Chapitre 4

Ensembles et applications.

Le but de ce chapitres est de se donner les principaux outils nécessaires pour manipuler des ensembles et des applications. Nous allons revoir puis approfondir les résultats vus dans le chapitre 0.

I Généralités sur les ensembles.

I.1 Premières définitions (rappel).

Définition.

- Un ensemble E est une collection d'objets appelés éléments de E.
- Si x est un élément de E, on note $x \in E$ et on dit que l'objet x appartient à E. Sinon, on note $x \notin E$ et on dit que l'objet x n'appartient pas à E.
- On appelle ensemble vide, et on note $\varnothing$, l'ensemble ne contenant aucun élément.

Comme dit dans le chapitre 0, nous nous contenterons d'une approche intuitive, qui a ses limites, cf. exercice 20.

I.2 Modes de définition d'un ensemble (rappel).

Il existe plusieurs façons de définir un ensemble. Les trois principales sont les suivantes :

1. Définir un ensemble grâce à des unions, intersections, passages au complémentaire (cf. partie IV) d'ensembles déjà existants. Par exemple, $E = [\,0\,;1\,] \cap \overline{\mathbb{Q}}$ est l'ensemble des irrationnels appartenant à $[\,0\,;1\,]$.
2. Définir un ensemble en donnant explicitement tous ses éléments entre accolades et séparés par des points virgules. Par convention un élément ne figure qu'une seule fois dans la liste.

On dit alors que l'ensemble est défini par extension.

Exemples :

- $\{0; 1\}$ est l'ensemble à deux éléments contenant uniquement 0 et 1.
- $\{1; 2; 3; 4; 5; 6\}$ est l'ensemble à six éléments contenant les entiers de 1 à 6. On le note également $[\![\,1\,;6\,]\!]$ (cf. chapitre 0).
- $\{0\}$ est l'ensemble à un élément contenant uniquement 0. On l'appelle « singleton 0 ». Plus généralement :

Définition. Un ensemble à un élément est appelé un singleton. Si on note x cet élément, cet ensemble (noté donc $\{x\}$) est appelé « singleton x ».

Remarque : Inconvénients de ce mode de définition :

- Il nécessite de connaître tous les éléments de l'ensemble.
- Il est difficilement maniable avec un grand nombre d'éléments : on le voit rien qu'avec 6 éléments !

3. Définir un ensemble par une propriété caractérisant ses éléments. Plus précisément, si P est une propriété, on note $\{x \in E \,|\, P(x)\}$ l'ensemble des éléments de E qui vérifient la propriété P. Détaillons cette notation :

On dit alors que l'ensemble est défini par compréhension.

$$\underbrace{\{}_{\text{accolade}}\ \underbrace{\dots}_{\substack{\text{Nature des éléments}\\ \text{de l'ensemble}\\ \text{(réels, entiers etc.)}}}\ \underbrace{|}_{\substack{\text{barre verticale}\\ \text{de séparation :}\\ \text{lire « tels que »}}}\ \underbrace{\dots}_{\substack{\text{Propriété(s) vérifiée(s)}\\ \text{par les éléments}\\ \text{de l'ensemble}}}\ \underbrace{\}}_{\text{accolade}}$$

Remarque : Nous nous intéressons seulement à la façon d'écrire un ensemble. Les notions mathématiques présentes ci-dessous (fonctions paires etc.) seront vues dans des chapitres ultérieurs.

Exemples :

- $P = \{f : \mathbb{R} \longrightarrow \mathbb{R} \,|\, \forall x \in \mathbb{R},\, f(-x) = f(x)\}$ est l'ensemble des fonctions f de $\mathbb{R}$ dans $\mathbb{R}$ telles que, pour tout $x \in \mathbb{R}$ $f(-x) = f(x)$. En d'autres termes, c'est l'ensemble des fonctions paires.
- $\{x \in \mathbb{R} \,|\, x^2 = 1\}$ est l'ensemble des réels dont le carré vaut 1. En d'autres termes, c'est l'ensemble à deux éléments $\{-1; 1\}$.
- $\pi\mathbb{Z} = \{x \in \mathbb{R} \,|\, \sin(x) = 0\}$ est l'ensemble des réels x tels que $\sin(x) = 0$. En d'autres termes, c'est l'ensemble des réels congrus à 0 modulo π.
- $\{f : \mathbb{R} \longrightarrow \mathbb{R} \,|\, f \text{ dérivable sur } \mathbb{R},\, f'(0) = f'(1) = 0\}$ est l'ensemble des fonctions f de $\mathbb{R}$ dans $\mathbb{R}$ dérivables et vérifiant $f'(0) = f'(1) = 0$.
- $S = \{x \in \mathbb{R} \,|\, x^5 - 3x - 1 = 0\}$ est l'ensemble des réels x solutions de l'équation $x^5 - 3x - 1 = 0$.

