes Cycliques

On appelle **groupe cyclique** un groupe G engendré par un unique élément qu'on notera g. Le but de ce chapitre est de classifier ces groupes et d'identifier leurs caractéristiques.

On considère tout d'abord le groupe quotient $\mathbb{Z}/n\mathbb{Z}$, on peut alors remarque que les éléments de ce groupe sont exactement **les classes de restes possibles par la division euclidienne par** n. Par exemple le groupe $\mathbb{Z}/6\mathbb{Z}$ représente le fait que $2 \equiv 8 \equiv 14[6]$.

Classification :

Dans cette section on retourne dans le case d'un groupe cyclique général G et on définit le morphisme surjectif suivant:

$$\begin{aligned} \phi : \mathbb{Z} &\longrightarrow G \\ n &\longmapsto g^n \end{aligned}$$

On raisonne sur la finitude de G et on peut alors caractériser tout les groupes cycliques trés simplement, en effet:

- Si G est infini, le morphisme ϕ est **injectif** et on a l'isomorphisme $G \cong \mathbb{Z}$
- Si G est fini, on utilise le **premier théorème d'isomorphisme** et on a l'isomorphisme $G \cong \mathbb{Z}/n\mathbb{Z}$

Il n'y a donc qu'un seul groupe cyclique d'ordre n (resp. d'ordre infini), celui des classes de congruences modulo n (resp. celui des entiers).

On remarque alors l'importance du groupe quotient $\mathbb{Z}/n\mathbb{Z}$, c'est le prototype de groupe cyclique fini.

Générateurs de $\mathbb{Z}/n\mathbb{Z}$:

On sait donc que ce groupe est **cyclique** d'ordre n, en particulier il est engendré par 1, mais aussi par toutes les classes dont le représentant est premier avec n, en effet si a est un tel élément alors d'aprés le théorème de Bézout, on a:

$$\exists u, v \in \mathbb{Z} \; ; \; au + bn = 1$$

Donc en particulier $[a] + \ldots + [a] = 1$ dans $\mathbb{Z}/n\mathbb{Z}$ et par la suite, a engendre tout le groupe. Il y a donc $\varphi(n)$ générateurs de ce groupe.

Théorème chinois :

Un des grands théorèmes sur les groupes cycliques est le suivant, si on considère $p_1, \ldots, p_k \in \mathbb{N}$ des nombres premiers entre eux, et qu'on note n leur produit, alors on peut montrer facilement qu'on a l'isomorphisme suivant:

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/p_1\mathbb{Z} \times \ldots \times \mathbb{Z}/p_k\mathbb{Z}$$

En particulier si $n = p_1^{\alpha_1} \ldots p_k^{\alpha_k}$ décompé en facteurs premiers, alors on a:

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \ldots \times \mathbb{Z}/p_k^{\alpha_k}\mathbb{Z}$$

C'est la **décomposition primaire d'un groupe cyclique**. Elle s'interprète en comprenant par exemple que la donnée du reste par 6 d'un entier est exactement equivalente à la donnée de son reste par 3 et 2.

II — Anneaux

Soit A un ensemble **non-vide** muni de deux lois de composition internes associatives notées $+, \times$ telles que:

- $(A, +)$ soit un groupe commutatif.
- La loi $\times$ est associative.
- La loi $\times$ est distributive sur la loi $+$.
- Il existe **un élément neutre** pour la loi $\times$.

Alors le triplet $(A, +, \times)$ est appellé **anneau**. Si la loi multiplicative est **commutative**, on dira alors que c'est un anneau commutatif. On définit aussi de nouveau types d'élément remarquable spécifiques au cas des anneaux:

- On dit qu'un élément $x \in A$ est **un diviseur de zéro**[1] si il existe y tel que $xy = 0$.
- On dit qu'un élément $x \in A$ est **un nilpotent** si il existe $n \in \mathbb{N}$ tel que $x^n = 0$.

On définit aussi les **inversibles** à droite ou à gauche pour la seconde loi. On dira qu'un anneau sans diviseurs de zéro est **intègre**, et dans ce cas on a la propriété suivante trés puissante:

$$\forall x, y \in A \ ; \ xy = 0 \implies x = 0 \text{ ou } y = 0$$

Exemples :

On peut alors considérer plusieurs anneaux remarquables:

- Les **entiers relatifs** muni des opérations usuelles forment un anneau intègre.
- Les **polynomes** muni de la somme et du produit forment un anneau intègre.
- Les **fonctions continues** muni de la somme et du produit forment un anneau.
- Les **matrices** muni de la somme et du produit forment un anneau.

Propriétés Algébriques:

Pour deux éléments $a, b \in A$ qui commutent, on a **la formule du binome de Newton**:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Sous-anneaux :

Les sous-structures dans le cas des groupes sont naturellement les **sous-anneaux**. Un cas remarquable est celui du **sous-anneau engendré** par H:

$$\langle H \rangle := \left\{ \sum_{k=1}^{n} \pm h_1^{k_1} h_2^{k_2} \dots h_n^{k_n} \ ; \ n \in \mathbb{N} \ , \ h_i \in H \ , \ k_i \in \mathbb{N} \right\}$$

On peut alors imagine généraliser la notion de sous-groupe normal, ie une sous-structure qui permet de quotient, mais il se trouve qu'alors la notion de sous-anneau engendré n'est pas la bonne notion, et on définit plutôt la notion **d'idéal** qui est un sous groupe additif de A qui soit stable par multiplication (à droite et à gauche) par n'importe quel élément de l'anneau.

[1] Ici c'est un diviseur de zéro **à droite**, on définit de même les diviseurs de zéro **à gauche**.

Caractéristique :

On définit la caractéristique d'un anneau non-nul par:

$$\mathrm{car}(A) := \min \left\{ n \in \mathbb{N} \, ; \, \underbrace{1 + \ldots + 1}_{n \text{ sommandes}} = 0 \right\}$$

Une autre formulation serait simplement que:

La caractéristique d'un anneau est l'ordre (additif) de l'unité multiplicative.

Anneaux à PGCD :

Anneaux Factoriels :

Anneaux Principaux :

Anneaux Euclidiens :

On appelle **anneau Euclidiens** tout anneau A principal qui possède une **division euclidienne**. Dans un tel anneau, on peut alors faire **de l'arithmétique** comme dans l'anneau des entiers naturels.

Schéma heuristique des structures d'anneaux :

Pour mieux visualiser la hierarchie des différents types d'anneaux, on peut représenter la structure logique sous la forme de la suite d'implications suivantes:

Euclidien $\Longrightarrow$ **Principal** $\Longrightarrow$ **Factoriel** $\Longrightarrow$ **PGCD** $\Longrightarrow$ **Intégre** $\Longrightarrow$ **Commutatif**

II — Corps

Soit A un anneau dont tout les éléments sauf 0 sont inversibles. Alors on dit que A est **un corps**.

Exemples :

On peut alors considérer plusieurs corps remarquables:

- Les **réels** muni des opérations usuelles.
- Les **complexes** muni des opérations usuelles.
- Les **quaternions**[1] muni des opérations usuelles.
- Les **corps finis** $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ pour p premier.
- Les **nombres constructibles** à la règle et au compas.

[1] C'est un exemple de corps non commutatif

II — Corps des Complexes

On définit le nombre imaginaire i dont le carré vaut -1, et on construit alors $\mathbb{C}$ comme l'extension du corps[1] $\mathbb{R}$ avec les deux lois usuelles, ie on définit:

$$\mathbb{C} := \mathbb{R}[i] = \left\{ a + ib \ ; \ a, b \in \mathbb{R} \right\}$$

On peut alors montrer que c'est un ensemble stable pour les lois usuelles et qu'il vérifie toutes les propriétés qui font de lui un **corps**.
Chaque nombre complexe se définit alors comme des sommes ou produits de réels et du nombre imaginaire et on appelle alors cette expression la **forme algébrique** d'un nombre complexe et on appelle a **la partie réelle** et b **la partie imaginaire** de ce nombre.

Géométriquement, on peut identifier les nombres complexes à des points du plan, en effet, $a + ib$ peut se comprendre comme une combinaison linéaire d'un nombre de l'axe réel, et d'un nombre de l'axe imaginaire.

Module :

On appelle **module** de $z \in \mathbb{C}$ le **prolongement** de la fonction valeur absolue à $\mathbb{C}$, c'est donc une **norme** et on la définit telle que :

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}$$

Dans la suite, on notera ϱ le module de z pour faciliter la lecture.

Forme trigonométrique :

L'interprétation géométrique permet alors de montrer par passage en coordonées polaires qu'il existe un unique angle θ (modulo 2π) qu'on appelle **argument** de z tel que:

$$z = \varrho(\cos\theta + i\sin\theta)$$

Forme exponentielle :

De même on définit alors **la forme exponentielle** de z l'expression:

$$z = \varrho e^{i\theta} := \varrho(\cos\theta + i\sin\theta)$$

On peut alors étendre les propriétés usuelles de l'exponentielle à $\mathbb{C}$ et on en déduit:

$$\begin{aligned} \arg(zz') &\underset{2\pi}{\equiv} \arg(z) + \arg(z') \\ \arg(\frac{z}{z'}) &\underset{2\pi}{\equiv} \arg(z) - \arg(z') \end{aligned}$$

Conjugué :

On appelle conjugaison **l'involution** qui à z associe son **conjugué**, noté $\overline{z}$ tel que:

$$\overline{z} := a - bi = \varrho(\cos\theta - i\sin\theta) = \varrho e^{-i\theta}$$

C'est une application **additive** et **multiplicative**, on montre alors les formules suivantes :

$$\Re(z) := \frac{z + \overline{z}}{2} \qquad\qquad \mathrm{Im}(z) := \frac{z - \overline{z}}{2i}$$

En utilisant ces formules pour z sous forme exponentielle, on a alors les **formules d'Euler** qui sont très importantes car elle permettent de **linéariser** des expression trigonométriques.

[1] La motivation principale de l'introduction de i et de cette construction est que $\mathbb{C}$ est algébriquement clos, ie tout les polynomes de degré n de $\mathbb{C}[X]$ ont n racines.

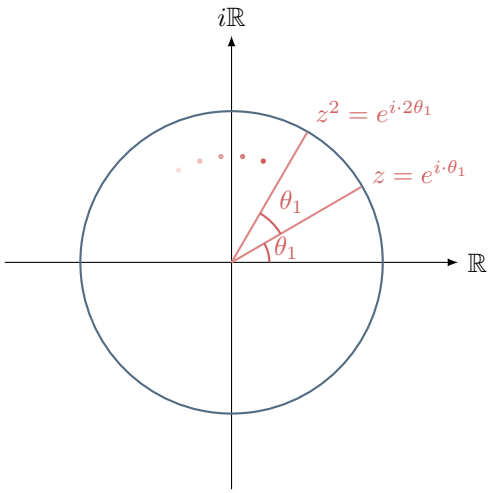
Formule de Moivre :

Un propriété importante des formes trigonométriques et exponentielles apellée **formule de Moivre**[1] est:

$$(e^{i\theta})^n = e^{n(i\theta)}$$
$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

Les différentes puissances d'un nombre complexe (de module 1) s'interprétent alors comme des points situés à equidistance sur un cercle.

Graphiquement:



Racines n-ièmes :

Soit $n \in \mathbb{N}$, une partie importante des problèmes impliquant des nombres complexes proviennent d'équations d'inconnue Z de la forme:

$$Z^n = z$$

On peut montrer que l'ensemble des solutions de ce type de problème est:

$$S = \left\{ \sqrt[n]{\varrho} e^{i\frac{\theta+2k\pi}{n}} \ ; \ k \in \{0, 1, \ldots, n-1\} \right\}$$

Cas particulier : Si on a une racine n-ième Z_0 de Z et qu'on connaît les racines n-ièmes de l'unité, alors on peut obtenir toutes les racines n-ièmes de Z grâce à:

$$\left\{ Z \in \mathbb{C} \ ; \ Z^n = z \right\} = \left\{ Z_0 u \ ; \ u \in \mathbb{U}_n \right\}$$

Le nombre complexe j :

On note j la première racine troisième de l'unité. Le nombre j est singulier, car il vérifie:

$$j^2 = j^{-1} = \overline{j}$$

Graphiquement, on peut observer que les affixes des nombres 1, j et $\overline{j}$ forment un triangle équilatéral inscrit dans le cercle trigonométrique.

[1] Ici, on a choisi de considérer $z \in \mathbb{U}$ mais ces propriétés sont vraies pour **tout nombre complexe**, il suffit alors d'appliquer la puissance au module.

II — Anneau des Polynômes

Définition :

Soit K un corps, et $n \in \mathbb{N}$, on appelle **polynômes** à coefficients dans K en l'indéterminée X les éléments de l'ensemble:

$$K[X] := \left\{ \sum_{i=0}^{n} a_i X^i \ ; \ a_i \in K \right\}$$

Degré et Valuation :

Soit $P, Q, R \in K[X]$.
On définit tout d'abord une propriété fondamentale appelée **degré** de P telle que $\deg(P)$ est le plus grand coefficient non nul de P. On a alors les propriétés du degré ci-dessous:

$$\deg(P+Q) \leq \max(\deg(P), \deg(Q))$$
$$\deg(PQ) = \deg(P) + \deg(Q)$$

La valuation est définie de manière analogue comme le plus petit coefficient non nul de P.

Opérations :

On considère ci-dessous que $\deg(P) = n$ et $\deg(Q) = m$ et on note a_i, b_i les coefficients de P (resp. Q).

On adjoint à cet ensemble une loi d'addition qui est simplement effectuée terme à termes.
On adjoint à cet ensemble une loi de multiplication se définit explicitement comme suit:

$$PQ := \sum_{k=0}^{n+m} \sum_{i+j=k} a_i b_i X^k$$

Muni de ces deux opérations, on donne à cet ensemble une structure **d'anneau**[1].

On ajoute aussi une opération appelée **dérivation formelle** d'un polynôme qui s'effectue comme la dérivation analytique usuelle.

Divisibilité :

On définit tout d'abord une **division euclidienne** de deux polynômes qui se comporte comme la division euclidienne usuelle, à la différence que la condition d'arrêt porte sur le **degré du reste** qui doit être inférieur à celui du diviseur.

Soit $U, V \in R[X]^2$, on peut aussi définir une relation de **divisibilité** entre deux polynômes, cette relation est **transitive et réflexive** et on a aussi:

$$D|A \wedge D|B \implies D|UA + VB$$

Racines et Factorisation :

On dit que α est une **racine** de P si $P(\alpha) = 0$.
On montre alors le théorème fondamental ci-dessous:

$$P(\alpha) = 0 \iff (X - \alpha)|P$$

[1] Intègre car les coefficients viennent d'un corps.

On appelle **multiplicité d'une racine** α l'entier m tel que:

$$\Big[(X-\alpha)^m | P\Big] \wedge \Big[(X-\alpha)^{m+1} \nmid P\Big]$$

Si on note P^m la dérivée n-ième de P, on a aussi:

$$P^m(\alpha) = 0 \ \wedge \ P^{m+1}(\alpha) \neq 0$$

On en déduit que pour une racine α de multiplicité m, on peut **factoriser** P par $(X-\alpha)^m$.

Si on considère maintenant plusieurs racines **distinctes** $a_0, a_1, \ldots, a_{n-1}, a_n$ de multiplicité respectivement $m_0, m_1, \ldots, m_{n-1}, m_n$, on peut montrer la propriété suivante:

$$\Big[\prod_{i=0}^{n}(X-\alpha_i)^{m_i}\Big] \Big| P \qquad \text{(On peut factoriser par le produit des } (X-\alpha_i)^{m_i}\text{)}$$

Décomposition :

On appelle décomposition d'un polynôme P une factorisation **irréductible** de P, cette décomposition dépend du corps considéré, en effet si on considère $\mathbb{C}[X]$, on peut montrer le **théorème fondamental de l'Algèbre** ci-dessous:

> Tout polynôme de degré n admet n racines dans $\mathbb{C}$.

Par suite, on peut montrer que tout les polynômes de $\mathbb{C}[X]$ sont **scindés**[1].
Par contre dans $\mathbb{R}[X]$, il existe évidemment des polynômes de degré 2 irréductibles.

Soit $z \in \mathbb{C}$, il existe une propriété très utile pour décomposer un polynôme **à coefficients réel** qui est:

$$P(z) = 0 \implies P(\overline{z}) = 0$$

Fonctions symétriques des racines :

On définit le k-ième **polynôme symétrique** à n indéterminées comme la somme des produits à k facteurs de ses indéterminées, et on le note σ_k.

Par exemple pour un polynôme en trois indéterminées a, b, c, on a successivement:

$\sigma_1 = a + b + c$ (Somme des produits à 1 facteur)
$\sigma_2 = ab + ac + bc$ (Somme des produits à 2 facteurs)
$\sigma_3 = abc$ (Somme des produits à 3 facteurs)

De manière générale on a:

$$\sigma_k(x_1, \ldots, x_n) = \sum_{1 \leq i_1 < \ldots < i_k \leq n} x_{i_1} x_{i_2} \ldots x_{i_k}$$

Soit $P \in K[X]$ un polynôme de degré n, on note $x_1, x_2, \ldots, x_{n-1}, x_n$ ses n racines et $a_1, a_2, \ldots, a_{n-1}, a_n$ ses n coefficients.
Alors, pour tout $k \in [\![0\ ;\ n]\!]$, on a le **théorème**:

$$\sigma_k(x_1, x_2, \ldots, x_{n-1}, x_n) = (-1)^k \frac{a_{n-k}}{a_n}$$

Ce théorème permet d'obtenir une relation **coefficient-racines**.

Exemple:
- Pour $k = 1$, on trouve que la somme des racines de P vaut $-\frac{a_{n-1}}{a_n}$
- Pour $k = 2$, on trouve que la somme des doubles produits des racines de P vaut $\frac{a_{n-2}}{a_n}$
- .....

[1] C'est à dire que tout les polynômes irréductibles de $\mathbb{C}[X]$ sont de degré 1.

III — Espaces Vectoriels

L'algèbre linéaire est une partie de l'algèbre générale s'intéressant à une structure particulière omniprésente en mathématiques, la structure **d'espace vectoriel**, c'est une structure tout comme les groupes et les anneaux et elle formalise la plupart des notions géométriques usuelles dans les espaces commes $\mathbb{R}^2$ ou $\mathbb{R}^3$. Tout les résultats d'algèbre générale s'appliquent bien évidemment à cette structure.

Définition :

Soit E un ensemble non-vide et $\mathbb{K}$ un corps commutatif. On dira alors que $(E, +, \cdot)$ est un **espace vectoriel sur** $\mathbb{K}$ si les conditions ci-dessous sont réunies:

- $(E, +)$ est un **groupe abélien**.
- La loi externe $\cdot$ est une **action de groupe** sur E qui vérifie la **distributivité mixte**.

On appelle alors vecteurs les éléments de E et scalaires les éléments de $\mathbb{K}$.

Sous-espaces vectoriels :

Soit $F \subseteq E$, on dit que F est un **sous-espace vectoriel** et on note $F \leq E$ si et seulement si:

- F est non-vide.
- F est stable par somme.
- F est stable par multiplication externe.

On montre alors facilement que **l'intersection** de deux sous-espaces vectoriels est aussi un sous-espace vectoriel mais que l'union de deux sous-espaces vectoriels n'est en général pas un sous-espace vectoriel.

