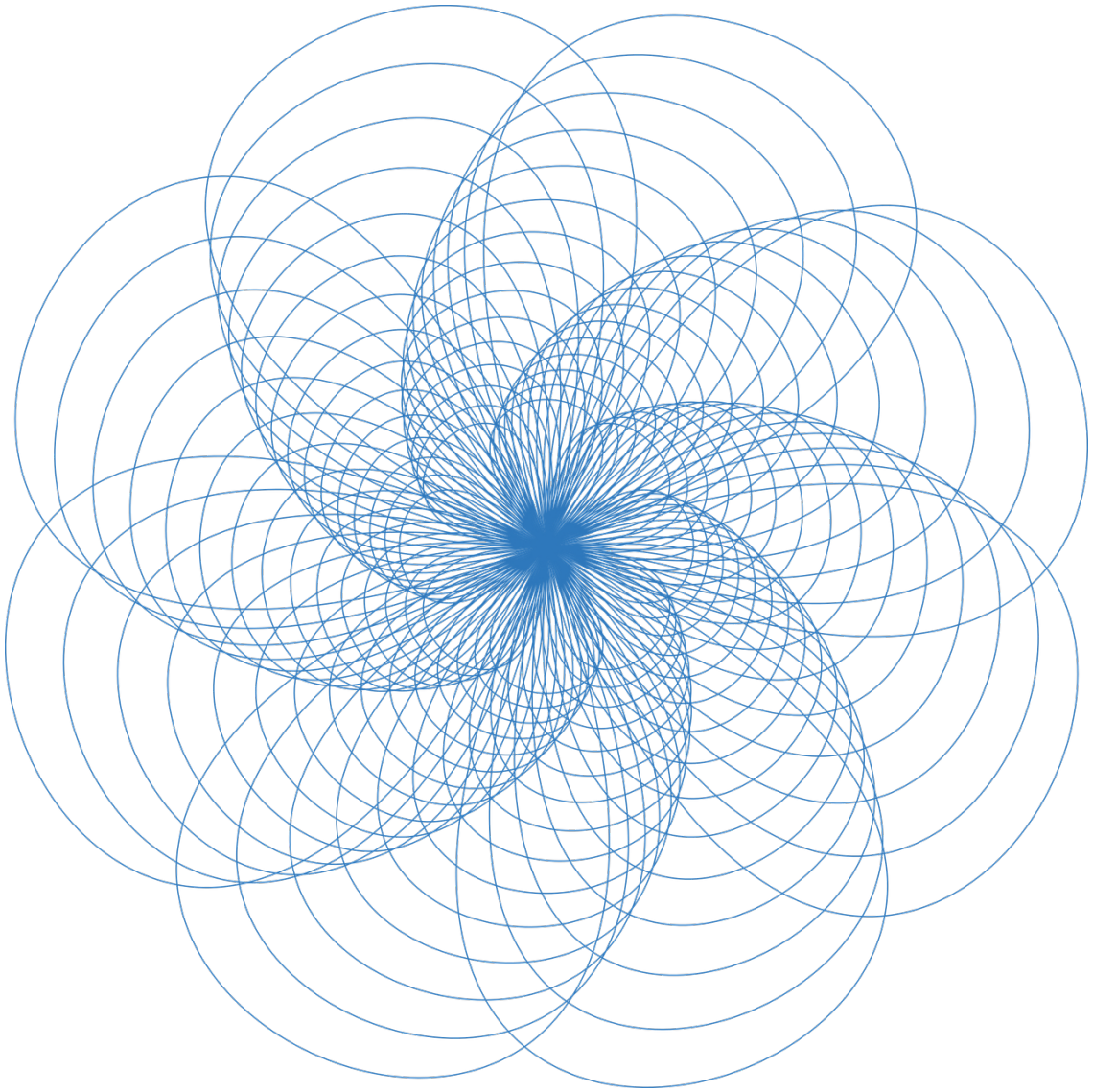
Mathématiques

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:

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

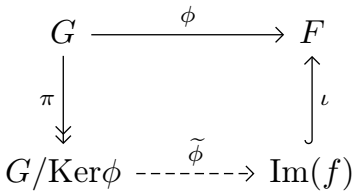
La propriété fondamentale de ces groupes, qui utilise le premier résultat du chapitre suivant est que les sous groupes normaux de G sont exactement les **noyaux** de morphismes de domaine G.

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.

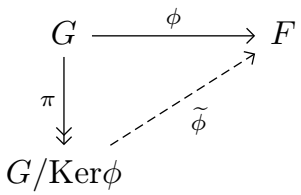
Premier théorème d'isomorphisme :

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]:



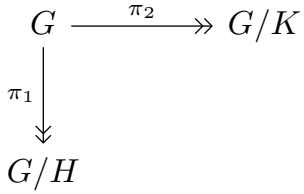
En effet, le passage au quotient rend le morphisme injectif, donc surjectif sur son image, et le diagramme commute, ie on a $\phi = \iota \circ \widetilde{\phi} \circ \pi$.