$\{x \in E \,|\, P(x)\}$ est l'ensemble des éléments x appartenant à E tels que $P(x)$ soit vraie. En d'autres termes, c'est l'ensemble des éléments de l'ensemble se situant à gauche de la barre verticale vérifiant la propriété à droite de cette barre. Ou enfin : si $x \in E$, alors x appartient à cet ensemble si et seulement si $P(x)$ est vraie. Avec le dernier exemple ci-contre : « soit $x \in \mathbb{R}$. Alors : $x \in S \iff x^5 - 3x - 1 = 0$ ». Une rapide étude de fonction couplée au théorème de la bijection prouve que cet ensemble comporte trois éléments car l'équation $x^5 - 3x - 1 = 0$ a trois solutions. Cependant, on ne les connaît pas, ce qui n'empêche pas de définir S !

Remarque : Avantages de ce mode de définition :

- Il permet de définir et de manipuler facilement des ensembles infinis.
- Il n'est pas nécessaire de connaître explicitement tous les éléments de l'ensemble pour le définir. C'est particulièrement frappant avec les deux derniers exemples ci-dessus : on serait particulièrement en peine de donner toutes les fonctions f dérivables sur $\mathbb{R}$ vérifiant $f'(0) = f'(1) = 0$, et on ne sait pas résoudre l'équation $x^5 - 3x - 1 = 0$. On ne sait même pas, au premier abord, combien de solutions cette équation possède !

Remarque : L'ensemble des entiers naturels pairs (parfois noté $2\mathbb{N}$) peut être écrit de la façon suivante :

$$2\mathbb{N} = \{n \in \mathbb{N} \,|\, \exists k \in \mathbb{N},\, n = 2k\}$$

mais c'est tout de même une notation assez lourde. On utilisera plutôt la notation plus concise $2\mathbb{N} = \{2k \,|\, k \in \mathbb{N}\}$. Plus généralement, si $f : E \longrightarrow F$ est une application (nous reverrons donc cette notation dans le chapitre 4) et si A est une partie de E, on note $\{f(x) \,|\, x \in A\}$ ou $f(A)$ l'ensemble des images par f des éléments de A. On peut lire : « $\{f(x) \,|\, x \in A\}$ est l'ensemble formé des $f(x)$ lorsque x parcourt A ». Là aussi, détaillons cette notation :

$$\underbrace{\{}_{\text{accolade}} \quad \underbrace{f(x)}_{\substack{\text{Éléments de} \\ \text{l'ensemble : images} \\ \text{d'un paramètre} \\ x \text{ par } f}} \quad \underbrace{|}_{\substack{\text{barre verticale} \\ \text{de séparation :} \\ \text{lire « lorsque »}}} \quad \underbrace{x \in A}_{\substack{\text{Le paramètre} \\ \text{et l'ensemble} \\ \text{qu'il parcourt}}} \quad \underbrace{\}}_{\text{accolade}}$$

Là aussi, précisons une chose (encore une fois, nous reverrons cette notation dans la partie V) : si $f : E \longrightarrow F$,

$$f(A) = \{f(x) \,|\, x \in A\}$$

est l'ensemble des images par f des éléments de A. En d'autres termes, un élément de F appartient à cet ensemble si et seulement s'il s'écrit sous la forme $f(x)$ avec $x \in A$. Avec des quantificateurs, cela donne : « Soit $y \in F$. Alors : $y \in f(A) \iff \exists x \in A, y = f(x)$ ». Cette équivalence est à connaître sur le bout des doigts ! Par exemple, avec l'ensemble E ci-contre : « Soit $x \in \mathbb{R}$. Alors : $x \in E \iff \exists n \in \mathbb{N},\, x = u_n$ ». Attention x n'est pas forcément un entier !