Sous-espace engendré :

On se donne une partie F de E, on peut alors caractériser le **sous-espace vectoriel engendré** comme défini dans le chapitre d'algèbre par:

$$\text{Vect}(F) := \left\{ \sum_{i=0}^{n} \lambda_i u_i \ ; \ (\lambda_i, u_i) \in \mathbb{K} \times F \ , \ n \in \mathbb{N} \right\}$$

*On dit que de telles combinaisons sont **des combinaisons linéaires** de vecteurs de F. Le sous-espace engendré est donc l'ensemble des combinaisons linéaires finies de vecteurs de F.*

Familles libre et génératrices :

Soit $\mathscr{F} := (u_1, \ldots, u_n)$ une famille de vecteurs de E.

On dit que $\mathscr{F}$ est **génératrice** de E si on a $E = \text{Vect}(\mathscr{F})$.
On dit que $\mathscr{F}$ est **libre** si tout combinaison linéaire nulle de vecteurs de $\mathscr{F}$ est à coefficient tous nuls.

Cette proposition signifie exactement que l'on ne peut pas obtenir un vecteur comme combinaison linéaire d'autres vecteurs, en effet si un des coefficients était non nul, il suffirait d'isoler le vecteur correspondant et il serait alors redondant. Formellement, on a:

$$\forall (\lambda_1, \ldots, \lambda_n) \in K^n \ ; \ \left[\sum_{i=0}^{n} \lambda_i u_i = 0_E \implies (\lambda_1, \ldots, \lambda_n) = (0, \ldots, 0) \right]$$

Bases :

On appelle **base** de E une famille **libre et génératrice**. Ce concept permet alors de caractériser le fait que tout élément de E peut s'écrire comme une **unique** combinaison linéaire des vecteurs de la base.

Soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base de E. On appelle **coordonnées** de u dans la base $\mathscr{B}$ les coefficients de la décomposition de u dans la base $\mathscr{B}$ et on note:

$$[u]^{\mathscr{B}} = [\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_{n-1} e_{n-1} + \lambda_{n-1} e_{n-1}]^{\mathscr{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Sommes et sommes directes:

Soit n un entier naturel et $(F_k)_{k \leq n}$ une famille finie de sous-espaces vectoriels de E, Alors on peut construire le plus petit sous espace vectoriel qui contient tout les (F_k) par:

$$S := \sum_{k \leq n} F_k := \left\{ \sum_{k \leq n} u_k \ ; \ u_k \in F_k \right\}$$

On dira alors que cette somme est **directe** si et seulement si la décomposition de zéro dans la somme est **unique**[1], ie:

$$\sum_{k \leq n} u_k = 0_E \ ; \ u_k \in F_k \implies (u_1, u_2, \ldots, u_n) = (0, 0, \ldots, 0)$$

Et on la note alors:

$$S = \bigoplus_{k \leq n} F_k$$

Si de plus la somme est égale à **l'espace entier**, alors on dira que les F_k sont **supplémentaires** dans E.

Caractérisation par les bases:

Soit $\mathscr{B}_k$ des bases de chacun des F_k, soit la famille $\mathscr{F} = (\mathscr{B}_1, \mathscr{B}_2, \ldots, \mathscr{B}_n)$, ie la famille constituée de bases des F_k concaténées. Alors on a alors le théorème suivant:

$$\mathscr{F} \text{ est une base de } S \iff \sum_{k \leq n} F_k = \bigoplus_{k \leq n} F_n$$

Espaces vectoriels quotient:

Les espaces vectoriels étant des groupes commutatifs par définition, tout ses sous groupes sont normaux, et la compatibilité par la loi externe est directe, on peut donc définir pour tout $F \leq E$, **l'espace vectoriel quotient** E/F.

- Exemple 1: Si $E = \mathbb{R}^2$ et $F = \text{Vect}(1, 0)$, alors E/F est **l'ensemble des droites parallèles à l'axe des abcisses**.
- Exemple 2: Si $E = \mathbb{R}_3[X]$ et $F = \text{Vect}(X^2)$, alors E/F est **l'ensemble des polynômes qui ne différent que d'un terme quadratique**.

[1] En particulier si $n = 2$, on montre directement qu'une condition nécessaire et suffisante pour que la somme soit directe est que:

$$F \cap G = \{0_E\}$$

III — Espaces Affines

Soit $\mathscr{E}$ un ensemble non-vide et V un $\mathbb{R}$-espace vectoriel de dimension n finie.

Définition :

On appelle **espace affine**[1] **de direction** V le couple $(\mathscr{E}, +)$ avec:

$$\begin{aligned} + : \quad \mathscr{E} \times V &\longrightarrow \mathscr{E} \\ (A,\, u) &\longmapsto A + u \end{aligned}$$

La loi $+$ doit vérifer les axiomes suivants [2]:

Existence d'un neutre	$\forall A \in \mathscr{E}\ ;\ A + 0_V = A$
Action de groupe	$\forall A \in \mathscr{E}\ ,\ \forall u, v \in V\ ;\ A + (u + v) = (A + u) + v$
Unicité du translaté	$\forall A, B \in \mathscr{E}\ ,\ \exists! u \in V\ ;\ A + u = B$

Etant donné deux points $A, B \in \mathscr{E}$, on note alors $\overrightarrow{AB}$ l'unique vecteur u qui vérifie:

$$A + u = B$$

On appelle alors A le **point initial** et B le **point final**

Propriétés :

Si $\mathscr{E}$ est un espace affine de direction V, on a alors les propriétés suivantes:

Relation de Chasles	$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$
Existence d'un neutre	$\overrightarrow{AB} + \overrightarrow{BA} = O_V$

Repères :

Un repère de $\mathscr{E}$ est un couple $\mathscr{R} = (O, \mathscr{B})$ formé d'un point $O \in \mathscr{E}$ et d'une base $\mathscr{B} = (e_1, \ldots, e_n)$ de V. On appelle alors le point O **origine** du repère, et les vecteurs $e_1, \ldots, e_n$ **vecteurs de base** du repère.

Soit $i \in [\![1\ ;\ n]\!]$, alors pour tout point $A \in \mathscr{E}$, on appelle **coordonées** de A dans le repère $\mathscr{R}$, les composantes $(x_i)_i$ du vecteur qui représente la translation de O vers A dans la base $\mathscr{B}$, et on a la caractérisation élémentaire suivante:

$$\overrightarrow{OA} = x_1 e_1 + \ldots + x_n e_n$$

Cas particulier des espaces vectoriels :

Il est important de noter que le couple $(V, +)$ forme un espace affine de **direction lui-même**. En effet, si on considère un élément de V comme un point, alors le couple $(V, +)$ forme un espace affine de direction V et la loi $+$ est alors exactement la loi de composition interne de V.

L'unique vecteur u tel que $A + u = B$ est exactement $B - A$ (ici A, B sont des points de V qui se trouvent être des vecteurs dans ce cas particulier).

---

[1] Ses éléments sont alors appelés des **points**.

[2] Une action d'un groupe $(G, \star)$ sur E est une application $+$ de $G \times E$ dans E qui vérifie:

$$\forall g, g' \in G\ ,\ \forall x \in E\ ;\ x + (g \star g') = (x + g) \star g'$$

En d'autres termes, additionner d'abord les vecteurs, ou d'abord le point avec le vecteur n'importe pas.

III — Théorie de la dimension

Dans ce chapitre, on considérera un espace vectoriel E qui admet une famille génératrice **finie**. On dira alors que E est de **dimension finie**.

Théorème de la base incomplète:

Soit $\mathcal{L}$ une famille libre et $\mathcal{G}$ une famille génératrice de E, le concept de dimension se définit grâce au **théorème de la base incomplète**:

- On peut **compléter** $\mathcal{L}$ en une base de E par ajouts de vecteurs de $\mathcal{G}$.
- On peut **extraire** de $\mathcal{G}$ une base de E.

Ce théorème permet alors d'assurer **l'existence** d'une base d'une espace vectoriel de dimension finie. Il se démontre par exhibition d'un algorithme qui complête $\mathcal{L}$ en une base.

Théorème de la dimension:

On peut alors montrer que le cardinal d'une partie libre est toujours inférieur au cardinal d'une partie génératrice[1], de cette considération, on peut alors montrer directement **le théorème de la dimension** qui énonce que toutes les bases d'un espace vectoriel de dimensions fini ont **même cardinal**.

Ce théorème permet alors d'assurer **l'unicité** du cardinal des bases d'une espace vectoriel de dimension finie.

Définition de la dimension:

Des deux théorèmes précédents, on a alors l'existence de bases d'une espace de dimension fini, et l'unicité de leur cardinal, on peut alors définir **la dimension d'une espace vectoriel** E comme ce cardinal et on la note $\dim(E)$.

Espaces de dimension finie:

Considérons maintenant E un espace vectoriel de dimension finie n est $\mathscr{F}$ une famille de n vecteurs de E. Alors par déduction immédiate de la définition de dimension, on a:

$$\mathscr{F} \text{ est libre} \iff \mathscr{F} \text{ est génératrice} \iff \mathscr{F} \text{ est une base.}$$

Soient F, G deux sous-espaces de E. La dimension permet aussi de prouver des **égalités** d'espaces vectoriels, grâce aux propriétés suivantes:

- Si $F \subseteq G$ et $\dim(F) = \dim(G)$, alors $F = G$.
- Si F, G sont en **somme directe** et que $\dim(F) + \dim(G) = \dim(E)$, alors ils sont **supplémentaires**.

Enfin on peut calculer la dimension d'une somme avec la **formule de Grassmann**:

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

Rang d'une famille de vecteurs:

Soit E un espace vectoriel de dimension finie et $\mathscr{F}$ une famille de vecteurs de cet espace, alors on appelle **rang** de $\mathscr{F}$ l'entier:

$$\mathrm{rg}(\mathscr{F}) = \dim(\mathrm{Vect}(\mathscr{F}))$$

C'est simplement **la dimension du sous-espace engendré par la famille**.

[1] Aussi appellé **lemme de Steiniz**.

III — Applications Linéaires

Soit deux $\mathbb{K}$-espaces vectoriels E et F, et $f : E \longrightarrow F$.

Définition:

On dit que f est une **application linéaire** si c'est un **morphisme d'espaces vectoriels**, ie si et seulement si pour tout couple de vecteurs $u, v \in E$ et pour tout scalaire λ elle vérifie:

- **Loi interne :** $f(u+v) = f(u) + f(v)$
- **Loi externe :** $f(\lambda u) = \lambda f(u)$

On note alors $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E vers F. Si $F = \mathbb{K}$, on dira que f est une **forme linéaire**.

Propriétés:

On s'intéresse au propriétés de l'ensemble $\mathcal{L}(E, F)$, c'est un ensemble de morphismes donc d'aprés le chapitre d'algèbre la composée de morphismes et l'inverse d'un morphisme bijectif est un morphisme.

En outre en considèrant les espaces vectoriels comme groupes additifs, on vérifie que le noyau d'un morphisme est un sous-espace bien défini et caractérise l'injectivité de ce dernier. De même, l'image de générateurs engendre l'image qui est aussi un sous-espace.

Caractérisations par les familles:

Soit $\mathscr{F} = (e_i)_{i \in \mathbb{N}}$ une famille de E et un endomorphisme de E, alors f est entièrement caractérisée par l'image de cette famille, en effet on a:

- L'image d'une famille libre est libre $\iff f$ est injective.
- L'image d'une famille génératice est génératice $\iff f$ est surjective.

Endomorphismes élémentaires remarquables:

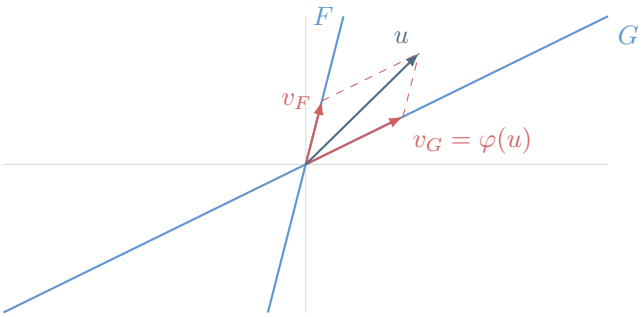
On définit ici des endomorphismes élémentaires remarquables comme:

- **Les homotéthies:** Elles sont caractérisées par $\varphi_k : u \longmapsto ku$
- **Les symétries:**Elles sont caractérisées par $\varphi \circ \varphi = \mathrm{Id}_E$
- **Les projecteurs:** Elles sont caractérisées par $\varphi \circ \varphi = \varphi$

Précisons le cas des projecteurs, en effet si on considère deux sous-espaces F, G supplémentaires dans E, alors chaque élément $u \in E$ admet une décomposition unique de la forme $u = v_F + v_G$. Cette décomposition définit canoniquement deux projecteurs, par exemple celui de direction F sur G qui est l'application φ telle que:

$$\varphi(u) = v_G$$

Graphiquement, pour φ le projecteur de direction F sur G:



Applications linéaires en dimension finie:

On étends la définition de rang d'une famille à celle du **rang d'une application linéaire**, qu'on note $\mathrm{rg}(f)$, qui correspond à **la dimension de son image**. Soit $f \in \mathcal{L}(E, F)$ une application de rang fini, alors d'après le premier théorème d'isomorphisme:

$$E/\ker(f) \cong \mathrm{Im}(f)$$

Et donc on a égalité des dimensions et aprés avoir montré que $\dim(E/F) = \dim(E) - \dim(F)$, on en déduis le **théorème du rang**:

$$\dim(E) = \mathrm{rg}(f) + \dim(\mathrm{Ker}(f))$$

Ce thèorème permet alors de caractériser l'injectivité et la surjectivité d'une application linéaire par:

- f est injective si et seulement si $\mathrm{rg}(f) = \dim(E)$
- f est surjective si et seulement si $\mathrm{rg}(f) = \dim(F)$

En particulier, si E et F sont deux espaces vectoriels **de même dimension** alors:

$$f \text{ est injective} \iff f \text{ est surjective} \iff f \text{ est bijective}$$

Ceci caractérise alors **les isomorphismes en dimension finie**.

Applications multilinéaires:

On peut généraliser le concept de linéarité à celui de **multilinéarité** ou **n-linéarité**. On considère alors une application de la forme $f : E^n \longrightarrow F$, alors on dit que f est multilinéaire si et seulement si elle est linéaire **en chacune des variables**, ie si:

$$\begin{aligned}
f(e_1 + \lambda u, e_2, \ldots, e_n) &= f(e_1, e_2, \ldots, e_n) + \lambda f(u, e_2, \ldots, e_n) \\
f(e_1, e_2 + \lambda u, \ldots, e_n) &= f(e_1, e_2, \ldots, e_n) + \lambda f(e_1, u, \ldots, e_n) \\
&\vdots \\
f(e_1, e_2, \ldots, e_n + \lambda u) &= f(e_1, e_2, \ldots, e_n) + \lambda f(e_1, e_2, \ldots, u)
\end{aligned}$$

Une telle application est dite:

- **Symétrique** si permuter deux variables préserve le résultat.[1]
- **Antisymétrique** si permuter deux variables change le signe du résultat.[2]
- **Altérnée** si elle s'annule à chaque fois qu'on l'évalue sur un k-uplet contenant deux vecteurs identiques.[3]

[1] Exemple: $f(e_1, e_2) = f(e_2, e_1)$

[2] Exemple: $f(e_1, e_2) = -f(e_2, e_1)$, en particulier, le signe du résultat aprés une permutation σ dépends alors de **la signature de la permutation** (la parité du nombre de permutations effectuées).

[3] Exemple: $f(e_1, e_1) = 0$

III — Espace des matrices

On appelle **matrice** à n lignes et p colonnes à coefficients dans un anneau $\mathbb{A}$ toute application de la forme:

$$M : [\![1\ ;\ n]\!] \times [\![1\ ;\ p]\!] \longrightarrow \mathbb{A}$$

Il s'agit d'une généralisation du concept de suite sous forme de suite à **deux indices**, qu'on peut alors voir comme un tableau de nombres tel qu'en chaque position (i, j), on ait un élément $a_{ij} \in \mathbb{A}$. A l'instar des suites, on notera $M = (a_{ij})$ pour faire référence à la matrice M.

On note alors $\mathcal{M}_{n,p}(\mathbb{A})$ l'espaces des matrices à n lignes et p colonnes à coefficients dans $\mathbb{A}$.

Structure:

On peut alors munir cete espace d'une structure **d'espace vectoriel** sur $\mathbb{K}$ par les opérations suivantes:

- On définit **la somme** de deux matrices par la matrice obtenue en sommant par composantes.
- On définit **la multiplication** d'une matrice par un scalaire comme la matrice dont tout les termes sont multipliés par ce scalaire.

Le vecteur nul de cet espace est alors la matrice nulle composée uniquement de zéros.

Produit matriciel:

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,m}(\mathbb{K})$ et $B = (b_{k,j}) \in \mathcal{M}_{m,p}(\mathbb{K})$. Alors on définit[1] la matrice $C := AB = (c_{i,j})$ comme étant la matrice telle que:

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

En particulier, on appelle ce produit **un produit ligne par colonne** qui se comprends visuellement par:

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{pmatrix} \times \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \end{pmatrix} = \begin{pmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{pmatrix}$$

Le coefficient à la troisième ligne, première colonne est obtenu en multipliant la troisième ligne par la première colonne.

Cas des matrices carrées à coefficients dans un corps:

Si $n = p$, et que $\mathbb{A} = \mathbb{K}$ est un corps, alors toutes les matrices en jeu sont carrées et cette loi est **interne**. On peut alors montrer que $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un **anneau unitaire non-commutatif et non-intègre**.

En particulier, le neutre pour cette loi est la matrice identité nulle partout et dont les termes diagonaux sont tous égaux à 1.

[1] Il faut que le nombre de colonnes de la première soit égal au nombre de lignes de la seconde pour que les matrices soient dites **compatibles**.

Transposition:

Soit $M = (x_{i,j}) \in \mathcal{M}_{n,m}(\mathbb{K})$, on définit **l'opération de transposition** d'une matrice notée M^{T}, c'est **une application linéaire involutive** définie par:

$$M^{\mathsf{T}} = (x_{j,i})$$

Intuitivement, cette application transforme **chaque ligne en colonne et inversement**. Par exemple:

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}^{\top} = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}$$

Il faut aussi noter son comportement par rapport au produit matriciel de deux matrices A, B, précisément on a:

$$(AB)^{\top} = B^{\top} A^{\top}$$

Trace:

On définit aussi une autre application linéaire appellée **trace d'une matrice**, et qui est définie comme **la somme des éléments diagonaux**. Formellement:

$$\mathrm{tr}(A) := \sum_{i=1}^{n} a_{i,i}$$

Elle est donc linéaire mais on a aussi:

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$

Matrice d'une famille de vecteurs:

Soit $\mathscr{F} = (e_i)_{i \in \mathbb{N}}$, alors on peut définir **la matrice de la famille** dans une base $\mathscr{B}$ par:

$$\boxed{\mathrm{Mat}_{\mathscr{B}}(\mathscr{F}) = ([e_1]_{\mathscr{B}}, [e_2]_{\mathscr{B}}, \ldots, [e_n]_{\mathscr{B}})}$$

La matrice d'une famille est donc constituée des coordonées des vecteurs dans la base donnée (en colonnes).

Matrice d'une application linéaire:

Soit E, F deux espaces vectoriels (de bases $\mathscr{B} = (e_i)_{i\in\mathbb{N}}$ et $\mathscr{C} = (f_i)_{i\in\mathbb{N}}$) et $f \in \mathcal{L}(E,F)$.

Alors on peut associer à l'application f **une unique matrice** dans les bases $\mathscr{B}, \mathscr{C}$, qu'on note alors $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f)$ qu'on construit comme suit:

$$\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f) = ([f(e_1)]_{\mathscr{C}}, [f(e_2)]_{\mathscr{C}}, \ldots, [f(e_n)]_{\mathscr{C}})$$

La matrice d'une application linéaire est donc constituée des coordonées dans la base d'arrivée de l'image des vecteurs de la base de départ (en colonnes).