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:

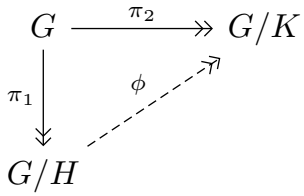


Deuxième théorème d'isomorphisme :

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

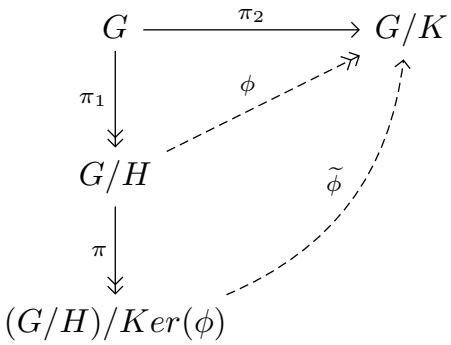


On peut alors utiliser la propriété universelle du quotient pour compléter le diagramme par un morphisme ϕ (par ailleurs surjectif):



[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.

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:

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

On remarque que ce groupe est doté d'une action naturelle sur $[\![1\ ;\ n]\!]$ donnée par $\sigma \cdot k = \sigma(k)$.

Support :

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

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

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

Cycles :

On appelle **k-cycle**[1] une permutation σ telle qu'il existe $k \geq 2$ et k éléments deux à deux distincts $a_1, \dots, a_k$ tels que:

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

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

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

Exemple: La permutation suivante est un 3-cycle:

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

Cycles et orbites :

Si σ est un k-cycle et que a n'est pas un point fixe, alors on en déduit que la donnée du support de σ est équivalente à la donnée de l'orbite de celle-ci (plus précisément du sous groupe engendré par celle-ci) pour son action naturelle, ie:

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

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] Si $k = 2$, on appellera un tel k-cycle une **transposition**.

Théorème de décomposition en cycles :

Une des problématiques principales à propos des groupes symétriques est la question de la décomposition d'une permutation en permutation plus simples. On peut tout d'abord 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. Il reste à choisir un représentant de chaque orbite et la décomposition en cycles est acquise.

Théorème de décomposition en transpositions :

Par la suite, on peut alors constater directement que pour tout k-cycle $\sigma = (a_1 \ \ldots \ a_k)$, on a une décomposition canonique en produit de transpositions:

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

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:

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

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** $[l_1, \ldots, l_k]$ 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 :

On considère une permutation σ de type $[l_1, \ldots, l_k]$, alors chacun de ses cycles se décompose en **le** produit de $l_i - 1$ transpositions défini ci-dessus. Il est alors naturel de définir alors la fonction suivante:

$$m(\sigma) = \sum_{i \leq k} (l_i - 1)$$

C'est le total du nombre de transpositions dans la décomposition en transpositions définie plus haut. On peut alors définir la **signature** d'une permutation par:

$$\varepsilon(\sigma) = (-1)^{m(\sigma)}$$

On montre alors facilement que cette fonction est un **morphisme de groupe** de $\mathfrak{S}_n \longrightarrow (\{-1, 1\}, \times)$ et que trivialement la signature d'une transposition est -1.

- Si $\varepsilon(\sigma) = 1$, on dira que cette permutation est **paire**.
- Si $\varepsilon(\sigma) = -1$, on dira que cette permutation est **impaire**.

L'ensemble des permutations de signature paire est alors un groupe (c'est le noyau de ε), qu'on appelle **groupe alterné** et qu'on note $\mathfrak{A}_n$.

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. On notera l'égalité dans ce contexte $a = b \pmod{n}$ en comprenant que ceci signife que $a + n\mathbb{Z} = b + n\mathbb{Z}$

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 $a + \ldots + a = 1 \pmod{n}$ 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:

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

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:

$$\begin{aligned} \arg(zz') &= \arg(z) + \arg(z') \pmod{2\pi} \\ \arg\Big(\frac{z}{z'}\Big) &= \arg(z) - \arg(z') \pmod{2\pi} \end{aligned}$$

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.