Exemples :

- $C = \left\{n^2 \,|\, n \in \mathbb{N}\right\}$ est l'ensemble des carrés parfaits.
- Plus généralement, si $(u_n)_{n \in \mathbb{N}}$ est une suite (donc une application de $\mathbb{N}$ dans $\mathbb{R}$, cf. chapitre 12), $E = \{u_n \,|\, n \in \mathbb{N}\}$ est l'ensemble des termes de la suite.
- $\{\sin(x) \,|\, x \in \mathbb{R}\} = [\,-1\,;1\,]$.

I.3 Comparaison d'ensembles

Définition. On dit que deux ensembles E et F sont égaux, et on note $E = F$, si ils ont les mêmes éléments.

Remarque : Avec des quantificateurs : $\forall x,\, x \in E \iff x \in F$. On peut donc prouver que deux ensembles sont égaux en raisonnant par équivalences, cf. partie IV (mais, en général, on raisonnera par double inclusion : voir ci-dessous).

Un élément appartient ou n'appartient pas à un ensemble, il n'y a pas de notion d'ordre ou de multiplicité (il appartient ou non à l'ensemble, il ne peut pas lui appartenir plusieurs fois). Par exemple, les trois ensembles $\{0;1\}, \{0;0;1\}$ et $\{1;0\}$ sont égaux : ils contiennent 0 et 1 et aucun autre élément, ils ont les mêmes éléments donc sont égaux.

Définition. Soient E et F deux ensembles.

- On dit que F est inclus dans E, et on note $F \subset E$, si : $\forall x \in F, x \in E$.
- Si F n'est pas inclus dans E, on note $F \not\subset E$.
- Si $F \subset E$ et $F \neq E$, on dit que l'inclusion est stricte et on note $F \subsetneq E$.

Pour quantifier $F \subset E$, on trouve aussi l'écriture suivante :

$$\forall x, x \in F \Rightarrow x \in E$$

En particulier, l'ensemble vide est inclus dans tous les ensembles. En effet, on a vu (cf. chapitre 0) que si A est fausse, alors $A \Rightarrow B$ est vraie. Par conséquent, si F est l'ensemble vide et si E est un ensemble quelconque, l'assertion « $x \in F$ » est fausse donc l'implication ci-dessus est vraie.

Remarque : Ne pas confondre $\subset$ et $\in$. Par exemple, si $E = \{0;1;2\}$, alors $1 \in E$ et $\{1\} \subset E$. On ne parle d'inclusion que pour des ensembles : un ensemble est **inclus** dans un autre si tous ses éléments **appartiennent** au second ensemble, cf. exercice 2.

Remarques :

- Tout ensemble est inclus dans lui-même.
- Pour montrer que F est inclus dans E, on se donne un élément x de F et on montre que x est un élément de E. On commence donc par écrire « Soit $x \in F$. Montrons que $x \in E$. »
- Par transitivité de l'implication, si E, F et G sont des ensembles tels que $G \subset F$ et $F \subset E$, alors $G \subset E$.

Exemples :

- Nous avons $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$.
- Notons $D(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions dérivables sur $\mathbb{R}$ et $C(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions continues sur $\mathbb{R}$. Si $f : \mathbb{R} \longrightarrow \mathbb{R}$ est dérivable sur $\mathbb{R}$, alors elle est continue sur $\mathbb{R}$ (cf. chapitre 14). Ainsi $D(\mathbb{R}, \mathbb{R}) \subset C(\mathbb{R}, \mathbb{R})$.

Il résulte de la définition l'équivalence (importante !) suivante :

Proposition (double inclusion). Si E et F sont deux ensembles, alors on a :

$$E = F \iff (E \subset F \text{ et } F \subset E).$$

Pour montrer que deux ensembles E et F sont égaux, on montrera souvent que l'un est inclus dans l'autre et vice versa. On dit qu'on procède par double inclusion. Raisonnement très classique !

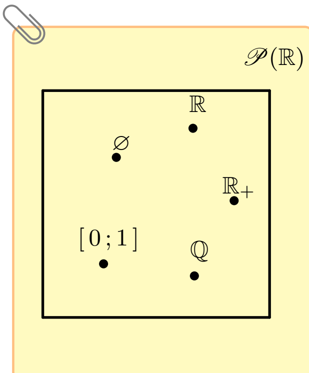
I.4 Parties d'un ensemble

Définition. Soient E et F deux ensembles. Si F est inclus dans E, on dit que F est une partie de E ou un sous-ensemble de E.