Exemple: Considérons un endomorphisme de $\mathbb{R}^3$ et sa base canonique notée (e_1, e_2, e_3), telle que $f(x,y,z) = (2x+1, 3x, z-1)$. Alors on calcule l'image des vecteurs de la base de départ, ie:

- $f(1,0,0) = (3,3,-1)$
- $f(0,1,0) = (0,0,-1)$
- $f(0,0,1) = (0,0,-2)$

Puis on calcule les coordonées de ces vecteurs dans la base d'arrivée, et on les range en colonne dans une matrice et on obtient:

$$\begin{array}{c} f(e_1) \quad f(e_2) \quad f(e_3) \\ \downarrow \quad\quad \downarrow \quad\quad \downarrow \\ \begin{pmatrix} 3 & 0 & 0 \\ 3 & 0 & 0 \\ -1 & -1 & -2 \end{pmatrix} \begin{array}{l} \dashrightarrow f_1 \\ \dashrightarrow f_2 \\ \dashrightarrow f_3 \end{array} \end{array}$$

Cette construction implique qu'il suffit donc, pour une base donnée, de **calculer les images des vecteurs de cette base**, puis leur coordonées pour décrire la transformation.

Si E et F sont de dimensions respectives n et p, on peut alors construire **un isomorphisme fondamental**:

$$\mathrm{Mat}_{\mathscr{B},\mathscr{C}} : \mathcal{L}(E,F) \xrightarrow{\sim} \mathcal{M}_{n,p}(\mathbb{K})$$

On peut alors considèrer, que du point de vue de l'algèbre linéaire, l'espace des application linéaires et l'espace des matrices sont identiques[1].

Dans la suite, pour plus de lisibilité, on considèrera un endomorphisme f de $\mathbb{R}^2$ de matrice M dans la base canonique $\mathscr{C}$. Alors, appliquer une transformation linéaire à un vecteur $u \in E$ revient à **multiplier ce vecteur par la matrice associée**[2], ie on a:

$$[f(u)]_{\mathscr{C}} = M \times [u]_{\mathscr{C}}$$

Exemple: Soit f de matrice $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $u = (5,3)$, calculer les coordonées de $f(u)$ revient à calculer:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 27 \end{pmatrix}$$

Enfin l'isomorphisme entre ces deux espaces nous permet de définir **le noyau et l'image d'une matrice**, comme simplement étant le noyau ou l'image de l'application linéaire associée à cette matrice.

[1] En dimension finie

[2] Il suffit de partir du produit matriciel et d'utiliser la définition de M comme étant la matrice de f et les règles de calculs sur les vecteurs colonnes.

Matrices inversibles:

Soit $M \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée, alors on dit que M est **inversible** si il existe un inverse pour la loi de multiplication des matrices, ie si il existe une matrice A^{-1} telle que:

$$AA^{-1} = \mathrm{Id}_n$$

En particulier, si on considère l'application linéaire associée à M, alors:

$$M \text{ inversible} \iff f \text{ est inversible}$$

On appelle l'ensemble des matrices carrées inversibles de taille n **le groupe linéaire** d'ordre n, qu'on note $\mathrm{GL}_n(\mathbb{K})$, c'est un groupe pour la multiplication matricielle, ie on a:

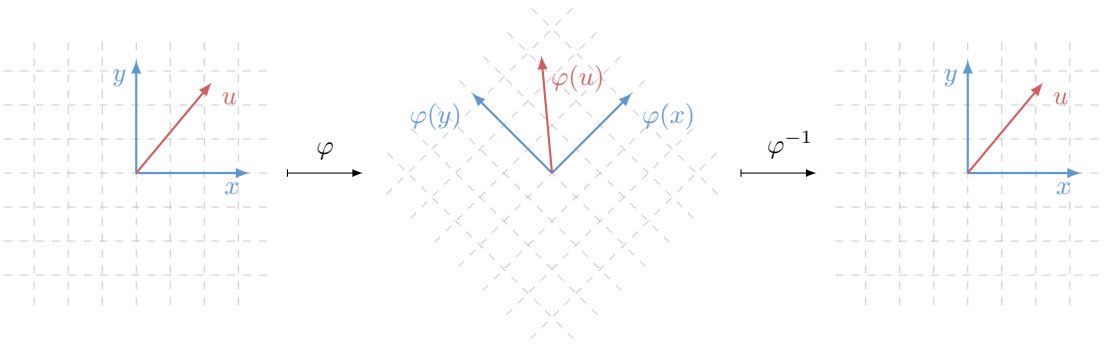
$$A, B \in \mathrm{GL}_n(\mathbb{K}) \iff AB \in \mathrm{GL}_n(\mathbb{K})$$

Et en particulier, on montre[1] que $(AB)^{-1} = B^{-1}A^{-1}$. Dans le cas d'une matrice inversible, on peut alors étendre notre définition de puissance d'une matrice au cas d'entiers négatifs.

Intuitivement:

*Le groupe $GL_n(\mathbb{K})$ est donc le groupe **des automorphismes** de E, c'est à dire le groupe **des transformations** (linéaires) de E qui sont **réversibles**.*

Visuellement, on comprends directement l'idée d'automorphisme d'un espace, son lien avec la bijectivité, la matrice inverse est alors la représentation algèbrique de la transformation inverse (ici avec une rotation de $\mathbb{R}^2$ de 45 degrés):



Enfin, on peut montrer que **la transposition** est compatible avec l'inverse, ie on a:

$$(M^{-1})^\top = (M^\top)^{-1}$$

Montrer qu'une matrice est inversible consiste simplement à calculer **le rang de la matrice**, comme expliqué dans la partie suivante.
Par ailleurs, calculer l'inverse peut se réaliser de plusieurs manières, la méthode la plus générale consiste à résoudre l'équation d'inconnue X de la forme:

$$AX = Y$$

On définira aussi une application fondamentale nommée **déterminant** qui sera définie dans le prochain chapitre et qui caractérise exactement toutes les matrices inversibles et donne une autre manière de calculer l'inverse.

[1] Intuitivement, c'est équivalent à l'inverse d'une composée car les matrices "sont" des applications au sens de l'algèbre linéaire.

Rang d'une matrice:

On a défini précédemment le concept de rang d'un famille et de rang d'une application linéaire, on peut généraliser cette définition en celle de **rang d'une matrice**, en effet on a simplement:

Le rang d'une matrice est la dimension de l'espace de ses colonnes.

Soit $M \in \mathcal{M}_n(\mathbb{K})$, dont on note les colonnes $\mathscr{F} = (C_1, C_2, \dots, C_n)$, alors si le rang de la matrice est n, on peut en déduire plusieurs propriétés:

- La matrice est inversible.
- La famille $\mathscr{F}$ est une base de $\mathbb{K}^n$.

On remarque donc que le rang nous donne **énormément d'informations** sur la matrice, la famille des colonnes, et l'application linéaire associée.

Matrice de passage:

On considère un espace vectoriel E et deux bases $\mathscr{F} = (e_1, \dots, e_n), \mathscr{F}' = (e'_1, \dots, e'_n)$ de cet espace.
On appelle **matrice de passage** de $\mathscr{F}$ à $\mathscr{F}'$ la matrice P qu'on construit comme suit:

$$\boxed{\text{Pass}(\mathscr{F}, \mathscr{F}') = ([e'_1]_{\mathscr{F}}, [e'_2]_{\mathscr{F}}, \dots, [e'_n]_{\mathscr{F}})}$$

La matrice de passage est donc constituée des coordonées dans l'ancienne base des vecteurs de la nouvelle base(en colonnes).

Exemple: Considèrons deux bases de $\mathbb{R}^2$ telles que $\mathscr{F} = [(1,2),(3,4)]$ et $\mathscr{F}' = [(5,6),(7,8)]$, alors on a:

$$[5,6]_{\mathscr{F}} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \qquad [7,8]_{\mathscr{F}} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

On range ensuite ces coordonées en colonnes dans la matrice, et on obtient:

$$\text{Pass}(\mathscr{F}, \mathscr{F}') = \begin{array}{c} \begin{matrix} e'_1 & e'_2 \\ \downarrow & \downarrow \end{matrix} \\ \begin{pmatrix} -1 & 3 \\ 2 & -2 \end{pmatrix} \begin{matrix} \rightarrow e_1 \\ \rightarrow e_2 \end{matrix} \end{array}$$

Une matrice de passage est alors **nécessairement inversible**[1], et si on note $U = [u]_{\mathscr{F}}$ et $U' = [u]_{\mathscr{F}'}$ on a alors le théorème fondamental suivant:

$$\boxed{U' = P^{-1}U}$$

La matrice de passage nous permet alors de représenter un même vecteur dans une base différente.

Considérons maintenant un endomorphisme f de matrices $M = \text{Mat}(\mathscr{F}, f), M' = \text{Mat}(\mathscr{F}', f)$ dans les deux bases respectivement, alors on a:

$$\boxed{M' = P^{-1}MP}$$

La matrice de passage nous permet alors de représenter un même endomorphisme dans une base différente.

[1] Car son rang est égal à sa dimension par construction.

Matrices semblables:

On dira alors que deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ sont **semblables** si il existe une matrice de passage telle que la relation ci-dessus soit vérifiée. Ce qui signifie exactement que:

> *Deux matrices semblables représentent* ***la même transformation géométrique*** *dans des bases différentes.*

En particulier on appelle alors **invariants de similitude** les propriétés qui **ne dépendent pas du choix de la base**, on peut alors montrer que:

> Le rang et la trace d'une matrice sont des invariants de similitude.

Un cas remarquable de matrices semblables est celui de la transposée, en effet dans un corps, une matrice est sa transposée sont semblables.

Matrices remarquables:

Dans ce chapitre nous allons présenter brièvement différentes matrices remarquables:

$$\begin{pmatrix} a & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & f \end{pmatrix}$$

Matrice diagonale

$$\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$$

Matrice triangulaire supérieure

$$\begin{pmatrix} a & 0 & 0 \\ b & d & 0 \\ c & e & f \end{pmatrix}$$

Matrice triangulaire inférieure

Puis on a deux types de matrices qui sont liées à la symétrie des coefficients et à **l'opération de transposition**[1]:

$$\begin{pmatrix} a & d & f \\ d & b & e \\ f & e & c \end{pmatrix}$$

Matrice symétrique

$$\begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix}$$

Matrice antisymétrique

On a aussi **les matrices élémentaires** qui sont obtenus par **opérations élémentaires**[2] sur la matrice identité:

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

Matrice de dilatation

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

Matrice de transvection

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

Matrice de permutation

Enfin, on a **les matrices orthogonales** qui sont des matrices telles que leur inverse soit **leur transposée**, ce sont par exemple les matrices de rotation:

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

Matrice de rotation

[1] En effet une matrice symétrique est égale à sa transposée et une matrice antisymétrique à **l'opposée** de sa transposée

[2] Voir la partie suivante.

Echelonnements & Calculs:

La théorie permettant de lier matrices, applications linéaires et familles de vecteurs, l'étude des espaces vectoriels engendrés par les colonnes ou les lignes d'une matrice revêt alors une importantce capitale qui est l'objet de cette partie.

On appelle **opérations élémentaires** sur une matrice M, l'une de ces 3 opérations:

- Multiplier une colonne par un scalaire
- Ajouter une colonnes à une autre
- Echanger deux colonnes

Il est alors possible de montrer que ces trois opérations préservent **le sous-espace engendré**[1] par les colonnes de la matrice, donc en particulier **le rang, le noyau et l'image** sont préservés.

Enfin, on appellera **matrice équivalente** à M et on notera $M' \sim M$ la matrice M à laquelle on a appliqué une ou plusieurs opérations élémentaires.
On dira qu'une matrice M est **échelonée** si la matrice obtenue aprés application d'opérations élémentaires est de la forme générale:

$$\begin{pmatrix} \alpha_1 & 0 & 0 & 0 & 0 \\ * & \alpha_2 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & \alpha_3 & 0 & 0 \end{pmatrix}$$

En d'autres termes, on **creuse** la matrice pour faire apparaître des zéros sur la partie triangulaire supérieure.
Il existe aussi une variante appellé **échelonnement avec mémorisation**, pour cette variante, on nomme chaque vecteur-colonne de la matrice et on reporte toutes les transformations réalisées sur ces vecteurs.

Présentons maintenant les différentes informations que nous pouvons tirer de ces échellonements:

1. On peut trés facilement trouver le rang d'une matrice, en effet aprés échelonnement, c'est simplement **le nombre de colonnes non-nulles** de la matrice. En particulier, il est donc trés facile de montrer qu'une matrice est inversible ou qu'une famille est une base.

2. On peut trouver une base d'un sous-espace engendré par une famille, pour cela on échelonne la matrice de cette famille, l'ensemble des colonnes non nulles consitue une base du sous-espace initial.

3. On peut trouver un supplémentaire d'une famille, on échelonne la matrice de la famille, et on complète la matrice par des vecteurs de la base canonique[2], ces vecteurs formeront alors un supplémentaire.

4. On peut trouver les équations cartésiennes d'un sous-espace engendré par une famille, pour cela on augmente la matrice de la famille par une matrice d'indéterminées, et on échelonne. Alors la dernière colonne fournira les equations cartésiennes du sous-espace.

5. On peut trouver le **noyau, l'image** d'une matrice, grâce à la variante **avec mémorisation**, alors les colonnes nulles aprés échelonnement permettent d'en déduire des vecteurs envoyés du noyau, et les colonnes non-nulles constituent une base de l'image (voir l'exemple ci-aprés).

Il peut être utile de noter que fondamentalement, on peut à la fois échelonner **sur les colonnes** ou sur **les lignes**. En effet, le rang est invariant par transposition, on aurait tout aussi bien pu développer des opération élémentaires sur les lignes, et les résultats seraient équivalents.

[1] L'échange ne change évidement rien, et le sous espace engendré ne change pas pas multiplication par un scalaire ou ajout d'un vecteur venant de la famille en question.
[2] Ou alors de telle sorte que la matrice reste échelonnée.

Soit $M = (C_1, C_2, C_3) = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 4 & 2 \\ 5 & 6 & 3 \end{pmatrix}$, nous allons présenter quelques exemples:

- Cherchons le rang de la matrice, on a:

$$M = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 4 & 2 \\ 5 & 6 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 2 \\ 5 & 4 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 5 & 4 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 3 & 2 \\ 5 & 4 \end{pmatrix}$$

  La matrice est échelonnée sur 2 colonnes, donc son rang est 2. En particulier, (C_1, C_2, C_3) n'est pas une base de $\mathbb{R}^3$, la matrice (donc l'endomorphisme associé) n'est pas inversible.

- Cherchons un sous-espace supplémentaire à (C_1, C_2, C_3) en reprenant l'échelonnement ci-dessus, on a simplement à rajouter des colonnes à M de sorte qu'elle soit échelonnée sur 3 colonnes, par exemple:

$$M' = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 5 & 4 & 1 \end{pmatrix}$$

  Et donc le sous-espace engendré par le vecteur $(0, 0, 1)$ est bien un supplémentaire de l'espace des colonnes.

- Cherchons des équations cartésiennes du sous-espace des colonnes, on a:

$$M = \begin{pmatrix} 1 & 2 & 1 & x \\ 3 & 4 & 2 & y \\ 5 & 6 & 3 & z \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 2 & 2 & y - 3x \\ 5 & 4 & 4 & z - 5x \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 5 & 4 & x - 2y + z & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 5 & 4 & x - 2y + z \end{pmatrix}$$

  Alors l'équation cartésienne du sous-espace des colonnes est $x - 2y + z = 0$.

- Cherchons le noyau et l'image de l'endomorphisme représenté par M, on doit utiliser la mémorisation, ie:

$$\begin{matrix} C_1 & C_2 & C_3 \end{matrix} \quad \begin{matrix} C_1 & C_2 - 2C_1 & 2C_3 - 2C_1 \end{matrix} \quad \begin{matrix} C_1 & C_2 - 2C_1 & 2C_3 - C_2 \end{matrix}$$
$$M = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 4 & 2 \\ 5 & 6 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 2 \\ 5 & 4 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 5 & 4 & 0 \end{pmatrix}$$

  On en conclut qu'une base de l'image est $[(1, 3, 5), (0, 2, 4)]$.

  Le noyau est de dimension 1 et par la mémorisation, une combinaison linéaire nulle est $0C_1 - C_2 + 2C_3$, le vecteur $(0, -1, 2)$ est bien dans le noyau et est donc une base du noyau.

III — Déterminant

Soit $A = (a_{i,j})$ une matrice carrée de taille n, nous cherchons à définir une application $\phi : \mathcal{M}_n(\mathbb{K}) \longrightarrow \mathbb{R}$ telle que:

$$A \in \mathrm{GL}_n(\mathbb{K}) \iff \phi(A) = 0$$

Définition:

On peut alors montrer[1] qu'une telle application existe qu'on appellera **déterminant**, qu'on note $|A|$ et qu'on définit alors par récurrence:

- Si $n = 1$ alors $|(a)| = a$
- Sinon $|A| := a_{1,1}|A_{1,1}| - a_{1,2}|A_{1,2}| + \ldots + (-1)^{n+1}a_{1,n}|A_{1,n}|$

Où la notation $A_{1,1}$ signifie la matrice A à laquelle on a retiré la première ligne et la première colonne.

Cette opération s'appelle aussi **développement selon la première ligne** du déterminant, elle se comprends visuellement par:

$$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} \square & \square & \square \\ \square & e & f \\ \square & h & i \end{vmatrix} - b \begin{vmatrix} \square & \square & \square \\ d & \square & f \\ g & \square & i \end{vmatrix} + c \begin{vmatrix} \square & \square & \square \\ d & e & \square \\ g & h & \square \end{vmatrix}$$

On peut ainsi **développer le déterminant** selon n'importe quelle ligne ou colonne en suivant cette règle. On régresse ainsi vers des déterminants de plus petite taille, et aprés plusieurs étapes, vers un réel.

Propriétés:

On peut montrer que cette application est **une forme multilinéaire alternée** en les colonnes ou lignes de la matrice, en particulier, on a les propriétés suivantes:

- Ajouter à une colonne une combinaison linéaire **des autres colonnes** ne change pas le déterminant.
- Echanger deux colonnes d'une matrice change le signe du determinant.
- Si deux colonnes sont égales ou qu'une des colonnes est nulle, le déterminant est nul.

En outre la multilinéarité sur les colonnes ou lignes implique:

$$\boxed{|\lambda A| = \lambda^n |A|}$$

Et on peut aussi montrer que le déterminant est multiplicatif, ie:

$$\boxed{|A \times B| = |A| \times |B|}$$

Si on considère une famille $\mathscr{F} = (e_1, \ldots, e_n)$ de vecteurs de coordonées $(C_1, \ldots, C_n)$ dans une certaine base $\mathscr{B}$, alors on peut définir le déterminant **d'une famille de vecteurs** dans cette base par:

$$|\mathscr{F}|_{\mathscr{B}} = |(C_1, \ldots, C_n)|$$

Le déterminant d'une famille par rapport à $\mathscr{B}$ est alors simplement le déterminant de la matrice construite avec les coordonées de ses vecteurs dans la base $\mathscr{B}$.

Soit deux bases $\mathscr{B}, \mathscr{B}'$, on pourrait alors considèrer **l'effet d'un changement de base** sur un tel déterminant et on a alors:

$$\boxed{|\mathscr{F}|_{\mathscr{B}'} = |\mathscr{F}|_{\mathscr{B}} \times |\mathscr{B}'|_{\mathscr{B}}}$$

[1] Particulièrement difficile..