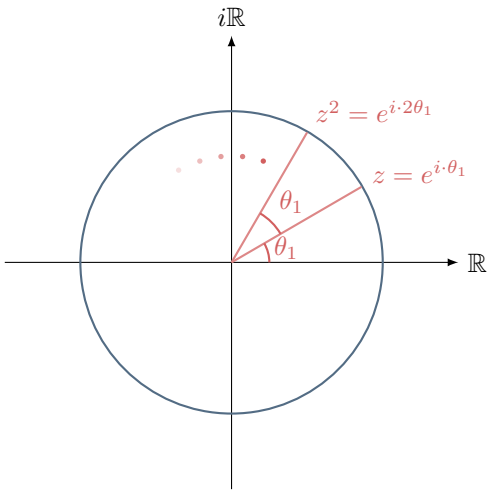
Formule de Moivre :

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

Soit $\mathbb{A}$ un anneau commutatif, on construit l'ensemble des **polynômes** à coefficients dans $\mathbb{A}$ comme l'ensemble des suites **nulles à partir d'un certain rang** d'éléments de $\mathbb{A}$, on peut alors munir cet ensemble des opérations naturelles suivantes:

- Une **addition** effectuée termes à termes.
- Une **multiplication par un scalaire** effectuée termes à termes.
- Une **multiplication polynomiale** définie par distributivité comme:

$$PQ = (p_0q_0, p_1q_0 + p_0q_1, \ldots) = \left(\sum_{k=0}^{n} P_kQ_{n-k}\right)_{n\in\mathbb{N}}$$

C'est bien une suite nulle à partir d'un certain rang et elle correspondra alors à la distributivité usuelle[1].

On note alors $X := (0, 1, 0, \ldots)$ et on appelle cette suite **indeterminée**, on remarque alors que:

$$\forall n, m \in \mathbb{N}\ ;\ X^nX^m = X^{n+m}$$

Et par suite que tout polynôme peut s'écrire comme combinaison linéaire de cette indeterminée, ce qui nous donne finalement l'expression canonique d'un polynôme et donc la définition canonique de l'ensemble des polynômes en l'indeterminée X donnée par:

$$\mathbb{A}[X] := \left\{\sum_{n\in\mathbb{N}} a_kX^k\ ;\ (a_k) \in \mathbb{A}^{\mathbb{N}}\right\}$$

Où la suite (a_n) est nulle à partir d'un certain rang.

Structure :

Ces opérations et la structure d'anneau commutatif des coefficients donnent alors une structure **d'anneau commutatif** à l'ensemble $\mathbb{A}[X]$, en outre on pourra aussi vérifier aprés avoir lu le chapitre correspondant que c'est un **espace vectoriel** de dimension infinie, et dont une base est donnée par:

$$(1, X, X^2, \ldots)$$

Finalement, si $\mathbb{A}$ est intégre (par exemple dans le cas usuel où c'est un corps), l'anneau $\mathbb{A}[X]$ l'est aussi.

Evaluation :

Soit R un anneau quelconque qui contient $\mathbb{A}$, et $x \in R$ alors on peut montrer qu'il existe une application fondamentale, dite application **d'évaluation** donné par:

$$\begin{aligned}\phi : \mathbb{A}[X] \times R &\longrightarrow R\\ (P, x) &\longmapsto P(x) = \sum a_kx^k\end{aligned}$$

En effet, si on fixe un élément $x \in R$, alors on a simplement $\forall P \in \mathbb{A}[X]\ ;\ \phi(P, x) = P(x)$ qui est simplement le polynôme inital dont on a substitué l'indeterminée par un élément de l'anneau. En particulier si $R = \mathbb{A}[X]$, on a directement l'identification $P = P(X)$.

[1] C'est un cas particulier de produit de convolution discret, voir le chapitre sur la convolution.

Dans la suite on se restreint au cas $\mathbb{A} = \mathbb{K}$ des polynômes à coefficients dans un corps. Cette contrainte supplémentaire nous permettra de développer une arithmétique plus riche des polynômes.

Degré et Valuation :

Soit $P, Q \in \mathbb{K}[X]$, on peut alors définir tout une propriété fondamentale appelée **degré** de P qui découle directement de la construction des polynômes:

$$\deg(P) := \max\{k \in \mathbb{N} \ ; \ a_k \neq 0\}$$

On a alors les propriétés opératoires du degré ci-dessous:

- **Degré de la somme:** $\deg(P+Q) \leq \max(\deg(P), \deg(Q))$
- **Degré du produit:** $\deg(PQ) = \deg(P) + \deg(Q)$

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

Divisibilité :

On peut naturellement définir une notion de **divisibilité** dans l'anneau $\mathbb{A}[X]$ qui vérifie toute les propriétés usuelles, mais la notion de degré nous permet aussi de définir 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**.

Plus précisément, on a le théorème suivant pour tout couple $A, B \in \mathbb{K}[X]$:

$$\exists!(Q, R) \in \mathbb{K}[X] \ ; \ A = BQ + R \text{ avec } \deg(R) < \deg(B)$$

La démonstration de ce théorème se fait de manière analogue à celui de $\mathbb{Z}$, ie en exhibant l'algorithme de division. Ce théorème donne alors à $\mathbb{K}[X]$ une structure **d'anneau Euclidien**, dont découle les conséquences suivantes:

- **Anneau de Bezout:** La relation de Bezout est vraie pour les polynômes.
- **Anneau principal:** Les idéaux sont principaux.
- **Anneau factoriel:** Il existe une décomposition en facteurs premiers.