Définition. Soit E un ensemble. On note $\mathscr{P}(E)$ l'ensemble des parties de E. On a alors :

$$A \in \mathscr{P}(E) \iff A \subset E.$$

Remarque : Là encore, il faut bien comprendre la différence entre appartenance et inclusion : un ensemble **appartient** à $\mathscr{P}(E)$ lorsqu'il est **inclus** dans E. La figure ci-contre (non rigoureuse !) illustre (dans le cas particulier où $E = \mathbb{R}$) cette différence. Ce qui, dans $\mathbb{R}$, sur la droite réelle, apparaît comme une partie (un ensemble d'éléments) apparaît dans $\mathscr{P}(\mathbb{R})$ comme un élément, un « point ». Une partie de $\mathscr{P}(\mathbb{R})$ serait donc un ensemble dont les éléments sont des parties de $\mathbb{R}$, voir des exemples ci-dessous.



Remarque : L'ensemble des parties de E est un ensemble d'ensembles, ce qui le rend difficilement maniable, et on ne sait pas l'expliciter en général (sauf si E est fini, voir ci-dessous) ce qui entraîne parfois des confusions entre parties de E et éléments de E. En particulier, puisqu'on n'explicite que rarement $\mathscr{P}(E)$, les élèves ont souvent du mal à répondre à cette simple question : quand un ensemble A appartient-il à $\mathscr{P}(E)$? L'équivalence ci-dessus permet de répondre simplement et sans erreur possible à cette question, et donc cette équivalence doit être sue sur le bout des doigts.

Exemples : Quand E est un ensemble fini, on peut facilement donner tous les éléments de $\mathscr{P}(E)$ (du moins en théorie... $\mathscr{P}(E)$ a beaucoup d'éléments : on peut montrer, cf. chapitre 17, que, si E a n éléments, alors $\mathscr{P}(E)$ a 2^n éléments) : il suffit de faire varier k de 0 à n et de donner toutes les parties de E à k éléments (et pour cela on prend k éléments de E de toutes les façons possibles) qu'on met entre accolades (définition par extension).

- $\mathscr{P}(\{0;1\}) = \{\varnothing,;\{0\};\{1\};\{0,1\}\}$: comme dit ci-dessus, $\mathscr{P}(\{0;1\})$ contient les parties à zéro élément (l'ensemble vide), les ensembles à un élément (les singletons $\{0\}$ et $\{1\}$) et les ensembles à deux éléments ($\{0;1\}$ lui-même). De même, $\mathscr{P}(\{P,F\}) = \{\varnothing;\{P\};\{F\};\{P,F\}\}$.
- $\mathscr{P}(\{0\}) = \{\varnothing;\{0\}\}$ et $\mathscr{P}(\{\text{machin}\}) = \{\varnothing;\{\text{machin}\}\}$.
- Le seul ensemble inclus dans l'ensemble vide est lui-même donc $\mathscr{P}(\varnothing) = \{\varnothing\}$.
- $\mathscr{P}(\{-1;0;1\}) = \{\varnothing;\{0\};\{-1\};\{1\};\{0;1\};\{0;-1\};\{1;-1\};\{0;1;-1\}\}$ (les ensembles à zéro élément, puis les ensembles à un élément, puis les ensembles à deux éléments, puis les ensembles à trois éléments).
- Si on lance un dé à 6 faces, on peut coder le résultat par un élément de $\Omega = [\![1\,;\,6]\!]$. En langage probabiliste on dira que Ω est l'univers associé à l'expérience aléatoire consistant à lancer un dé. On appellera toute partie de Ω un événement. Par exemple $\{1;3;5\}$ correspond à l'événement « obtenir un chiffre impair ». Ainsi $\mathscr{P}(\Omega)$ est l'ensemble des événements (cf. chapitre 26). On ne l'explicite pas ici car il a 64 éléments !

Remarques :

- $\varnothing \in \mathscr{P}(E)$ et $E \in \mathscr{P}(E)$.
- $\{x\} \in \mathscr{P}(E) \iff \{x\} \subset E \iff x \in E$,

II Généralités sur les applications.

Dans le chapitre 2, nous avons étudié les fonctions de A (une partie de $\mathbb{R}$) dans $\mathbb{R}$. Cela consiste à associer à tout élément de A un unique élément de $\mathbb{R}$. Dans ce chapitre, nous allons généraliser cela : en associant à tout élément d'un ensemble non vide quelconque E un unique élément d'un (autre) ensemble non vide F. Par exemple :

> Nous allons donc revoir et approfondir les résultats vus dans les chapitres 0 et 2.