Cofacteurs:

Reprenons en simplifant la définition donnée plus haut du développement selon la première ligne:

$$|A| := a_{1,1}|A_{1,1}| - a_{1,2}|A_{1,2}| + \ldots + (-1)^{n+1}a_{1,n}|A_{1,n}| = \sum_{j=1}^{n} a_{1,j}(-1)^{j+1}|A_{1,j}|$$

On appelle alors **cofacteur** de l'élément $a_{i,j}$ le scalaire:

$$\boxed{\text{cof}(a_{i,j}) := (-1)^{j+i}|A_{i,j}|}$$

Intuitivement en reprenant la visualisation exposée plus haut:

$$a\begin{vmatrix} \square & \square & \square \\ \square & e & f \\ \square & h & i \end{vmatrix} - b\begin{vmatrix} \square & \square & \square \\ d & \square & f \\ g & \square & i \end{vmatrix} + c\begin{vmatrix} \square & \square & \square \\ d & e & \square \\ g & h & \square \end{vmatrix}$$

Ce sont simplement les déterminants mineurs que l'on calcule à chaque itération **en tenant compte du signe qui précède le coefficient**.
En pratique pour trouver le signe du cofacteur, on ne calcule pas $(-1)^{i+j}$, mais on le détermine par **la règle de l'échiquier** qu'on peut se représenter comme suit:

$$\begin{vmatrix} + & - & + & - & \ldots \\ - & + & - & + & \ldots \\ + & - & + & - & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix}$$

Exemple: Pour $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, le cofacteur du coefficient en position $(1,2)$ est $-\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix}$

Comatrice:

On peut alors définir **la comatrice** d'une matrice A donnée, c'est en fait simplement **la matrice des cofacteurs** de A, ie le terme en position (i,j) de la comatrice est exactement le cofacteur de l'élement en position (i,j).

L'intérêt principal de la comatrice est de calculer l'inverse de matrices inversibles, en effet, on peut montrer l'identité:

$$A \times \text{com}(A)^\top = |A|I_n$$

Et en particulier si le déterminant est non-nul (donc si A est inversible), on a:

$$\boxed{A^{-1} = \frac{1}{|A|}\text{com}(A)^\top}$$

Cette formule est néanmoins particulièrement inexploitable pour $n > 3$ du fait de la quantité de calcul à réaliser.

Formules de Cramer:

On considère un systême linéaire de n équations à n inconnues, alors on peut l'écrire matriciellement sous la forme:

$$AX = B$$

Pour $A \in \mathcal{M}_n(\mathbb{K})$, X un vecteur indéterminé, et B le vecteur fixé du second membre des équations.

Un tel système est dit **système de Cramer** si la matrice A est inversible, il possède alors une unique solution qui s'écrit matriciellement $X = A^{-1}B$.

Alors on montre que sa i-ème inconnue x_i (et par suite la solution tout entière X = $(x_1, \ldots, x_n)$) est donnée par:

$$\boxed{x_i = \frac{|A_i|}{|A|}}$$

Dans laquelle A_i désigne la matrice obtenue en remplaçant la i-ème colonne de A par le second membre B

Exemple:

$$\begin{cases} x + 4y + 8z = 2 \\ 2x + 5y + 8z = 2 \\ 3x + 6y + 9z = 3 \end{cases} \iff \begin{pmatrix} 1 & 4 & 8 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$$

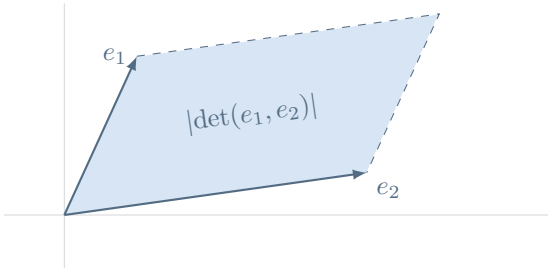
Alors on a:

$$y = \frac{\begin{vmatrix} 1 & 2 & 8 \\ 2 & 2 & 8 \\ 3 & 3 & 9 \end{vmatrix}}{|A|} = \frac{6}{-3} = -2$$

Volume orienté:

Le déterminant admet un interprétation géométrique intéressante liée au concept d'aire, de volumes, et d'hypervolumes de manière générale.

Considérons le cas simple d'une base $\mathscr{B} = (e_1, e_2)$ de vecteurs de $\mathbb{R}^2$, alors le déterminant de cette base correspond **à l'aire algébrique[2] du parallélogramme formé par les deux vecteurs**, ie:



Plus généralement, le déterminant d'une famille de n vecteurs est **l'hypervolume algébrique** du parallélotope formé par ces n vecteurs.

Orientation:

On peut alors définir **l'orientation d'un espace vectoriel**, en effet, on dira que deux bases $\mathscr{B}, \mathscr{B}'$ ont meme orientation si et seulement si:

$$\det(\text{Pass}(\mathscr{B}, \mathscr{B}')) > 0$$

Si on fixe alors une base orientée canoniquement, on qualifie son orientation (et celles de toutes les bases de meme orientation) de **directe** et les bases d'orientation opposée sont alors d'orientation **indirecte**.

Exemple: Dans le cas de $\mathbb{R}^n$, la base canonique est la base qui, par convention, est d'orientation directe.

On peut alors définir le **groupe spécial linéaire**, noté SL($\mathbb{R}$), des endomorphismes qui respectent l'orientation de l'espace.

[2] C'est une aire "signée", ie l'aire géométrique est donc la valeur absolue de cette aire algébrique.

III — Dualité

On définit dans ce chapitre une notion fondamentale en algèbre linéaire, trés liée à celle de produit scalaire, qui est celle **d'espace dual** d'un espace vectoriel E, qu'on notera E^* et qu'on définit par:

L'espace dual d'un espace vectoriel est l'ensemble des formes linéaires sur cet espace.

Attention, on parle ici de dual **algébrique**, on peut aussi définir un dual **topologique** en requierant que les formes linéaires considérées soit continues. Dans la plupart des exemples, on considèrera $E = \mathbb{R}^2$ et donc par exemple un élément de E^* est:

$$\phi : (x, y) \mapsto 2x + y$$

Dans toute la suite, on se placera dans un espace E de dimension finie[1]

Notations:

On introduit de nouvelles notations pratique, tout d'abord on utilisera souvent la notation **delta de Kronecker** qui pour tout $n, m \in \mathbb{N}$ donne:

$$\begin{cases} \delta_n^m = 1 \text{ si } n = m \\ \delta_n^m = 0 \text{ sinon} \end{cases}$$

Propriétés:

On souhaite caractériser la forme d'une forme linéaire ϕ de l'espace dual. On a directement par linéarité que:

$$\forall x \in E \ ; \ \phi(x) = \phi\left(\sum_i x^i e_i\right) = \sum_i x^i \phi(e_i)$$

Donc en particulier, on a que évidemment que ϕ est caractérisée par ses images des vecteurs de base, et réciproquement si une application ϕ' est telle que $\phi = \sum_i x_i a_i$, alors ϕ' est une forme linéaire sur E.

Base duale:

En dimension finie, on a alors le théorème fondamental suivant:

L'espace vectoriel et son dual sont isomorphes.

En effet, on a tout d'abord égalité des dimensions[2]. Fixons une base $\mathscr{B} = (e_1, \dots, e_n)$ de E, alors si on considère la famille de formes linéaires suivantes:

$$\begin{aligned} e_i^* : E &\longrightarrow \mathbb{R} \\ \sum_i x^i e_i &\longmapsto x_i \end{aligned}$$

Alors elle forme alors une base de E^* qu'on appelera **base duale** de $\mathscr{B}$ et qu'on notera $\mathscr{B}^*$.

C'est une famille de projections qui a chaque vecteur associe la coordonée correspondante.

Elle vérifie en outre la relation suivante:

$$e_i^*(e_j) = \delta_i^j$$

Base antéduale:

On considère alors le problème inverse, et on se donne une base $\mathscr{B} = (\phi_1, \dots, \phi_n)$ du dual de E, alors il existe une unique base $(e_1, \dots, e_n)$ de E telle que:

$$\forall i \in [\![1 \ ; \ n]\!] \ ; \ \phi_i = e_i^*$$

Ou dit autrement il existe une base $\mathscr{C}$ de E telle que $\mathscr{C}^* = \mathscr{B}$.

[1] Tout ce qui suit est en général faux en dimension infinie, l'idée prinicipale étant qu'en dimension infinie, des considérations topologiques sont **nécéssaires**.

[2] Car $\dim \mathcal{L}(E, F) = \dim \mathcal{L}(E) \dim \mathcal{L}(F)$

Hyperplans:

On rapelle alors que d'aprés le cours de première année, on appelle hyperplan tout sous-espace dont le supplémentaire est une droite, on peut alors caractériser les hyperplans en termes de formes linéaires par la proposition suivante:

Les hyperplans sont exactements les noyaux de formes linéaires.

Vecteurs covariants et contravariants:

On deux bases $\mathscr{C}, \mathscr{B}$ de E et un vecteur u de coordonées respectives U, U', on a la propriété de changement de base suivante pour des notations évidentes:

$$U' = P^{-1}U$$

On voit alors ici que les coordonées aprés le changement de base dépendent de **l'inverse de la matrice de passage**. On dira dans ce cas qu'un vecteur est **contravariant par rapport aux bases de** E.

Maintenant on considère un vecteur $u \in E^*$, alors on peut calculer les bases duales $\mathscr{C}^*, \mathscr{B}^*$ et utiliser la même formule pour montrer que pour des notations évidentes on a:

$$U' = \text{Pass}(\mathscr{B}^*, \mathscr{C}^*)U$$

Or, on peut alors montrer la propriété suivante fondamentale suivante:

$$\text{Pass}(\mathscr{B}^*, \mathscr{C}^*) = {}^tP$$

En particulier on se ramène alors à une matrice de passage entre les bases de E et on a alors la formule de changement de base suivantes (pour les coordonées placées en ligne):

$${}^tU' = {}^tUP$$

On voit alors ici que les coordonées aprés le changement de base (en ligne) dépendent de **la matrice de passage**. On dira dans ce cas qu'un vecteur est **covariant par rapport aux bases de** E. Finalement moralement on a les formules de changement de bases suivantes:

- Si X, Y sont des vecteurs : $Y = P^{-1}X$
- Si X, Y sont des covecteurs : ${}^tY = {}^tXP$

La notion de vecteurs contravariants et covariants est centrale en algèbre multilinéaire, elle permet la définition d'objets généraux appelés tenseurs qui généralisent ce concept.

III — Introduction à la réduction

Dans ce chapitre, nous étudirons un domaine vaste de l'algèbre linéaire appellé **réduction des endormorphismes**.
En effet, sous une forme quelconque un endomorphisme représenté par une matrice présente plusieurs problèmes:

- Il est couteux de calculer les puissances d'une matrice quelconque.
- Une représentation quelconque donne peu d'informations sur l'endomorphisme.

L'objectif sera donc de réduire (comprendre simplifier) la représentation de l'endormorphisme, et de le représenter par une matrice plus simple.

Soit $f \in \mathcal{L}(E)$ et une famille (E_i) de sous-espaces **supplémentaires**[1], alors on peut construire une base $\mathcal{B}$ dite **base adaptée à la décomposition** en concaténant des bases respectives des E_i.

Elements Propres :

Soit $\lambda \in \mathbb{R}$, on dit que λ est une **valeur propre**[2] de l'endomorphisme f si et seulement si il existe un vecteur non-nul $u \in E$ tel que:

$$f(u) = \lambda u$$

On dira alors qu'un tel vecteur est **vecteur propre** de l'endomorphisme et on appelle **sous-espace propre** associé à la valeur propre λ l'ensemble des vecteurs propres associés, qu'on note E_λ dont on déduit une expression[3]:

$$E_\lambda = \mathrm{Ker}(f - \lambda \mathrm{Id})$$

Enfin, on peut montrer une propriété trés importante pour la suite:

Toute somme de sous-espaces propres est **directe**.

L'image des sous-espaces propres par l'endomorphisme se réduit à une homotéthie de rapport la valeur propre.

Sous-Espaces Stables :

On dit que F est **stable par l'endormorphisme** si et seulement si:

$$f(F) \subseteq F$$

On considère maintant une base de F qu'on compléte en une base de E via le théorème de la base incomplète, alors dans une telle base, l'endormorphisme est représenté par la matrice par blocs:

$$\left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$$

[1] Comme tout sous-espace admet un supplémentaire (théorème de la base incomplète), on peut définir une base adaptée à **un seul** sous-espace comme étant une base adaptée à la somme directe de ce sous-espace et de son supplémentaire, qui revient à compléter la base en une base de E.
[2] On appelle **specte**, noté $\mathrm{Sp}(f)$ l'ensemble des valeurs propres de f.
[3] Directement d'après la définition d'un vecteur propre associé à λ.

Plus généralement, si on a une famille (E_i) de sous-espaces **stables et supplémentaires**, ie $E = \bigoplus_{i\in\mathbb{N}} E_i$, et $\mathcal{B}$ est une base adaptée à cette décomposition, alors dans cette base, f est représentée par la matrice diagonale par blocs:

$$\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_n \end{pmatrix}$$

On remarque donc que la stabilité des sous-espaces nous permet de reprénsenter notre transformation de manière plus simple, en particulier, on peut alors montrer une propriété fondamentale:

Tout les sous-espaces propres sont stables.

Polynôme caractéristique :

On peut montrer[1] que λ est valeur propre si et seulement si:

$$E_\lambda \neq \{0_E\} \Longleftrightarrow \det(f - \lambda \mathrm{Id}) = 0$$

On définit alors le **polynôme caractéristique** d'un endormorphisme par:

$$P_f = \det(f - X\mathrm{Id})$$

En particulier, on peut donc montrer que:

Les valeurs propres sont exactement les racines du polynôme caractéristique.

Ceci nous donne donc une méthode systématique pour trouver les valeurs propres d'un endomorphisme. En particulier, on peut alors montrer les identités suivantes, utiles dans la recherche de valeurs propres:

- La somme des valeurs propres est égale à **la trace de la matrice**.
- Le produit des valeurs propres est égale au **déterminant de la matrice**.

Diagonalisation :

Les endomorphismes qu'on peut représenter le plus simplement sont ceux qui ceux réduisent (dans une base bien choisie) à une homotéthie des vecteurs de la base. On dira alors que ces endomorphismes sont **diagonalisables**, formellement:

Un endormorphisme est diagonalisable si il existe une base de E constituée de vecteurs propres de f.

De manière équivalente:

Un endormorphisme est diagonalisable si ses matrices sont semblables à une matrice diagonale.

Exemple: Soit f un tel endormorphisme de $\mathbb{R}^3$ et $(e_{\lambda_1}, e_{\lambda_2}, e_{\lambda_3})$ des tels vecteurs propres, alors dans cette base, f est représenté par la matrice:

$$D = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}$$

Ou encore si A est la matrice de f dans la base canonique, alors $A = PDP^{-1}$ avec:

$$P = ([e_{\lambda_1}]_{\mathscr{C}}, [e_{\lambda_2}]_{\mathscr{C}}, [e_{\lambda_3}]_{\mathscr{C}}) \text{ et } D = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}$$

Diagonaliser un endomorphisme revient à ***décomposer l'espace en somme directe de droites stables****.*

[1] E_λ est un noyau, il suffit de caractériser le fait qu'il soit non vide en termes de la bijectivité d'un certain endomorphisme.

Critères de diagonalisabilité :

On sait qu'un endomorphisme f est diagonalisable si et seulement si il admet une base de vecteurs propres, alors on peut montrer[1] que f est diagonalisable si et seulement si:

$$\boxed{E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} E_\lambda}$$

En particulier on remarque alors que montrer la supplémentarité revient à montrer que **la somme des dimensions des sous-espaces propres est égale à la dimension totale** car toute somme de sous-espaces propres est directe.

On peut alors montrer que si λ est une valeur propre de multiplicité α pour le polynome caractéristique, alors:

$$1 \leq \dim(E_\lambda) \leq \alpha$$

On a alors le théorème fondamental suivant:

Un endomorphisme est diagonalisable si et seulement si son polynôme est scindé sur $\mathbb{K}$ et que la dimension de chaque sous-espace propre est égale à la multiplicité de la valeur propre associée.

On peut donc étudier si un endormorphisme est diagonalisable en calculant les valeurs propres et les dimensions des sous-espaces propres associés, ce qui revient à un calcul de polynôme caractéristique suivi de calculs de noyau.

Exemple: Diagonalisons la matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}$

On peut calculer $P_A = \det(A - XI_3) = X^2(X-6)$, on a alors deux sous-espaces propres, E_0 et E_6 et on sait que E_0 est de dimension 1, il suffit alors de vérifier que E_6 est bien de dimension 2 pour conclure que $\sum \dim(E_\lambda) = \dim(E)$ et donc que la matrice est diagonalisable.

Pour trouver une base de vecteurs propres, il suffit alors de trouver une base de E_0, E_6 et de la concaténer en une base de E.

[1] En effet, si tel est le cas, alors il suffit de prendre une base pour chaque E_λ (qui est bien constituée de vecteurs propres par définition), et de les concaténer pour obtenir une base de E constituée de vecteurs propres.

III — Polynomes d'endomorphismes

L'objet principal de ce chapitre est l'étude des polynômes d'endomorphismes et de matrices, en effet, les matrices est les endomorphismes formant une algèbre, on peut en calculer des puissances, des sommes, et effectuer un multiplication externe, on peut donc définir des **polynômes de matrices/d'endomorphismes**, en effet si on a $P = \sum_{k=0}^{n} a_k X^k \in \mathbb{K}[X]$ et u un endomorphisme, on définit alors:

$$P(u) = \sum_{k=0}^{n} a_k u^k$$

Propriétés :

On définit alors un **morphisme d'algèbre** pour un endomorphisme donné par:

$$\begin{aligned} \phi_u : \mathbb{K}[X] &\longrightarrow \mathcal{L}(E) \\ P &\longmapsto P(u) \end{aligned}$$

En particulier, on a donc $PQ(u) = P(u) \circ Q(u)$, on peut alors en déduire la proposition suivante:

Deux polynôme d'un même endomorphisme commutent.

On peut alors montrer que les polynômes d'endomorphismes ont un bon comportement vis-à-vis des changements de bases, en particulier si $A = PBP^{-1}$, pour tout polynôme Q, on montre facilement que:

$$Q(A) = PQ(B)P^{-1}$$

Enfin, on montre aussi que si u est représenté par A dans une base, alors u^k est représenté par A^k dans cette même base, et donc par linéarité $P(u)$ est représenté par $P(A)$ dans cette base. En particulier, le polynôme d'un endomorphisme ne dépends alors par de la représentation choisie.

Valeurs propres d'un polynôme d'endomorphisme :

Pour un endormorphisme u admettant une valeur propre λ de vecteur propre associé v, on peut alors étudier le lien entre les polynômes d'endomorphisme et les valeurs propres, et en particulier, on peut montre que λ^k est valeur propre de u^k et donc par linéarité que:

$$P(u)(V) = P(\lambda)(V)$$

Et donc que $P(\lambda)$ est valeur propre de $P(u)$.

Polynômes annulateurs :

On considère $u \in \mathcal{L}(E)$, et on définit l'ensemble des **annulateurs** de u par:

$$\mathscr{A}_u := \left\{ P \in \mathbb{K}[X] \; ; \; P(u) = 0_{\mathcal{L}(E)} \right\}$$

Ce sont l'ensemble des polynômes qui annulent u. On définit de même les polynômes annulateurs de matrices. Une propriété fondamentale est alors que cette ensemble n'est jamais vide[1], en effet on a que:

Tout endormoprhisme admet un polynôme annulateur non-nul.