Racines:

Soit $\alpha \in \mathbb{K}$ et $P \in \mathbb{K}[X]$, alors on peut montrer le théorème fondamental ci-dessous:

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

On dira alors que α est **racine** de P si une de ces conditions est vérifiée.

Multiplicité:

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

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

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 $m_0, m_1, \ldots, m_{n-1}, m_n$, le lemme de Gauss nous permet de montrer qu'alors:

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

Caractérisation de la multiplicité:

Si on note P^m la dérivée n-ième de P, on peut caractériser le fait que α soit de multiplicité m par:

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

Facteurs premiers :

On appelle **facteurs premiers** de $\mathbb{K}[X]$ les polynômes (non-constants) qui n'admettent pas de **diviseurs stricts** (non-constants), ces éléments dépendents du corps considéré, en effet par exemple:

- **Dans** $\mathbb{R}[X]$**:** X^2+1 est premier.
- **Dans** $\mathbb{C}[X]$**:** $X^2+1=(X-i)(X+i)$ n'est pas premier.

On dira qu'un polynôme est **scindé** sur $\mathbb{K}[X]$ si ses facteurs sont tous de degré 1.

Décomposition :

Un des grans thèmes de l'étude des polynômes est alors la recherche de la décomposition de ceux-ci en facteurs premiers, par exemple:

- Si on considère l'anneau $\mathbb{C}[X]$, on peut alors montrer le **théorème fondamental de l'Algèbre**:

  Tout polynôme non-constant admet une racine.

  Et donc en particulier par récurrence tout les polynômes de $\mathbb{C}[X]$ sont **scindés**.
- Si on considère l'anneau $\mathbb{R}[X]$, il existe donc des polynômes de degré 2 irréductibles. Mais on sait alors qu'il ceux ci ont des racines complexes, et par évaluation on trouve la propriété intéressante suivante:

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

- Si on considère l'anneau $\mathbb{F}_2[X]$, on peut par exemple remarquer la factorisation: $X^2+1=(X+1)^2$

Polynômes en plusieurs indeterminées :

On peut alors généraliser la construction des polynômes en une indeterminée X en un anneau de polynômes en plusieurs indeterminées $\mathbb{A}[X_1,\dots,X_n]$ qu'on définit par récurrence par:

$$\mathbb{A}[X_1,\dots,X_n]=\left(\mathbb{A}[X_1,\dots,X_{n-1}]\right)[X_n]$$

C'est aussi un **anneau commutatif**, aussi intègre si $\mathbb{A}$ l'est et ses éléments sont alors de la forme:

$$\mathbb{A}[X_1,\dots,X_n]:=\left\{\sum_{i_1,\dots,i_n\in\mathbb{N}} a_{i_1,\dots,i_n}X_1^{i_1}\dots X_n^{i_n}\right\}$$

Où la somme est finie, ie où l'ensemble des coefficients $(a_{i_1,\dots,i_n})_{i_1,\dots,i_n\in\mathbb{N}}$ est une famille finie. Quelques exemples:

- **Dans** $\mathbb{R}[X,Y]$**:** $P=2X^2+3XY-Y$
- **Dans** $\mathbb{C}[X,Y,Z]$**:** $P=iX^2Y^2+2iX-5Z$

Relations coefficients racines :

On peut trouver une relation entre les coefficients et les racines d'un polynôme qui peut souvent nous permettre de nous ramener à la résolution d'un système et potentiellement trouver les racines, en effet on suppose la décomposition acquise alors on a:

$$P=\sum_{k=0}^{n}a_kX^k=a_n(X-\alpha_1)\dots(X-\alpha_n)$$

En développant on trouve alors des relations pour la somme, la somme des doubles produits, la somme des triples produits, etc, et le produit des racines:

- **La somme des racines:** $\sum_{1\leq i\leq n}\alpha_i=(-1)^1\frac{a_{n-1}}{a_n}$
- **La somme des k-produits des racines:** $\sum_{1\leq i_1<\dots<i_k\leq n}\alpha_{i_1}\dots\alpha_{i_k}=(-1)^k\frac{a_{n-k}}{a_n}$
- **Le produit des racines:** $\alpha_1\dots\alpha_n=(-1)^n\frac{a_0}{a_n}$

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:

$$\begin{cases} \|X\|_p := \left(\int_\Omega |X|^p d\mathbb{P}\right)^{\frac{1}{p}} = \left(\mathbb{E}(|X|^p)\right)^{\frac{1}{p}} \\ \|X\|_\infty := \sup \operatorname{ess}\{|X|\} \end{cases}$$

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.