- la fonction f_1 qui à un événement associe l'année en lequel il a eu lieu.
- la fonction f_2 qui à une fonction φ de $\mathbb{R}$ dans $\mathbb{R}$ dérivable associe le couple $(\varphi(0), \varphi'(0))$.
- la fonction f_3 qui à un couple (x,q) de réels associe la suite $(xq^n)_{n\in\mathbb{N}}$.
- la fonction f_4 qui à un élève de la classe associe son prénom.

> Nous reverrons les couples, les familles, les ensembles produits dans la partie III.

On se donne dans la suite de ce chapitre trois ensembles non vides E, F et G.

Définition. Une application ou une fonction f de E dans F est la donnée pour chaque élément x de E d'un unique élément de F, noté $f(x)$. L'ensemble E est appelé ensemble de départ de f et F est appelé son ensemble d'arrivée.

> Nous nous contentons d'une définition intuitive de fonction. De plus, il y a en fait une distinction entre fonction et application, distinction que nous ne ferons pas dans cet ouvrage car elle dépasse le cadre du programme.

Définition. Soit f une application de E dans F. Soit $y \in F$.

- Soit $x \in E$. Si $y = f(x)$, on dit que y est l'image de x par f et que x est **un** antécédent de y par f.
- Si y admet un antécédent par f (c'est-à-dire s'il existe $x \in E$ tel que $y = f(x)$), on dit que y est atteint par f.

Exemples : Regardons de plus près les exemples de l'introduction :

- L'application f_1 va de l'ensemble des événements historiques dans $\mathbb{Z}$. Par exemple $f_1(\text{prise de la Bastille}) = 1789$.
- L'application f_2 va de $D(\mathbb{R}, \mathbb{R})$ (notation usuelle pour l'ensemble des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$, cf. chapitre 2) dans $\mathbb{R}^2$. Par exemple $f_2(\exp) = (1,1)$, $f_2(\sin) = (0,1)$ et $f_2(x \mapsto 3-2x) = (3,-2)$.
- L'application f_3 va de $\mathbb{R}^2$ dans $\mathbb{R}^{\mathbb{N}}$. Par exemple $f_3((0,7))$ est la suite nulle et $f_3((1,-2)) = ((-2)^n)_{n\in\mathbb{N}}$.
- L'application f_4 va de l'ensemble des élèves de la classe dans l'ensemble des prénoms.

Remarques :

- Ne pas confondre F, l'ensemble d'arrivée de f, avec l'image de f (cf. partie V). En particulier (voir ci-dessous), tous les éléments de F ne sont pas forcément atteints : il peut exister des éléments de F n'admettant aucun antécédent.
- Par définition, un élément de E n'a qu'une seule image, mais un élément de F peut très bien admettre plusieurs antécédents (voire même une infinité).

Exemples :

- Tout couple de réels (λ, μ) admet une infinité d'antécédents par la fonction f_2 de l'introduction : notamment les fonctions $x \longmapsto \alpha x^2 + \mu x + \lambda$ pour tout $\alpha \in \mathbb{R}$.
- Si f désigne la fonction nulle de $\mathbb{R}$ dans $\mathbb{R}$, alors tout réel est un antécédent de 0, tandis que si y est un réel non nul, alors y n'admet aucun antécédent.

Attention, dans l'écriture ci-contre, la variable x est muette. La fonction f ne dépend pas de x donc on peut très bien écrire

$$f : t \longmapsto f(t)$$

sans modifier en rien la fonction f. En particulier, la fonction f ne dépend pas de x, écrire (par exemple) « la fonction f est croissante pour tout x » n'a aucun sens ! Une fonction est un objet qui, à un élément de E, lui associe une image. Et c'est tout ! Et c'est déjà pas mal...

Définition. Une fonction f de E dans F est notée de la façon suivante :

$$f : \left\{ \begin{array}{ccc} E & \longrightarrow & F \\ x & \longmapsto & f(x) \end{array} \right.$$

Cette notation signifie : f est la fonction qui va de E dans F et qui à x associe $f(x)$. On écrit aussi souvent « soit $f : E \longrightarrow F$ » pour signifier : « soit f une fonction de E dans F ».