On peut alors étudier le lien entre les valeurs propres d'un endomorphisme et ses annulateurs, et on peut alors montrer la propriété suivante:

$$\mathrm{Sp}(u) \subseteq \left\{ \alpha \in \mathbb{K} \; ; \; P(\alpha) = 0 \right\}$$

Les seules valeurs propres possibles sont les racines de l'annulateur.

[1] Il suffit de considèrer la dimension de E, et une famille plus grande que cette dimension, donc liée, et on peut alors trouver un polynôme en u qui s'annule.

Néanmoins il n'y a pas équivalence, plus précisément, si $A_u \in \mathscr{A}_u$, et si on définit $Q_u = (X-\lambda_1)\dots(X-\lambda_k)$ où les (λ_i) sont toutes les valeurs propres de u, alors on a:

$$\boxed{Q_u \mid A_u}$$

Polynôme minimal :

On considère l'ensembe des annulateurs d'un endomorphisme u, alors il est non-vide comme énoncé ci-dessus, et il admet aussi **un plus petit élément unitaire** (au sens du degré) et il est unique.

On appelle alors ce plus petit élément **le polynôme minimal** de u qu'on note M_u

Théorème de Cayley-Hamilton :

On peut alors énoncer le théorème fondamental de la réduction des endormorphismes, ie le **théorème de Cayley-Hamilton**:

Le polynôme caractéristique est un annulateur.

La démonstration, non-triviale, se fait par un argument topologique et par la continuité de la fonction polynôme caractéristique. On a donc la relation avec le polynôme minimal suivante:

$$M_u \mid P_u$$

Lemme des noyaux :

On s'intéresse finalement aux **noyaux de polynômes d'endomorphismes** pour pouvoir énoncer le dernier théorème de cette partie. On peut tout d'abord montrer facilement le résultat suivant:

Le noyau d'un polynôme d'endomorphisme est stable par celui-ci.

Soit P, Q deux polynômes **premiers entre eux**, on peut alors montrer[1] le **lemme des noyaux**, c'est à dire que:

$$\boxed{\mathrm{Ker}PQ(u) \;=\; \mathrm{Ker}P(u) \oplus \mathrm{Ker}Q(u)}$$

En particulier, pour P un polynôme annulateur de u qui se décompose en $P_1, \dots, P_k$, on a la décomposition suivante de l'espace tout entier:

$$E = \bigoplus_{k=1}^{n} \mathrm{Ker}P_k(u)$$

Caractérisations via les annulateurs :

On peut caractériser la diagonalisabilité via les annulateurs, en effet, on peut montre via le lemme des noyaux qu'on a:

Un endomorphisme est diagonalisable si et seulement si il admet un annulateur scindé à racines simples.

On peut aussi caractériser la trigonalisabilité par:

Un endomorphisme est trigonalisabilité si et seulement si il admet un annulateur scindé.

[1] La démonstration est non-trivial et fait appel à la relation de Bezout pour les polynômes.

III — Trigonalisation

Les endomorphismes qu'on ne peut représenter sous forme diagonale nous posent alors problème, on cherche alors dans ce chapitre à mobiliser la théorie des polynômes d'endomorphismes pour comprendre les conditions pour représenter de tels endomorphismes sous une forme plus simple triangulaire, ou sous **forme de Dunford** qui sera présentée ci-dessous.

Critère de trigonalisation :

On peut montrer le critère suivant:

Un endomorphisme u est trigonalisable sur $\mathbb{K}$ si et seulement si son polynôme caractéristique est scindé sur $\mathbb{K}$.

En particulier tout les endormorphismes sont trigonalisables dans $\mathbb{C}$.

Néanmoins, on comprends vite qu'une forme triangulaire quelconque sera peu utile car on ne pourra calculer ses puissances facilement, on peut alors montrer que si u est trigonalisable, il admet une forme plus simple encore appellée **forme de Dunford**:

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

C'est une matrice **triangulaire par blocs triangulaires**. Et les puissances de telles matrices sont alors facile à calculer via le produit par blocs et le binôme de Newton. En effet chaque bloc et de la forme $\lambda I_n + N$ avec N nilpotente, donc le binôme simplifie grandement les calculs.

Structure des noyaux itérés :

Soit u un endomorphisme, alors on peut montrer que les noyaux des puissances de u forment la structure suivante:

$$\mathrm{Ker} u \subsetneq \mathrm{Ker} u^2 \subsetneq \ldots \subsetneq \mathrm{Ker} u^k$$

Et cette suite de noyaux itérés est **stationnaire**, en particulier, si u est nilpotent, elle est stationnaire et le dernier sous espace est E tout entier.

Sous-espaces caractéristiques :

Soit u un endomorphisme de polynôme caractéristique $P_u = (X-\lambda_1)^{\alpha_1} \ldots (X-\lambda_k)^{\alpha_k}$, alors on appelle **sous-espace caractéristique** associé à la valeur propre λ_k le sous-espace suivant:

$$F_{\lambda_k} = \mathrm{Ker}(u - \lambda_k \mathrm{Id})^{\alpha_k}$$

On sait que les $(X-\lambda_k)^{\alpha_k}$ sont premiers entre eux, donc d'aprés le théorème de Cayley-Hamilton et le lemme des noyaux, on a:

$$E = F_{\lambda_1} \oplus \ldots \oplus F_{\lambda_k}$$

Et donc en particulier on a $\dim(F_{\lambda_1}) = \alpha_1$.

Les sous-espaces caractéristiques sont des sous-espaces propres "sympathiques".

C'est sont aussi des noyaux de polynômes d'endomorphismes donc en particulier, ils sont stables par u. Par ailleurs, d'aprés la structure des noyaux itérés, on a:

$$E_{\lambda_k} = \text{Ker}(u - \lambda_k \text{Id})^1 \subsetneq \text{Ker}(u - \lambda_k \text{Id})^2 \subsetneq \ldots \subsetneq \text{Ker}(u - \lambda_k \text{Id})^{\alpha_k} = F_{\lambda_k}$$

Ce sont ces sous-espaces qui nous permettront de construire un base de E dans laquelle u est représenté par une matrice de Dunford.

Trigonalisation de Dunford :

On peut alors définir une méthode générale de trigonalisation de Dunford, on considère un endomorphisme u et son polynôme caractéristique, alors on obtient une base de trigonalisation de Dunford par l'algorithme suivant:

- Si la dimension du sous-espace propre E_λ est égale à la multiplicité, le bloc associé à λ est diagonal, et la base recherchée est une base du sous-espace propre
- Sinon, on calcule une **base adaptée** aux noyaux itérés $E_\lambda \subsetneq \text{Ker}(u - \lambda \text{Id})^2 \subsetneq \ldots \subsetneq F_\lambda$ via le théorème de la base incomplète puis les coordonées de l'image de cette base par u pour obtenir le bloc associé à λ.

Exemple: Dans toute la suite nous considérerons l'exemple de l'endomorphisme de $\mathbb{R}^5$ de polynôme caractéristique $P_u = (X-2)^2(X-3)^3$, alors d'après le lemme des noyaux et le théorème de Cayley-Hamilton, on a:

$$E = F_2 \oplus F_3$$

Les valeurs propres sont $2, 3$ et on supposera que la première valeur propre est telle que la dimension du sous-espace propre est égale à la multiplicité, alors on trouve aisément une base e_1, e_2 du bloc associé à 2, il sera diagonal.

Il nous suffit alors de trouver une base de $F_3 = \text{Ker}(u - 3\text{Id})^3$ qu'on va calculer de la manière suivante:

- On calcule une base de $E_3 = \text{Ker}u - 3\text{Id}$
- On la complète en une base de $\text{Ker}(u - 3\text{Id})^2$
- On la complète en une base de $F_3 = \text{Ker}(u - 3\text{Id})^3$

Finalement, on a F_3 de dimension 3 et donc une base e_3, e_4, e_5 de F_3. La base finale recherchée est donc $(e_1, e_2, e_3, e_4, e_5)$ et dans cette base la matrice est de la forme:

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & * & * \\ 0 & 0 & 0 & 3 & * \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Où les coefficients $*$ sont donnés par le calcul des cordonnées des images des vecteurs dans la base.

III — Applications de la réduction

Dans cette dernière partie, on va maintenant pouvoir développer les applications possibles de la réduction dans la résolution de problèmes variés. On considère ici le cas d'une matrice

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & * \\ 0 & 0 & 3 \end{pmatrix} P^{-1}$$

Dans le cas plus simple de matrice diagonalisable, tout les calculs sont plus simples et les mêmes méthodes s'appliquent.

Calculs de puissances :

La première étape pour calculer une puissance de matrice via la réduction est de remarquer que:

$$A^k = (PTP^{-1})^k = PT^kP^{-1}$$

Or, T^k se calcule alors par blocs et on a:

$$T^k = \begin{pmatrix} B_1^k & \\ & B_2^k \end{pmatrix} = \begin{pmatrix} 1 & \\ & B_2^k \end{pmatrix}$$

Et on a alors $B_2 = 3\mathrm{Id} + N$ avec N strictement triangulaire donc nilpotente, et donc on calcule facilement sa puissance via le binôme de Newton car Id commute toujours.

Suites récurrentes :

Soit 3 suites u_n, v_n, w_n telles que $u_1 = 1, v_1 = 1, w_1 = 1$on considère maintenant le **système de suites récurrentes** suivant:

$$\begin{cases} u_n = 2u_{n-1} + v_{n-1} + w_{n-1} \\ v_n = u_{n-1} + 2v_{n-1} + w_{n-1} \\ w_n = 3w_{n-1} \end{cases}$$

On pose alors $U_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$ et le système se réécrit alors sous la forme matricielle suivante:

$$U_n = AU_{n-1}$$

Par récurrence on trouve alors que $U_n = A^nU_1$, donc en particulier sachant U_1, il nous suffit alors de calculer A^k comme précédemment ainsi que la matrice de passage et son inverse pour réussir à trouver le terme général de u_n, v_n et w_n.

Plus subtilement, cette méthode s'applique aussi aux suites récurrentes d'ordre multiple, considérons par exemple la suite u_n de premiers termes $u_1 = 1$ et $u_2 = 2$:

$$u_n = u_{n-1} + u_{n-2}$$

En effet si on pose $U_n = \begin{pmatrix} u_{n-2} \\ u_{n-1} \\ u_n \end{pmatrix}$ alors on a l'expression matricielle:

$$U_n = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} U_{n-1}$$

Alors si la matrice est diagonalisable[1], on peut alors trouver une expression de U_n en fonction de U_0 et alors une expression de u_n simplement en fonction de n.

[1] La matrice associée à la suite de Fibonacci n'est pas réductible dans $\mathbb{R}$ donc on ne peut pas trouver une expression de son terme général.

Systèmes différentiels :

On considère trois fonctions réelles f, g, h de classes $\mathcal{C}^1$ et on cherche à résoudre le système différentiel suivant:

$$\begin{cases} f'(x) = 2f(x) + g(x) + h(x) \\ g'(x) = f(x) + 2g(x) + h(x) \\ h'(x) = 3h(x) \end{cases}$$

On pose alors $F(x) = \begin{pmatrix} f(x) \\ g(x) \\ h(x) \end{pmatrix}$ et le système se réécrit alors sous la forme matricielle suivante:

$$F'(x) = AF(x) = PTP^{-1}F(x)$$

On pose alors $G(x) = P^{-1}F(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \\ g_3(x) \end{pmatrix}$, alors on se ramène à l'équation matricielle suivante:

$$G'(x) = TG(x)$$

Alors on s'est ramené à un système triangulaire que l'on sait résoudre en partant du bas.

On peut donc résoudre pour $G(x)$, et alors $F(x)$ est égal à $PG(x)$ (en utilisant la définition de $G(x)$) et on a donc trouvé les fonctions qui satisfont le système. Si on a des conditions initiales, on peut alors résoudre pour trouve l'unique triplet qui le satisfait.

IV — Espaces Quadratiques

On se donne un espace vectoriel E, on dira que c'est un **espace quadratique** si et seulement si on peut définir une **forme bilinéaire symétrique** sur cet espace.

Alors on pourra alors définir une **forme quadratique** sur cet espace qui est une application $q : E \to \mathbb{K}$ telle que pour une certaine forme bilinéaire symétrique f, on ait:

$$q(x) = f(x, x)$$

On appelera alors le membre de droite **forme polaire** de la forme quadratique q. On remarquera par la suite que moralement ces deux applications s'interpètent de la manière suivante:

- Une forme bilinéaire symétrique mesure des "angles" dans l'espace.
- Une forme quadratique mesure des "longeurs" dans l'espace.

Mais dans le cas général d'un espace quadratique et sans plus d'hypothèses, ces "angles" et "longueurs" ne correspondent pas vraiment aux concepts géométriques que l'on connaît. Tout l'algèbre bilinéaire développée dans ce chapitre peut se résumer à étudier des formes primitives qui par la suite permettront d'axiomatiser la notion de **produit scalaire** qui incarnera les propriétés géométriques recherchées.

Formules de polarisation:

On considère une forme quadratique q quelconque et on souhaiterait reconstituer la forme polaire de q, alors on peut montrer qu'elle vérifie les **identités de polarisation** ci-dessous:

$$\begin{cases} f(x, y) = \frac{1}{4}(q(x + y) - q(x - y)) \\ f(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)) \end{cases}$$

C'est identités se retrouvent aisément en considérant la forme quadratique triviale sur $\mathbb{R}$ associée à $f(x, y) = xy$ qui est donc $q(x) = x^2$.

Expression matricielle:

En dimension finie, on peut représenter les vecteurs $x, y \in E$ dans une base $\mathscr{B} = (e_i)$, on peut alors développer par bilinéarité et symétrie pour obtenir:

$$\boxed{f(x, y) = \sum_{1 \leq i,j \leq n} x_i y_j f(e_i, e_j)}$$

En particulier on peut alors remarquer que la forme bilinéaire est parfaitement déterminée par la donnée des n^2 images des paires des vecteurs de la base. En particulier, en notant X, Y les vecteurs colonnes des coordonées de x, y dans la base, alors on peut montrer qu'il existe une matrice A telle que:

$$\boxed{[f(x, y)]^{\mathscr{B}} = X^\top A Y}$$

Et cette matrice est alors de la forme suivante:

$$(a_{i,j}) = f(e_i, e_j)$$

En particulier cette matrice est symétrique et représente parfaitement la forme bilinéaire symétrique et la forme quadratique[1].

[1] Il suffit de calculer les coordonées de $f(x, x)$ pour le voir, on trouve alors qu'elle sont égales à $X^\top AX$

Règle du dédoublement:

Supposont que l'on connaisse une expression analytique de $q(x)$ dans une base $(e_i, \dots, e_n)$, on a alors:

$$q(x) = \sum_{1 \leq i,j \leq n} x_i x_j f(x_i, x_j) = \sum_{1 \leq i,j \leq n} x_i x_j b_{i,j}$$

Et donc par symétrique du produit, on peut alors retrouver la matrice $A = (a_{i,j})$ de q par la règle dite du **dédoublement des termes**, ie:

$$\begin{cases} a_{i,j} & = b_{i,j} \text{ si } i = j \\ a_{i,j} & = \frac{1}{2} b_{i,j} \text{ si } i \neq j \end{cases}$$

Exemple: On considère la forme quadratique suivante sur $\mathbb{R}^2$:

$$q(x, y) = 3x^2 + 5y^2 + 8xy$$

Alors la règle du dédoublement des termes nous donne que la matrice de q dans la base canonique est:

$$M = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}$$

Orthogonalité:

On peut alors définir une notion d'orthogonalité[1] pour la forme bilinéaire f, qu'on appelera f-orthogonalité, et on dira alors:

Deux vecteurs sont f-orthogonaux si et seulement si $f(x, y) = 0$.

On notera alors $x \perp y$. On peut alors définir le concept de famille f-orthogonale, ainsi que le concept de base f-orthogonale, comme une famille telle que tout les vecteurs soient deux à deux orthogonaux.

Attention: Dans le cas général, une famille orthogonale n'est pas forcément libre ! Il suffit de considérer $q(x, y) = x^2$ et la famille $(0, 1), (0, 1)$. C'est une nouvelle conséquence du fait que cette notion d'orthogonalité n'est **pas géométrique**.

Orthogonal d'une partie:

Soit A une partie de E, alors on peut définir l'orthogonal de A comme **le sous-espace vectoriel** défini par:

$$A^\perp := \left\{ u \in E \ ; \ \forall v \in A \, , \, u \perp v \right\}$$

En dimension finie, on a alors une base $(e_1, \dots, e_n)$ de F et on a la caractérisation suivante:

$$F^\perp := \left\{ u \in E \ ; \ \forall i \in [\![1 \ ; \ n]\!] \, , \, u \perp e_i \right\}$$

On peut alors montrer les propriétés suivantes pour l'application qui a une partie associe son orthogonal:

- La décroissance du passage à l'orthogonal.
- L'inclusion de la partie dans son double orthogonal.

Noyau:

Pour une forme bilinéaire quelconque, il est possible que l'ensemble $E^\perp$ soit non-trivial, on appelle alors cet ensemble **noyau** de la forme bilinéaire, cette dénomination[2] venant de la raison ci-dessous, pour M la matrice de la forme bilinéaire:

$$E^\perp = \ker(M)$$

[1] **Attention:** Pour l'instant ceci n'est **pas** une notion géométrique, mais purement algébrique.

[2] **Attention:** Ici on montre que le noyau de la forme bilinéaire est défini par le noyau de l'application linéaire associée à sa matrice, c'est non-trivial.

Avec ces définitions il est alors possible de montrer un analogue à la **formule du rang**:

$$\dim(\mathrm{Ker}q) + \dim(\mathrm{Im}q) = \dim(E)$$

Si ce noyau est trivial, on dira alors que la forme bilinéaire est **non-dégénérée** et en **dimension finie** on a alors les égalités suivantes:

$$\begin{cases} \dim(F^{\perp}) + \dim(F) = \dim(E) \\ F = F^{\perp^{\perp}} \end{cases}$$

Attention, ici il est important de noter que même pour une forme non-dégénérée, on n'a **pas la supplémentarité à priori**.

Cône isotrope:

Pour une forme bilinéaire quelconque, il est aussi possible que l'ensemble $\{x \in E; f(x,x) = q(x) = 0\}$ soit non-trivial, ici il s'agit des vecteurs dont la "longueur" est nulle, on appelera alors ces vecteurs **vecteurs isotropes**. On a alors directement l'inclusion suivant:

Le noyau est inclu dans le cône isotrope.

Si le cône isotrope est trivial, alors on dira que la forme est **définie**.

Cas particulier des formes réelles:

Soit f une forme bilinéaire symétrique réelle, alors on classifie ces formes par:

- Si $\forall x \in E, f(x,x) \geq 0$, on dira que la forme est **positive**.
- Si $\forall x \in E, f(x,x) \leq 0$, on dira que la forme est **négative**.

Réduction:

On cherche maintenant à trouver une base $\mathscr{B}$ de E telle que l'expression de $f(x,y)$ soit plus simple, en particulier on essaye de trouver une base telle que sa matrice soit **diagonale**, et on peut alors facilement montrer que c'est équivalent à **chercher une base f-orthogonale**. On a alors dans une telle base:

$$q(x) = \sum_{i \in I} q(e_i) x_i^2 = \sum_{i \in I} q(e_i) e_i^*(x)^2$$

On peut montrer le théorème suivant:

Il existe des formes linéaires indépendantes telles que $q(x) = q(e_1)l_1(x)^2 + \ldots + q(e_n)l_n(x)^2$

En particulier, ce théorème se démontre de manière constructive via un algorithme qui nous permettra de réduire tout forme bilinéaire en somme de carrés de formes linéaires indépendantes, c'est la **réduction de Gauss**, une fois les formes linéaires explicitées, il faut alors compléter la famille de formes linéaires en une base du dual, et la base orthogonale recherchée sera **la préduale de cette base**.

Algorithme de Gauss:

On se donne une forme quadratique suivante à décomposer:

$$q(x) = \sum_{i,j} c_{i,j} x_i x_j$$

L'algorithme est récursif et comporte deux cas:

- La forme quadratique comporte un terme carré de la forme $c_{i,i}x_i^2$
- La forme quadratique ne comporte pas de termes carrés de la forme $c_{i,i}x_i^2$

Traitons ces deux cas séparément:

- Dans le premier cas, on regroupe tout les termes qui comportent le terme carré et on applique la formule $A^2 + BA = (A + \frac{B}{2})^2 - \frac{B}{2}^2$ ce qui fait apparaître un carré de forme linéaire.
- Dans le second cas on regroupe tout les termes qui contiennent deux variables choisies et la formule $AB = \frac{1}{4}((A+B)^2 - (A-B)^2)$ ce qui fait apparaître des carrés de formes linéaires.

Exemple: On pose la forme quadratique suivante:

$$q(x, y, z, t) = x^2 + 2xy + y^2 - 4yz$$

On commence par isoler les termes en x et appliquer la formule du premier cas, on a donc notre premier carré de forme linéaire:

$$q(x, y, z, t) = (x^2 + 2xy) + y^2 - 4yz = (x + y)^2 - y^2 + y^2 - 4yz = l_1(x, y, z, t)^2 - 4yz$$

Maintenant on réapplique l'algorithme à la forme quadratique restante $q(x, \tilde{y}, z) = -4yz$, on applique la formule du second cas et on a:

$$q(x, \tilde{y}, z, t) = -4yz = -((y + z)^2 - (y - z)^2) = -l_2(x, y, z, t)^2 + l_3(x, y, z, t)^2$$

Finalement on trouve:

$$q(x, y, z, t) = l_1(x, y, z, t)^2 - l_2(x, y, z, t)^2 + l_3(x, y, z, t)^2 = (x + y)^2 - (y + z)^2 + (y - z)^2$$

On a alors trouvé les formes linéaires recherchées l_1, l_2, l_3 et pour trouver une base f-orthogonale de $\mathbb{R}^4$, il reste encore la dernière étape:

On complète (l_1, l_2, l_3) en une base du dual de E et la base recherchée est alors la préduale de celle-ci.

Signature:

On définir alors la **signature d'une forme bilinéaire** comme étant le couple d'entiers (p, q) avec:

- L'entier p est **le nombre de valeurs propres strictement positives**.
- L'entier q est **le nombre de valeurs propres strictement négatives**.

On peut alors montrer **la loi d'inertie de Sylvester** et si on a la forme réduite:

$$q(x) = \alpha_1 l_1(x)^2 + \ldots + \alpha_n l_n(x)^2$$

Alors on cette loi nous donne que p est exactement le nombre de coefficients α_i strictement positifs, et q le nombre de coefficients négatifs.

IV — Espaces Préhilbertiens Réels

On appelle **espace préhilbertien réel** un $\mathbb{R}$-espace vectoriel muni d'une forme **bilinéaire symétrique définie et positive**, c'est à dire une forme qui vérifie:

Symétrie	$f(u,v) = f(v,u)$
Définie	$f(u,u) = 0 \implies u = 0$
Positivité	$f(u,u) \geq 0$

On dira alors qu'une telle forme est un **produit scalaire** sur E et on notera:

$$f(x,y) = \langle x \,|\, y \rangle$$

Cet espace, qui est un cas particulier d'espace quadratique, est celui où se réalisera la signification **géométrique** des formes bilinéaires et quadratiques, grâce aux nouvelles contraintes sur ces formes.

Exemples :

- On définit sur $\mathbb{R}^n$ le produit scalaire défini par:

$$\langle u \,|\, v \rangle = \sum_{k=1}^{n} u_k v_k$$

- On définit sur $\mathcal{C}^1([0\,;\,1], \mathbb{R})$ le produit scalaire défini par:

$$\langle f \,|\, g \rangle = \int_0^1 f(t)g(t)dt$$

- On définit sur $\mathbb{R}_n[X]$ le produit scalaire défini pour (x_n) $n+1$ points fixés par:

$$\langle P \,|\, Q \rangle = \sum_{k=0}^{n} P(x_k)Q(x_k)$$

- On définit sur $\mathcal{M}_n(\mathbb{R})$ le produit scalaire défini par:

$$\langle A \,|\, B \rangle = \mathrm{tr}(A^\top B)$$

Norme :

A partir de la définition d'un tel espace, on alors montrer que $\langle \cdot \,|\, \cdot \rangle$ **induit une norme** sur E donnée par la **forme quadratique associée**:

$$\forall u \in E \;;\; \|u\|^2 = \langle u \,|\, u \rangle$$

Cela fait donc de E un espace vectoriel normé.

Angle:

A partir de ces définitions, on peut alors définir **l'angle non-orienté** $\theta \in [0\,;\,\pi]$ entre deux vecteurs u, v par:

$$\theta := \cos^{-1}\left(\frac{\langle u \,|\, v \rangle}{\|u\|\,\|v\|}\right)$$

Ce qui nous permet de caractériser **l'orthogonalité** du produit scalaire comme un orthogonal **géométrique**, en effet $\theta \equiv \frac{\pi}{2}$ dans cette définition ssi $\langle u \,|\, v \rangle = 0$

L'interprétation de cet "angle" ou de "l'orthogonalité" entre deux vecteurs diffère selon le contexte, elle peut alors signifier une corrélation en probabilité, ou un réel angle géométrique dans $\mathbb{R}^n$ par exemple.

Inégalité de Cauchy-Schwarz:

Dans tout espace préhilbertien réel, pour tout $u, v \in E$ on a l'inégalité[1] suivante:

$$\boxed{|\langle u \,|\, v \rangle| \leq \|u\| \, \|v\|}$$

Avec cas d'égalité quand u, v sont **liés**.

Formules géométriques:

Dans tout espace préhilbertien réel, pour tout $u, v \in E$ on a les identités suivantes:

- **Identité du parallélogramme :** $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$.
- **Théorème de Pythagore :** $x \perp y \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2$.

La première caractérise les espaces normés telles que leur norme soit issue d'un produit scalaire.

Orthogonalité:

Dans le cadre des espaces préhilbertiens, l'orthogonal obtient alors une partie des propriétés géométriques intuitives qu'on lui connait, en particulier:

- Une famille orthogonale de vecteurs non-nuls est toujours libre.
- Une partie et son orthogonal sont toujours en somme directe.

Attention dans un espace préhilbertien quelconque, ils ne sont pas toujours supplémentaires, on verra que c'est le cas en dimension finie !

Théorème de représentation:

On se donne un élément ϕ du dual d'un espace préhilbertien alors, dans ce cadre, et même de manière générale dans celui des espaces de Hilbert, on peut montrer **le théorème de représentation de Riesz** qui caractérise une forme linéaire ϕ par le produit scalaire:

$$\exists w \in E \ , \ \forall x \in E \ ; \ \phi(x) = \langle w \,|\, x \rangle$$

Simplement, cela signifie que toute forme linéaire est **exactement représentée** par le produit scalaire pour un certain vecteur w, en particulier, on a donc une bijection entre E et son dual.

Exemple: On prends la forme linéaire $\phi(x, y, z) = 5x + 4y + 3z$ et on munit $\mathbb{R}^3$ de son produit scalaire canonique, alors pour tout u on a directement que:

$$\phi(u) = \langle (5, 4, 3) \,|\, u \rangle$$

Transposition:

On se donne $f \in \mathcal{L}(E, F)$ représentée par une matrice $M \in \mathcal{M}_{n,p}(\mathbb{R})$, alors on définit **l'application transposée** de f par:

$$\begin{aligned} f^\top : F^* &\longrightarrow E^* \\ \phi &\longmapsto \phi \circ f \end{aligned}$$

On vérifie alors que cette application est bien définie est on a alors la propriété suivante:

L'application transposée est représentée dans les bases correspondates par la matrice transposée.

Ce qui donne finalement une interprétation fonctionnelle de la matrice transposée. A FINIR, LA TRANSPOSEE EST EXACTEMENT LADJOINT, LIEN AVEC TOUT LE RESTE A FAIRE, UNIQUE APPLICATION QUI VERFIE <f(x), y> = <x, tf(y)>, DUALITE.

[1] Trés puissante et permet d'obtenir des majorations dans des cas trés variés, la preuve parte de l'étude du polynôme $P(t) = \|x + ty\|^2$

IV — Espaces Euclidiens

On appelle **espace euclidien** tout espace préhilbertien réel **de dimension finie**. Dans toute la suite on prendra $(e_1, \dots, e_n)$ une base de E.
En particulier, dans une base **orthonormée**, on a $\langle x | e_i \rangle = x_i$ et donc:

$$x = \sum_{k=1}^{n} \langle x | e_i \rangle e_i$$

Orthogonalité:

En dimension finie, on a finalement l'ensemble des propriétés géométriques de l'orthogonalité qui deviennent vraies, en effet on a::

$$F \oplus F^{\perp} = E$$

Tout sous-espace admet un unique supplémentaire orthogonal.

On peut alors en déduire qu'en dimension finie on a:

$$(F^{\perp})^{\perp} = F$$

Projection orthogonale:

L'existence d'une unique décomposition nous permet alors de définir **la projection** sur F de direction $F^{\perp}$ par:

$$\text{proj}_F : x = x_F + x_{F^{\perp}} \mapsto x_F$$

En particulier, on en déduit par l'unicité de la décomposition que $\text{proj}_F(x)$ est **l'unique vecteur** de F qui vérifie:

$$x - \text{proj}_F(x) \in F^{\perp}$$

Le projeté orthogonal a une signification géométrique importante, en effet on a:

$$\left\| x - \text{proj}_F(x) \right\| = \min_{y \in F}(\|x - y\|)$$

C'est le vecteur "le plus proche" de F au sens du produit scalaire utilisé.

Calcul de projeté orthogonal:

On considère un sous-espace F de bases $(e_1, \dots, e_p)$ et $x \in E$, on sait d'aprés les propriétés précédentes que $\text{proj}_F(x)$ est l'unique vecteur de F tel que $x - \text{proj}_F(x) \in F^\perp$, ce qui est équivalent à dire que:

$$\forall j \in [\![1\ ;\ p]\!]\ ;\ \langle x - \text{proj}_F(x) \,|\, e_j \rangle = 0 \iff$$
$$\forall j \in [\![1\ ;\ p]\!]\ ;\ \langle \text{proj}_F(x) \,|\, e_j \rangle = \langle x \,|\, e_j \rangle$$

On raisonne alors par coefficient indeterminés avec l'écriture de $\text{proj}_F(x) = \sum_{i=1}^{n} \alpha_i e_i$ dans la base de F pour obtenir l'expression suivante:

$$\forall j \in [\![1\ ;\ p]\!]\ ;\ \sum_{i=1}^{n} \alpha_k \langle e_i \,|\, e_j \rangle = \langle x \,|\, e_i \rangle$$

Enfin on obtient alors **le système des équations normales**:

$$\begin{cases} \alpha_1 \langle e_1 \,|\, e_1 \rangle + \alpha_2 \langle e_2 \,|\, e_1 \rangle + \ldots + \alpha_p \langle e_p \,|\, e_1 \rangle = \langle x \,|\, e_1 \rangle \\ \vdots \\ \alpha_1 \langle e_1 \,|\, e_p \rangle + \alpha_2 \langle e_2 \,|\, e_p \rangle + \ldots + \alpha_p \langle e_p \,|\, e_p \rangle = \langle x \,|\, e_p \rangle \end{cases}$$

Ou de manière équivalente pour M la matrice du produit scalaire dans la base de $(e_1, \dots, e_p)$:

$$MY = \begin{bmatrix} \langle x \,|\, e_1 \rangle \\ \vdots \\ \langle x \,|\, e_p \rangle \end{bmatrix}$$

Dans le cas d'une base **orthogonale**, le système ci-dessus est beaucoup plus simple, en effet presque tout les produits scalaires sont nuls, et on obtient un système **diagonal** et donc dans ce cas précis, le projeté orthogonal s'obtient simplement par la formule:

$$\boxed{\text{proj}_F(x) = \sum_{k=1}^{p} \frac{\langle x \,|\, e_i \rangle}{\langle e_i \,|\, e_i \rangle} e_i}$$

Finalement, une remarque importante permet de comprendre la projection sur un sous espace doté d'une base orthogonale, en effet:

Projeter un vecteur sur un sous-espace revient à ajouter les projetés de ce vecteur sur les vecteurs de la base du sous-espace.

Exemple: Le projeté de $x = (1, 2, 3)$ sur le plan $\text{Vect}((1, 0, 0), (0, 1, 0))$ est donné par $\text{proj}_{(1,0,0)}(x) + \text{proj}_{(0,1,0)}(x)$

Procédé de Gramm-Schmidt:

Soit $(e_1, \dots, e_n)$ une base de E, on cherche alors à élaborer un procédé permettant **d'orthogonaliser cette base** en une base $(\varepsilon_1, \dots, \varepsilon_n)$, on pose $\varepsilon_1 = e_1$ et $H_i = \text{Vect}(e_1, \dots, e_i)$ et on définit par récurrence:

$$\boxed{\varepsilon_i = e_i - \text{proj}_{H_{i-1}}(e_i) = e_i - \left(\sum_{k=0}^{i-1} \text{proj}_{\varepsilon_k}(e_k) \right) = e_i - \left(\sum_{k=0}^{i-1} \frac{\langle e_i \,|\, \varepsilon_i \rangle}{\langle \varepsilon_i \,|\, \varepsilon_i \rangle} \varepsilon_i \right)}$$

Moralement, on "redresse" chaque vecteur de la base initiale en lui retirant son défaut d'orthogonalité représenté par sa projection sur le sous-espace précédent.

IV — Espaces Hemitiens

On peut généraliser la notion de produit scalaire au cas des espaces vectoriels sur $\mathbb{C}$, en particulier, on dira demandera alors que la forme $f : H \times H \to \mathbb{C}$ soit:

Linéaire à gauche	$f(x+\lambda y, z) = f(x,z) + \lambda f(y,z)$
Symétrie Hermitienne	$f(u,v) = \overline{f(v,u)}$
Définie	$f(u,u) = 0 \implies u = 0$
Positivité	$f(u,u) \geq 0$

On dira alors que f est un **produit hermitienne**. Elle est alors dite **sesquilinéaire** car on a:

$$f(x, \lambda y) = \overline{\lambda} f(x,y)$$

On définit aussi pour toute matrice dans $\mathcal{M}_n(\mathbb{C})$, sa **matrice ajointe** donnée par:

$$M^* = {}^t\overline{M}$$

Exemples :

- On définit sur $\mathbb{C}^n$ le produit hermitien défini par:

$$\langle u \,|\, v \rangle = \sum_{k=1}^{n} u_k \overline{v_k}$$

- On définit sur $\mathcal{C}^1([0\,;1], \mathbb{C})$ le produit hermitien défini par:

$$\langle f \,|\, g \rangle = \int_0^1 f(t)\overline{g(t)}dt$$

- On définit sur $\mathbb{C}_n[X]$ le produit hermitien défini pour (x_n) $n+1$ points fixés par:

$$\langle P \,|\, Q \rangle = \sum_{k=0}^{n} P(x_k)\overline{Q(x_k)}$$

- On définit sur $\mathcal{M}_n(\mathbb{C})$ le produit hermitien défini par:

$$\langle A \,|\, B \rangle = \mathrm{tr}(AB^*)$$

Expression matricielle:

En dimension finie, on peut représenter les vecteurs $x, y \in H$ dans une base $\mathscr{B} = (e_i)$, on peut alors développer par sesquilinéarité:

$$f(x,y) = \sum_{1 \leq i,j \leq n} x_i \overline{y_j} f(e_i, e_j)$$

En particulier on peut alors remarquer que la forme bilinéaire est parfaitement déterminée par la donnée des n^2 images des paires des vecteurs de la base. En particulier, en notant X, Y les vecteurs colonnes des coordonées de x, y dans la base, alors on peut montrer qu'il existe une matrice A telle que:

$$[f(x,y)]^{\mathscr{B}} = X^* A Y$$

Et cette matrice est alors de la forme suivante:

$$(a_{i,j}) = f(e_i, e_j)$$

En particulier cette matrice est égale à son adjointe et représente parfaitement la forme bilinéaire symétrique et la forme quadratique[1].

[1] Il suffit de calculer les coordonées de $f(x,x)$ pour le voir, on trouve alors qu'elle sont égales à X^*AX

Orthogonalité:

On peut alors définir la même notion d'orthogonalité et montrer que pour un produit hermitien, toute les propriétés de l'orthogonalité sont conservées sauf une, en effet le **théorème de Pythagore** n'est plus vrai dans un espace hermitien et on a seulement:

$$x \perp y \implies \|x+y\|^2 = \|x\|^2 + \|y\|^2$$

Théorème de représentation:

On se donne un élément ϕ du dual d'un espace hermitien alors, dans ce cadre, on peut aussi montrer **le théorème de représentation de Riesz** qui caractérise une forme linéaire ϕ par le produit scalaire:

$$\exists w \in H \ , \ \forall x \in H \ ; \ \phi(x) = \langle x \,|\, w \rangle$$

A nouveau, cela signifie que toute forme linéaire est **exactement représentée** par le produit scalaire pour un certain vecteur w, en particulier, on a donc une bijection entre H et son dual.

IV — Endomorphismes Remarquables

Aprés avoir défini les espaces euclidiens et hermitiens, on cherche maintenant à s'intéresser aux endomorphismes qui on des propriétés intéressantes en regard du produit scalaire, on sera alors amené à les définir et les étudier. Dans tout la suite, le corps de base peut être $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ et l'espace est muni d'un produit scalaire correspondant.

Isométries:

Soit $f \in \mathcal{L}(E)$, on dira que f est une **isométrie** et on note $f \in O(E)$ si elle **préserve les angles**, ie si:

$$\forall x, y \in E \; ; \; \langle f(x) \,|\, f(y) \rangle = \langle x \,|\, y \rangle$$

On en déduit directement qu'elle **préserve aussi les longeurs**[1].

On peut alors facilement montrer que la composée de deux isométries est une isométrie et que la réciproque l'est aussi. Aussi on peut déduire de la définition les propriétés suivantes:

$$\mathrm{Sp}(f) \subseteq \{-1, 1\}$$

Ainsi que comme corollaire immédiat:

$$\det(f) \in \{-1, 1\}$$

Matrices orthogonales:

On considère la matrice d'une isométrie dans une base orthonormée, on a alors d'aprés l'expression matricielle du produit scalaire:

$$\forall X, Y \in \mathcal{M}_{n,1}(\mathbb{K}) \; ; \; {}^*(MX)MY = {}^*X{}^*MMY = {}^*XY$$

On remarque donc f est une isométrie si et seulement si sa matrice dans une base orthonormée vérifie ${}^*M = M^{-1}$, on appelera de telles matrices **matrices unitaires** et on notera ces matrices $\mathbb{U}_n(\mathbb{K})$. On peut alors montrer que:

Les matrices unitaires forment un sous-groupe des matrices inversibles qu'on appelle groupe unitaire.

En particulier, la matrice de passage entre deux bases orthonormées est une matrice unitaire, et les colonnes d'une telle matrice sont de norme 1 et deux à deux orthogonales.

Classification des isométries:

On peut alors classifier les isométries selon leur action sur l'orientation de l'espace:

- Si $\det f = 1$ on dira que l'isométrie est directe, et on note l'ensemble de ces isométries $\mathrm{SO}(E)$, appelé **groupe spécial orthogonal**.
- Si $\det f = -1$ on dira que l'isométrie est **indirecte**, mais leur ensemble ne possède pas de structure particulière.

Géométriquement, les isométries directes sont donc celles qui préservent l'orientation de l'espace, c'est un sous-groupe du groupe spécial linéaire. Dans le chapitre suivant, on classifie plus précisément les isométries dans le cas d'une petite dimension.

[1] En particulier, elle est bijective et l'image d'une base orthonormée par une isométrie est toujours orthonormée.

Adjoint d'un endomoprhisme:

On considère un endomorphisme $f \in \mathcal{L}(E)$, alors le théorème de représentation de Riesz nous permet d'affirmer qu'il existe un unique endomorphisme f^* qui vérifie:

$$\forall x, y \in E \; ; \; \langle x \,|\, f(y) \rangle = \langle f^*(x) \,|\, y \rangle$$

On appelle alors cet endomorphisme **l'adjoint** de f, en particulier si on représente f par une matrice M dans une base orthonormée, alors on défini de meme la **matrice adjointe** de M par:

$$\forall X, Y \in \mathcal{M}_{n,1}(\mathbb{R}) \; ; \; {}^tXMY = {}^t(M^*X)Y$$

On peut alors facilement montrer que dans le cas présent de matrices, l'adjoint d'un endormorphisme est simplement sa **transposée**. On verra plus tard que la notion de matrice adjointe est une généralisation de la transposition. On peut alors entrevoir le role spécial que vont jouer les matrices symétriques. A FINIR, LA TRANSPOSEE EST EXACTEMENT LADJOINT, LIEN AVEC TOUT LE RESTE A FAIRE, UNIQUE APPLICATION QUI VERFIE <f(x), y> = <x, tf(y)>, VOIR DUALITE.

Endomorphismes auto-adjoints:

On appelle **endomorphisme auto-adjoint** (ou encore endomorphisme symétrique) tout endomorphisme qui est égal à son adjoint et on a donc les propriétés suivantes, cas particuliers des définitions ci-dessus:

$$\forall x, y \in E \; ; \; \langle x \,|\, f(y) \rangle = \langle f(x) \,|\, y \rangle$$

Puis matriciellement dans une base ortonormée:

$$\forall X, Y \in \mathcal{M}_{n,1}(\mathbb{K}) \; ; \; {}^*XMY = {}^*(MX)Y$$

En particulier dans le cas réel, le fait d'etre auto-adjoint est caractérisé par la propriété simple suivante sur la matrice de l'endomorphisme dans une base orthonormée:

La matrice de l'endomorphisme est symétrique.

L'ensemble des endomorphismes auto-adjoints, qu'on note $\mathcal{S}(E)$ forme un **sous-espace vectoriel** de $\mathcal{L}(E)$. Dans la suite on va étudier les propriétés de ces endomorphismes.

Théorème Spectral:

Une propriété fondamentale de ces endomorphismes est la suivante:

Leurs sous-espaces propres sont orthogonaux.

On peut alors énoncer un théorème puissant de réduction pour les endomorphismes auto-adjoints, en effet soit f un tel endomorphisme, alors:

Il existe une base orthonormée formée de vecteurs propres.

On a alors le corollaire matriciel pour la matrice M de f dans une base, donné par l'existance d'une matrice de passage P dans le groupe unitaire et d'une matrice diagonale D telles que:

$$M = PDP^*$$

Technique de réduction:

En particulier, cela nous donne une nouvelle méthode pour réduire les formes bilinéaires, on peut alors diagonaliser dans une base orthonormée et obtenir une base orthogonale pour la forme, en pratique, on effectue l'algorithme suivant:

- On trouve une base pour un sous-espace propre.
- On l'orthogonalise par Gram-Schmidt.

Lien avec les formes bilinéaires:

On sait donc que les endomorphismes autoadjoints sont représentés par des matrices symétriques, on a donc la propriété suivante:

On peut associer **une forme bilinéaire** symétrique à chaque[1] endomorphisme autoadjoint et **réciproquement**.

En particulier, les concepts relevant de l'études des formes peuvent alors se transposer dans l'étude des endomorphismes comme les sections suivantes le démontreront.

Endomorphismes auto-adjoints positifs:

On définit alors les endomorphismes autoadjoints **positifs** qu'on note $\mathcal{S}^+(E)$ définis par:

$$\mathcal{S}^+(E) := \{f \in \mathcal{S}(E) \ ; \ \mathrm{Sp}(f) \subseteq \mathbb{R}_+\}$$

Cette définition est alors équivalente à la propriété suivante:

$$\forall x \in E \ \left\langle x \,|\, f(x) \right\rangle \geq 0$$

Les matrices symetriques positives définissent alors des formes bilinéaires symétriques positives.

Endomorphismes auto-adjoints définis positifs:

On définit alors enfin les endomorphismes autoadjoints **définis positifs** qu'on note $\mathcal{S}^{++}(E)$ définis par:

$$\mathcal{S}^{++}(E) := \{f \in \mathcal{S}(E) \ ; \ \mathrm{Sp}(f) \subseteq \mathbb{R}_+^*\}$$

Cette définition est alors équivalente à la propriété suivante:

$$\forall x \in E\backslash\{0_E\} \ \left\langle x \,|\, f(x) \right\rangle > 0$$

Les matrices symetriques définies positives définissent alors des nouveaux produits scalaires.

[1] En effet soit $f \in S(E)$, alors $\phi(x,y) = \left\langle x \,|\, f(y) \right\rangle$ est une telle forme, et réciproquement Riesz nous donne que $\phi(x,y) = \left\langle x \,|\, f(y) \right\rangle$ pour une certaine fonction f qui est alors un endormorphisme autoadjoint.

IV — Isométries en petite dimension

On va maintenant étudier le cas particulier des isométries dans le cas de la petite dimension, c'est à dire dans le cas où E est de dimension 2 ou 3, on introduira un outil pratique dans ce contexte qui est le **produit vectoriel** et on classifiera les isométries dans cet espace.

Bases directes:

Pour ce chapitre nous auront besoin du concept **d'orientation** de l'espace défini dans le chapitre sur les déterminants, en effet on appelera **base orthonormée directe** toute base ayant même orientation que la base canonique des espaces considérés. Sinon on dira que la base est **indirecte**.

Cas de la dimension 2:

Soit θ un réel, on définit les deux matrices orthogonales suivantes:

$$R_\theta := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \qquad S_\theta := \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$$

On appele alors R_θ **matrice de rotation** d'angle θ et S_θ est un **symétrie axiale**[1]. Alors on a le théorème fondamental suivant, pour $f \in O(E)$ et tout base orthonormée $\mathscr{B}$, alors il existe θ réel tel que:

$$[f]_{\mathscr{B}} \in \{R_\theta, S_\theta\}$$

En particulier si f préserve l'orientation, alors nécéssairement, f est une **rotation**. Et donc les matrices de passages entre bases orthonormées directes sont des rotations.
En particulier si f ne préserve pas l'orientation, alors nécéssairement, f est une **symétrie axiale**.

Cas de la dimension 3:

Dans le cas de la dimension trois, on pose $\varepsilon = \pm 1$ et on définit la matrice suivante:

$$M_\theta := \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}$$

Alors on a le théorème fondamental suivant, pour $f \in O(E)$ et tout base orthonormée **directe** $\mathscr{B}$, alors il existe θ réel tel que:

$$[f]_{\mathscr{B}} = M_\theta$$

La classification des isométries dans ce cas est alors plus complexes et repose sur l'étude de la dimension de l'espace des point fixes, qu'on notera F, et on peut alors les classifier selon le tableau[2] ci-dessous:

$\dim(F)$	Orientation	Matrice	Nature
3	Directe	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Identité
2	Indirecte	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$	Symétrie par rapport à F
1	Directe	$\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Rotation d'axe F
0	Indirecte	$\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & -1 \end{pmatrix}$	Composée d'une symétrie et d'une rotation d'axe F

[1] D'axe la bissectrice de l'angle θ, faire un dessin.
[2] Le dernier cas peut se ramener par des propriétés trigonométriques au cas d'un rotation, en effet $-f = R_{\theta+\pi,u}$

Produit mixte:

On considère une famille de vecteurs u, v, w et deux bases orthonormées directes $\mathscr{B}, \mathscr{B}'$ de E, alors on a:

$$\det([u]_{\mathscr{B}}, [v]_{\mathscr{B}}, [w]_{\mathscr{B}}) = \det([u]_{\mathscr{B}'}, [v]_{\mathscr{B}'}, [w]_{\mathscr{B}'})$$

On appelle alors le déterminant de cette famille de vecteur **le produit mixte** de ces trois vecteurs, et on le note:

$$\det([u]_{\mathscr{B}}, [v]_{\mathscr{B}}, [w]_{\mathscr{B}}) = [u, v, w]$$

C'est donc le volume orienté du parallélotope formé par les trois vecteurs.

Produit vectoriel:

On considère une famille de vecteurs u, v, et un vecteur x quelconque, alors d'aprés le théorème de représentation de Riesz, on a:

$$\exists w \in E \ ; \ [u, v, x] = \langle w \,|\, x \rangle$$

On appelle alors ce vecteur **produit vectoriel** de u et v et on le note $u \times v$. On peut alors à partir des propriétés du determinant, montrer que:

- Le produit vectoriel est une application **bilinéaire altérnée**.
- Le produit vectoriel $u \times v$ est **orthogonal** à u et v.
- Si (u, v) est une famille orthonormée, $(u, v, u \times v)$ est une base orthonormée directe.
- Si (u, v, w) est une base orthonormée directe $w = u \times v, u = v \times w$ et $v = w \times u$.

Le produit vectoriel est donc un moyen trés pratique de **construire des bases directes** ou de tester la colinéarité. Analytiquement, on peut le calculer en cordonnées par:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \det \begin{pmatrix} x_2 & y_2 \\ x_3 & y_3 \end{pmatrix} \\ -\det \begin{pmatrix} x_1 & y_1 \\ x_3 & y_3 \end{pmatrix} \\ \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \end{pmatrix}$$

Recherche des éléments caractéristiques:

Pour réussir à trouver le réel θ, on utilise alors le faire que dans une base adaptée, on a:

$$\mathrm{tr}(f) = 2\cos(\theta) + 1$$

Et donc à partir de la trace de la matrice dans une base quelconque, on peut retrouver le cosinus de l'angle θ dont il reste à déterminer le sinus pour le caractériser.

Arrive alors l'intérêt du produit mixte, en effet, on peut alors montrer que le sinus de l'angle θ est du même signe que la quantité ci-dessous, pour x un vecteur quelconque (souvent c_1) de notre choix:

$$[x, f(x), u]$$

Ce qui caractérise alors parfaitement l'isométrie.

XII — Espaces probabilisés

Le domaine des probabilités chercher à modéliser des **expériences aléatoires** ie des expériences dont toutes les **issues** possibles sont connues à priori mais dont le résultat peut varier lorsqu'on la répète (lancer de dés, tirage dans une urne...).

Le cadre de la théorie de la mesure, nous permet de formaliser la théorie axiomatique des probabilités, ainsi que les différents objets en jeu, en particulier, on considère un espace mesurable $(\Omega, \mathcal{A})$ muni d'une mesure $\mathbb{P}$ à valeurs dans $[0\,;\,1]$ et telle que $\mathbb{P}(\Omega) = 1$. On apelle une telle mesure **loi de probabilité**.

Le triplet $(\Omega, \mathcal{A}, \mathbb{P})$ est alors apellé **espace probabilisé**. Dans ce cadre l'ensemble Ω des issues possibles de l'expérience est apellé **univers**, les parties mesurables sont apellées **évenements** et deux parties mesurables disjointes seront dites **incompatibles**.

Espace probabilisé discret et continu

La mesure de l'espace doit être égale à 1 donc en particulier, on doit avoir $\int_\Omega d\mathbb{P} = 1$, ceci étant dit, on peut alors distinguer deux grands cas d'espaces probabilisés:

- Le cas où les parties non-négligeables de Ω sont **dénombrables**, alors par applications de la relation de Chasles et l'invisibilité des parties négligeables, on obtient que:

$$\int_\Omega d\mathbb{P} = \int_{\bigcup x_n} d\mathbb{P} = \sum_n \mathbb{P}(x_n)$$

  On remarque alors que la loi de probabilité est entièrement déterminée par la probabilité **d'évenements élémentaires** d'une certaine famille (x_n) dont la série vaut 1. On appele cette famille **distribution** de $\mathbb{P}$ et de tels espaces **espaces probabilisés discrets**.

- Le cas où elles ne le sont pas et que $\Omega \subseteq \mathbb{R}^n$, alors il est toujours possible de définir la **fonction de répartition** de la mesure de probabilité par la fonction suivante qui caractérise la loi:

$$F : x \longrightarrow \mathbb{P}(]-\infty\,;\,x])$$

  Si de plus les non-boréliens sont négligeables pour $\mathbb{P}$ (on dit aussi que $\mathbb{P}$ est **absolument continue** par rapport à la mesure de Lebesgue) alors on peut montrer que $\mathbb{P}$ admet une densité, c'est à dire une fonction intégrable f dont l'intégrale vaut 1 et qui caractérise alors la probabilité:

$$\mathbb{P}(A) = \int_A f(x)dx$$

  On dira alors que ces lois sont des **lois à densité**. Si la dernière condition n'est pas vérifiée, on dira alors que la loi est **mixte ou singulière**.

Espace probabilisé produit

Si on se donne une famille de n espaces probabilisés $(\Omega_i, \mathscr{A}_i, \mathbb{P}_i)$, on peut alors conformément à la théorie de la mesure définir l'espace produit $(\prod \Omega_i, \mathscr{A}_\otimes, \mathbb{P}_\otimes)$ avec la tribu et la loi produit. Dans tout la suite on considère simplement le cas $n = 2$ pour simplifier.
Selon le cas discret ou a densité, on a alors que la loi est caractérisée par:

- **Cas dénombrable:**

$$\mathbb{P}_\otimes(A) = \sum_{\mathbb{N}\times\mathbb{N}} \mathbb{P}_\otimes(x_n, y_m)$$

- **Cas absolument continu:**

$$\mathbb{P}_\otimes(A) = \int_{\Omega_1\times\Omega_2} f(x,y)d\mathbb{P}_\otimes$$

Lois marginales

Sachant la loi produit $\mathbb{P}$, une question intéressante est alors de determiner les lois **marginales** des espaces composantes, on montre alors qu'on a:

- **Dans le cas dénombrable:**
$$\mathbb{P}_1(\{k\}) = \sum_{\mathbb{N}} \mathbb{P}(\{k\}, y_m)$$
- **Dans le cas absolument continu:**
$$f_1(x) = \int_{\Omega_1} f(x,y)dy$$

Malheureusement on peut montrer que la donnée des lois marginales ne caractérise pas la loi produit, en effet les lois marginales dans le cas fini par exemple correspondent aux sommes des lignes ou colonnes du tableau des probabilités et deux sommes peuvent être égales sans que les valeurs individuelles soient toutes égales.

Exemples

Plusieurs exemples de différentes natures:

- **Discret fini:** Si on cherche à modéliser 3 tirages successifs à pile ou face avec une pièce non truquée, on peut modéliser cette expérience par l'espace probabilisé suivant:
$$\left(\left\{(r_1,r_2,r_3)\ ;\ (r_i) \in \{P,F\}\right\}, \mathcal{P}(\Omega), \mathbb{P}(A) = \frac{|A|}{|\Omega|}\right)$$
- **Discret infini:** Si on cherche à modéliser le nombres de visiteurs qui se présentent dans un musée, on peut alors modéliser ce phénomène par un espace probabilisé dénombrable et une probabilité rapidement décroissante, ie on pose par exemple:
$$\left(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathbb{P}(\{n\}) = \frac{1}{2^{n+1}}\right)$$
Alors ceci est bien un espace probabilisé discret infini.
- **Absolument continu:** Si on cherche à modéliser un tir de fléchette sur le disque unité $D \subseteq \mathbb{R}^2$où la probabilité suit une densité uniforme, alors on peut poser:
$$\left(D, \mathcal{B}(D), \mathbb{P}(D) = \frac{1}{\pi}\int_A d\mu = \frac{\mu(A)}{\pi}\right)$$
- **Mixte:** Si on cherche à modéliser une loterie où l'on tire un nombre de $[0\,;\,10]$ avec $\mathbb{P}(\{0\}) = 0.1$ qui correspond au jackpot et densité uniforme pour le reste des nombres. Alors on a naturellement une structure d'espace probabilisé mixte.
- **Espace produit:** Si on cherche à modéliser le choix uniforme d'un point dans $[\![1\ ;\ n]\!]^2$, on modèlise ceci par l'espace produit:
$$([\![1\ ;\ n]\!]^2, \mathcal{P}([\![1\ ;\ n]\!]^2), \mathbb{P}(\{(k,l)\}) = \frac{1}{n^2})$$
Alors les distributions marginales sont facilement $\mathbb{P}_1(\{k\}) = \mathbb{P}_2(\{k\}) = \frac{1}{n}$, on a en fait le tableau des probabilités décrit par la matrice de taille n suivante:
$$\begin{pmatrix} \frac{1}{n^2} & \cdots & \frac{1}{n^2} \\ \vdots & \ddots & \vdots \\ \frac{1}{n^2} & \cdots & \frac{1}{n^2} \end{pmatrix}$$
Les colonnes (resp. lignes) correspondant aux probabilités $\mathbb{P}_1(\{k\})$ (resp. $\mathbb{P}_2(\{k\})$)

XII — Probabilités conditionnelles

Lorsque l'on dispose d'informations sur le résultat d'une expérience donnée, il est possible d'affiner nos prédictions.
Soit X un évenement qui n'est pas négligeable, alors on définit l'application:

$$\begin{aligned}\mathbb{P}(\cdot|X) : \mathcal{P}(\Omega) &\longrightarrow [0\,;\,1] \\ A &\longmapsto \frac{\mathbb{P}(A \cap X)}{\mathbb{P}(X)}\end{aligned}$$

On peut montrer que c'est une mesure de probabilité sur Ω et on l'appelle **probabilité de A sachant X**. De la symétrie de l'intersection on peut alors en déduire la **formule de Bayes** qui permet alors **d'inverser le conditionnement**:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A)\mathbb{P}(B|A)}{\mathbb{P}(B)}$$

Formule des probabilités composées

On en déduit directement la **formule des probabilités composées**:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A) = \mathbb{P}(B)\mathbb{P}(A|B)$$

Qui se généralise pour une famille finie d'évenements $(A_n)_{n\in I}$ d'intersection non nulle:

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}_{A_1}(A_2)\mathbb{P}_{A_1\cap A_2}(A_3)\ldots\mathbb{P}_{A_1\cap\ldots\cap A_{n-1}}(A_n)$$

Formule des probabilités totales

On considère une partition $(A_n)_n$ de Ω en évenements disjoints (on appelle une telle partition **systême complet d'évenements**), alors on peut montrer la **formule des probabilités totales**:

$$\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}(A_k)\mathbb{P}_{A_k}(B)$$

Indépendance

On dit que deux évenements A, B sont **indépendants** si et seulement si la donnée de la réalisation d'un des évenements n'influence pas l'autre, ie:

$$\mathbb{P}(A|B) = \mathbb{P}(A)$$

Ou encore par la formule conditionelle:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Si deux évenements sont indépendants, alors n'importe quelle paire de $A, B, \overline{A}, \overline{B}$ est indépendante.

XII — Lois usuelles

Dans ce chapitre, on énumère les lois usuelles en probabilité et leurs cas d'utilisation. Comme vu précédement, on distingue le cas discret et absolument continu. Dans tout la suite on considère un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.

Lois discretes usuelles

On appelle **épreuve de Bernoulli** est une expérience aléatoire qui n'a que deux issues, usuellement nommées **succés et échec**.

On dit que la loi est **uniforme** si on a la distribution:

$$\forall A \in \mathscr{A} \ ; \ \mathbb{P}(A) = \frac{|A|}{|E|}$$

On dit que la loi est **binomiale** de paramêtres n, p si $\Omega = \{1, \dots, n\}$ et si on a la distribution:

$$\forall k \leq n \ ; \ \mathbb{P}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}$$

On dit que la loi est **géométrique** de paramêtre p si $\Omega = \mathbb{N}^*$ et si on a la distribution:

$$\forall k \geq 1 \ ; \ \mathbb{P}(\{k\}) = p(1-p)^{k-1}$$

On dit que la loi est **hypergéométrique** de paramêtres (p, n, N) si $\Omega = \mathbb{N}^*$ et si on a la distribution:

$$\forall k \geq 1 \ ; \ \mathbb{P}(\{k\}) = \frac{\binom{pN}{k}\binom{(1-p)N}{n-k}}{\binom{N}{n}}$$

On dit que la loi est **de Poisson** de paramêtres λ si $\Omega = \mathbb{N}^*$ et si on a la distribution:

$$\forall k \geq 1 \ ; \ \mathbb{P}(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda}$$

La **loi binomiale** est utilisée pour déterminer la probabilité d'obtenir exactement k succés aprés n itérations d'une épreuve de Bernoulli.
La **loi géométrique** est utilisée pour déterminer la probabilité d'un temps d'attente k avant le le premier succés d'une épreuve de Bernoulli.
La **loi hypergéométrique** est utilisée pour déterminer la probabilité d'obtenir k succés aprés n itérations d'une épreuve de tirage sans remise dans une urne contenant N boules, dont pN boules gagantes, et $(1-p)N$ boules perdantes, avec la contrainte que pN soit un entier.
La **loi de Poisson** est utilisée pour déterminer le nombre d'événements se produisant dans un intervalle de temps fixé, si ces événements se produisent avec une fréquence moyenne connue, et indépendamment du temps écoulé depuis l'événement précédent[1].

[1] C'est une loi qui s'obtient asymptotiquement à partir d'une loi binomiale de paramêtres $T, \frac{\lambda}{T}$ en faisant tendre T vers l'infini.

Lois à densité usuelles

On dit que la loi est **uniforme** si sa densité f est constante sur un intervalle $[a\,;\,b]$ et nulle en dehors.
On dit que la loi est **exponentielle** de paramètre λ si on a la densité:

$$f(x) = \lambda \exp(-\lambda x)\ ;\ x \geq 0$$

On dit que la loi est **normale** dee paramêtres[1] μ, σ si on a la densité:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

La **loi exponentielle** est utilisée pour modéliser le temps d'attente d'un phénomène sans mémoire, en particulier, c'est l'analogue continue de la loi géométrique[2].
La **loi normale** est fondamentale en probabilités du fait de son omniprésence dans les sciences expérimentales, en effet, un théorème fondamental montrera que la somme d'une suite de variables aléatoires (comprendre expériences) convergera vers une certaine loi normale. Elle est donc d'importance capitale en statistiques.

[1] Ces paramêtres correspondent alors à l'éspérance et l'écart type de la loi.
[2] Elle s'obtient asymptotiquement à partir d'une loi géométrique de paramètre λT en faisant tendre T vers 0.

XII — Variables aléatoires

Trés souvent, il se trouve que l'espace probabilisé de l'expérience est inconnu, trop grand ou trop complexe, on considérera alors simplement son existence et on étudiera celui ci via des fonctions définies sur cet espace, appelées **variables aléatoires**. Ces fonctions induiront un nouvel espace probabilisé correspondant à notre expérience précise (souvent un espace probabilisé numérique).

On considère alors souvent $(\Omega, \mathscr{A}, \mathbb{P})$ comme un espace probabilisé abstrait et on l'oublie même complétement trés souvent. Par exemple:

- On considère une expérience aléatoire qui tire au hasard un gateau dans une chaîne de fabrication, on peut alors définir une variable aléatoire sur l'espace probabilisé naturellement défini qui à chaque évenement associe le volume du gateau, son taux de sucre, le nombre de raisins secs ... Et faire alors des suppositions sur la loi de ces variables aléatoires par exemple on pourra supposer que le taux de sucre d'un gateau choisi au hasard suit une loi normale.
- Si on considère un groupe de N personnes vivant un épisode épidémique, alors il est trés compliqué de modéliser l'état épidémique du groupe à un instant donné du fait des différentes intéractions et dépendances, on préfère alors éudier des espaces probabilisés plus simples induits par des variables aléatoires comme le nombre de personnes infectées, le temps mis par l'épidémie pour atteindre une certaine taille etc ..

Définition

On dira que X est une **variable aléatoire** de $(\Omega_1, \mathscr{A}, \mathbb{P})$ vers un espace mesurable $(\Omega_2, \mathscr{B})$ si et seulement si c'est une **fonction mesurable** sur cet espace. Elle définit alors une loi naturelle sur Ω_2 définie par la **mesure image**:

$$\begin{aligned} \mathbb{P}_X : \mathscr{B} &\longrightarrow [0\,;\,1] \\ B &\longmapsto \mathbb{P}(X^{-1}(B)) \end{aligned}$$

On note alors plus simplement $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$. Trés souvent, on considérera $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ comme espace d'arrivée et donc la variable aléatoire sera dite **réelle** et définira une loi sur les boréliens.

Cas réel

Dans le cas de variables aléatoires **réelles** on peut aussi créer une notation qui s'applique si B est un intervalle et on a:

$$X^{-1}\big(]a\,;\,b[\big) = \Big\{\omega \in \Omega\,;\ a < X(\omega) < b\Big\} \overset{\text{notation}}{=} (a < X < b)$$

Propriétés

La famille $((X = a))_{a \in X(\Omega)}$ est un **système complet d'évenements**, en effet car si on considère une issue $\omega \in \Omega$, on a:

$$\begin{cases} X(\omega) = x \implies \omega \in (X = x) \\ X(\omega) \neq x \implies \omega \notin (X = x) \end{cases}$$

On peut donc partitionner les éléments de Ω selon leur image par X

On peut aussi noter que si f est une application mesurable, alors $f \circ X$ **est une variable aléatoire** sur les espaces correspondants.

Indépendance

On dira alors que deux variables aléatoires X, Y sont indépendantes si et seulement si pour tout couple x, y, les évenements correspondants sont indépendants:

$$\mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Par ailleurs, si X, Y sont deux variables aléatoires, il existe une mesure de la dépendance (corrélation) de deux variables aléatoires appellée **covariance** définie dans la dernière partie.

Cas des vecteurs aléatoires

Dans le cas où la variable aléatoire $X = (X_1, \ldots, X_n)$ est à valeurs dans $\mathbb{R}^n$, alors elle définit une loi produit (appelée dans ce cadre **loi conjointe** de X) sur les boréliens et on l'appelle **vecteur aléatoire**. En particulier les lois marginales sont alors les lois des variables composantes X_i. Par exemple si on prends un vecteur aléatoire choisissant uniformément un point dans $[\![1\ ;\ n]\!]^2$, alors on a que:

$$\mathbb{P}(X = (k, l)) = \frac{1}{n^2}$$

Et les loi marginales sont données par:

$$\mathbb{P}(X_1 = k) = \sum_{i=1}^{n} \mathbb{P}(X = (k, i))$$

On retrouve alors les même lois marginales que dans l'exemple analogue sans variable aléatoire, en particulier les lois conjointe et marginales sont exactement les lois produits et marginales sur $\mathbb{R} \times \mathbb{R}$.

Intégrabilité et formule de transfert

On se donne une variable aléatoire réelle intégrable par rapport à la mesure $\mathbb{P}$ ie telle que:

$$\int_\Omega X(\omega) d\mathbb{P} < \infty$$

Alors on peut montrer l'identité suivante par les propriétés de la mesure image $d\mathbb{P}_X$:

$$\int_\Omega X(\omega) d\mathbb{P} = \int_{\mathbb{R}} x d\mathbb{P}_X$$

Et même plus généralement, on a le **théorème de transfert** pour tout fonction ϕ telle qu'une des intégrale ait un sens:

$$\int_\Omega \phi(X(\omega)) d\mathbb{P} = \int_{\mathbb{R}} \phi(x) d\mathbb{P}_X$$

Et les intégrales du membre de droite se calculent souvent facilement, par exemple dans les deux cas classiques:

- **Cas dénombrable:** On a

$$\int_{\mathbb{R}} x d\mathbb{P}_X = \sum_{\mathbb{N}} \int_{y_n} x d\mathbb{P}_X = \sum_{\mathbb{N}} y_n \mathbb{P}(X = y_n)$$

- **Cas absolument continu:** On a

$$\int_{\mathbb{R}} x d\mathbb{P}_X = \int_{\mathbb{R}} x f(x) dx$$

XII — Indicateurs

Dans tout la suite, on considère $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace probabilisé fini et X, Y deux variables aléatoires **intégrables** pour la mesure $\mathbb{P}$.

On appelle **indicateur de position** un nombre réel permettant de situer les valeurs d'une série statistique, par exemple l'esperance et la médiane sont des indicateurs de position.
On appelle **indicateur de dispersion** un nombre réel permettant de mesurer la variabilité des valeurs d'une série statistique autour d'une valeur (généralement autour de la moyenne), par exemple la variance, l'écart-type ou l'écart interquartile sont des indicateurs de dispersion.

Esperance

L'espérance mathématique correspond à la moyenne théorique du résultat qu'on peut espérer avoir en répétant une expérience aléatoire un grand nombre de fois, c'est **la moyenne des valeurs de la variable aléatoire, pondérées par leur probabilités respectives**, ou c'est aussi le centre de masse de la densité, on définit alors celle ci par:

$$\mathbb{E}(X) := \int_\Omega X d\mathbb{P}$$

L'esperance prends alors la forme d'une somme pondérée dans le cas discret, ou d'une intégrale pondéree dans le cas absolument continue. Elle existe toujours dans le cas d'une variable aléatoire **finie** mais ce n'est pas le cas en général, et il faut alors étudier l'intégrabilité de la variable aléatoire.

L'espérance possède plusieurs propriétés remarquables, elle est **linéaire et croissante** et l'espérance d'une constante est cette constante.
Mais en général, l'espérance **n'est pas multiplicative**, c'est néanmoins le cas quand les deux variables aléatoires sont **indépendantes**.

Variables centrées

On rappelle qu'on a le théorème de transfert donc $\mathbb{E}(\phi(X))$ existe si $\phi(X)$ est intégrable. Aussi, on appelle **variable centrée** une variable aléatoire d'espérance nulle. On peut alors centrer une variable aléatoire par la translation $X' = X - \mathbb{E}(X)$.

Moments d'ordre k

On généralise cette définition et on définit le **moment d'ordre k** de la variable X, si il existe par la quantité suivante:

$$m_k(X) = \mathbb{E}(X^k)$$

On remarque alors que cette quantité existe si et seulement si X^k est intégrable. De manière générale, on comprends facilement qu'on a la propriété suivante (où la disjonction dépends de la discrétude ou non de X):

$$X \text{ admet un moment d'ordre } k \iff X \in L^k(\Omega) \text{ ou } \ell^k(\Omega)$$

Variance

On manque alors d'informations sur les valeurs de X, elles peuvent tout aussi bien rester toujours très proches de $\mathbb{E}(X)$, ou s'en éloigner beaucoup, on a donc besoin de mesurer la distance moyenne entre X et $\mathbb{E}(X)$, qui serait alors $\mathbb{E}(|X - \mathbb{E}(X)|)$.

Cette formule est techniquement impraticable du fait de la valeur absolue, on utile donc **la moyenne des carrés des distances** entre X et $\mathbb{E}(X)$, et on définit alors la variance:

$$\mathbb{V}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

La deuxième expression apellée **formule de Koenig-Huygens** se déduit facilement de la première par les propriétés de l'espérance et on en déduit qu'une variable aléatoire admet une variance si et seulement si elle admet un moment d'ordre 2.
On voit directement que la variance est **positive** (ou nulle si X est constante presque partout). Elle vérifie aussi les propriétés suivantes:

- **Invariante par translation** $\mathbb{V}(X + a) = \mathbb{V}(X)$
- **Quadratique** $\mathbb{V}(\lambda X) = \lambda^2 \mathbb{V}(X)$

Ecart type

On peut alors définir **l'écart-type** de la variable X qui est défini par $\sigma(X) = \sqrt{\mathbb{V}(X)}$

Enfin, on appelle **variable réduite** une variable aléatoire d'écart type 1 et on peut alors définir la **variable centrée réduite** associée à X:

$$X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}$$

Covariance

Dans le cas d'un couple aléatoire et contrairement à l'espérance, le concept de variance perds son sens. Plutôt que de chercher un écart par rapport à la moyenne, on va préférer chercher **un écart moyen entre les deux variables**[1] qui se définit par:

$$\mathrm{Cov}(X, Y) = \mathbb{E}\big[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\big] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Gràce à la covariance on peut aussi définir la variance d'une somme:

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\mathrm{Cov}(X, Y)$$

La covariance de X, Y n'existe bien sûr que si X, Y et XY sont intégrables.

Propriétés de la covariance

La covariance vérifie plusieurs propriétés intéressantes:

- Si les deux variables aléatoires sont indépendantes, alors on a $\mathrm{Cov}(X, Y) = 0$, la reciproque étant **fausse**.
- Si X admet une variance alors on retrouve celle ci comme $\mathrm{Cov}(X, X)$

On peut même montrer que la covariance est **bilinéaire, symétrique et positive**. Informellement c'est un *"pseudo produit scalaire"* sur les variables aléatoires, néanmoins suffisament proche du produit scalaire pour avoir **l'inégalité de Cauchy-Schwartz**:

$$|\mathrm{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$$

Plus précisément, considérons l'espace des variables aléatoires centrées réduites, alors c'est **un espace pre-hilbertien**, et la covariance définit son produit scalaire, l'écart type définit alors la **norme** associée et on peut définir le coefficient de corrélaition linéaire par:

$$\varrho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

Qui s'intérpréterait alors comme *l'angle* entre les variable aléatoires.

[1] On remarque que la variance n'est alors que la covariance de X avec elle-même. Aussi si les deux variables aléatoires sont indépendantes, dans ce cas l'espérance est multiplicative et on a $\mathrm{Cov}(X, Y) = 0$, la reciproque étant **fausse**.

Espérance & variances usuelles

Il est intéressant de considérer les différents indicateurs de position et de dispersion des lois usuelles[1]

Lois	Espérance	Variance
$X \sim \mathcal{U}(E)$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
$X \sim \mathcal{B}(n,p)$	$n \cdot p$	$n \cdot p(1-p)$
$X \sim \mathcal{G}(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$X \sim \mathcal{H}(p,n,N)$	$n \cdot p$	$n \cdot p(1-p)\frac{N-n}{N-1}$

Inégalités

On souhaite majorer la probabilité d'avoir des valeurs "extrèmes", ie éloignées de l'espérance, alors si X est une variable aléatoire positive presque partout et $\alpha \in \mathbb{R}^{+*}$ on peut montrer **l'inégalité de Markov**:

$$\mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}(X)}{\alpha}$$

Si la variable aléatoire admet une variance, on a alors **l'inégalité de Bienaymé-Tchebychev**:

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \alpha) \leq \frac{\mathbb{V}(X)}{\alpha^2}$$

Cette dernière est un cas particulier de la première mais est en général *plus fine* que la majoration donnée par l'inégalité de Markov.

[1] Pour la loi uniforme on considère X à valeurs dans $[\![1\ ;\ n]\!]$

XII — Convergences Stochastiques

On peut alors tenter d'appliquer les résultat sur les espaces L^p à la théorie des probabilité, en particulier étudier des intégrales de variables aléatoires, des espérances, etc .. revient à étudier la finitude d'un norme dans $L^p(\Omega)$ en particulier pour tout $p \in [1\,;\infty]$, on en déduis que les normes p s'appliquent aux variables aléatoires, ie on a:

$$\begin{cases} \|X\|_p := \left(\int_\Omega |X|^p d\mathbb{P}\right)^{\frac{1}{p}} = \left(\mathbb{E}(|X|^p)\right)^{\frac{1}{p}} \\ \|X\|_\infty := \sup \text{ess}\{|X|\} \end{cases}$$

Où ici X est supposée à densité ou à distribution $f(x)$.

Cette approche sera alors trés fructueuse, en effet par l'étude et la définition de différent *modes de convergences* bien choisis sur des suites de variables aléatoires (X_n), on pourra alors démontrer les grands théorèmes probabilistes.

Convergence en loi

On dira que X_n **converge en loi** vers X si et seulement si pour tout $x \in \mathbb{R}$, la loi de X_n est abitrairement proche de la loi de X, ie si on a:

$$\mathbb{P}(X_n \leq x) \longrightarrow \mathbb{P}(X \leq x)$$

On notera alors:

$$(X_n) \xrightarrow{\mathcal{L}} X$$

Convergence en probabilité

On dira que X_n **converge en probabilité** vers X si et seulement si pour tout $\varepsilon > 0$, la probabilité que $(X_n)_n$ s'éloigne de X tends vers 0, ie si on a:

$$\mathbb{P}(|X_n - X| > \varepsilon) \longrightarrow 0$$

On notera alors:

$$(X_n) \xrightarrow{\mathcal{P}} X$$

Convergence presque partout

On dira que X_n **converge presque partout** vers X si et seulement si elle converge sauf sur un domaine de mesure nul, ie:

$$\mathbb{P}\left(\left\{\omega \in \Omega\,;\ \lim_{n\to+\infty} X_n(\omega) \neq X(\omega)\right\}\right) = 0$$

On notera alors:

$$(X_n) \longrightarrow X \ p.p.$$

Convergence L^p

On dira qu'une suite (X_n) converge en norme p vers une variable aléatoire X si et seulement si:

$$\lim_{n\to\infty} \|X_n - X\|_p = \lim_{n\to\infty} \left(\mathbb{E}(|X_n - X|^p)\right)^{\frac{1}{p}} = 0$$

Relations entre les convergences

On peut montrer les différentes implications suivantes:

$$\textbf{Convergence p.p.} \implies \textbf{Convergence en probabilité} \implies \textbf{Convergence en loi}$$

Et pour $p > q \geq 1$

$$\textbf{Convergence } L^p \implies \textbf{Convergence } L^q \implies \textbf{Convergence en probabilité}$$

XII — Théorèmes limites

Munis de nos nouveaux outils et modes de convergences des suites de variables aléatoires, on peut alors énoncer et démontrer les grands théorèmes probabilistes.

Lois des grands nombres

On se donne une suite de variables aléatoires indépendantes et identiquement distribuées (communément noté iid) (X_n) admettant une espérance μ, et on pose une variable aléatoire appellée **moyenne**[1] **empirique**:

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

Alors on peut montrer la loi **faible** des grands nombre, ie:

$$\overline{X}_n \xrightarrow{\mathcal{P}} \mu$$

Pour les memes hypothèses que précédemment la loi **forte** des grands nombres nous assure une convergence plus forte, ie on a:

$$\overline{X}_n \xrightarrow{p.s.} \mu$$

Théorème central limite

On se donne une suite de variables aléatoires iid (X_n) admettant une espérance μ et un écart-type σ finis, alors on considère à nouveau la moyenne empirique:

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

Mais cette fois on considère cette variable sous sa forme **centrée réduite**, qu'on notera $\overline{X}^*_n$, alors le théorème central limite nous donne que:

$$\overline{X}^*_n \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

[1] Elle correspond à la moyenne faite sur les réalisations d'une expérience par exemple.