Remarques :

- On peut parfois ne pas expliciter E et F quand il n'y a aucune ambiguité. Par exemple, quand on parlera de la fonction $x \longmapsto x + \cos(x)$, il sera sous-entendu (sauf indication contraire) qu'elle va de $\mathbb{R}$ dans $\mathbb{R}$. De même, quand on parlera de la fonction $x \longmapsto \dfrac{x+1}{x-1}$, il sera sous-entendu (encore une fois sauf indication contraire) qu'elle va de $\mathbb{R}\backslash\{1\}$ (son domaine de définition) dans $\mathbb{R}$.
- Ne pas confondre la fonction f avec l'image $f(x)$! Ces deux objets n'ont pas la même nature et donc ne sauraient être confondus : f est une fonction et $f(x)$ est un élément de F, cela n'a rien à voir ! On ne parlera donc jamais de la fonction $f(x)$, tout simplement car ce n'est pas une fonction !
- Une fonction associe, à tout élément de E, un unique élément de F que l'on note $f(x)$, mais on n'a pas toujours une expression explicite de $f(x)$ en fonction de x. Si à tout élément x de E correspond un unique élément y de F (dépendant de x), cela définit bien une fonction (en posant, pour tout x, $f(x) = y$), même si on a aucune idée de la valeur de $f(x)$ (on dit alors qu'on définit f de façon implicite). Un exemple fréquent de définition de fonction implicite est quand on utilise le théorème de la bijection (cf. partie VII). Par exemple (cf. exercice 66) : si $\alpha > 0$ alors, pour tout réel x positif, il existe un unique réel positif noté $f(x)$ tel que $f(x)e^{f(x)} = x^\alpha$.

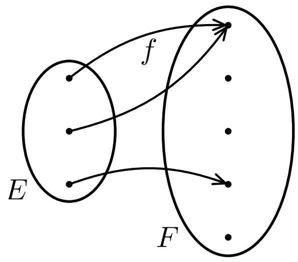
Quand on écrit que f va de E dans F, cela signifie que f prend ses valeurs dans F, cela ne signifie pas du tout que tous les éléments de F sont atteints. Par exemple, on peut dire que l'exponentielle et même la fonction nulle vont de $\mathbb{R}$ dans $\mathbb{R}$!

En d'autres termes : pour tout $x \geqslant 0$, il existe un unique $y \geqslant 0$ tel que $ye^y = x^\alpha$, et on pose $y = f(x)$. Cela définit bien une fonction f définie sur $\mathbb{R}_+$ car y existe et est unique.

Définition. Deux fonctions f et g de E dans F sont égales si, pour tout $x \in E$, $f(x) = g(x)$.

Remarque : Par contraposée : deux fonctions f et g sont distinctes **s'il existe** $x \in E$ tel que $f(x) \neq g(x)$. Deux fonctions sont distinctes si elles diffèrent en un point ! En particulier, deux fonctions distinctes peuvent être égales en certains points !

Remarque : Si E et F sont finis, on peut représenter f sous la forme d'un diagramme sagittal (patatoïde pour les intimes) : on représente les éléments de E d'un côté, ceux de F de l'autre et on relie par des flèches les éléments de E et leur image.

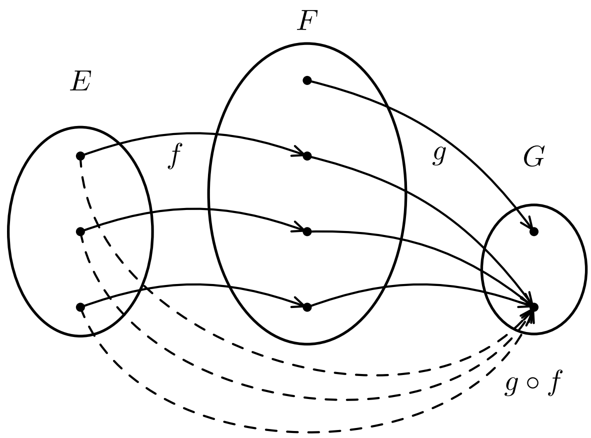


Définition. On note F^E ou $\mathscr{F}(E, F)$ l'ensemble des fonctions de E dans F.

Exemple : Une suite réelle (cf. chapitre 12) est une fonction de $\mathbb{N}$ dans $\mathbb{R}$. On note donc $\mathbb{R}^{\mathbb{N}}$ l'ensemble des suites réelles.

Définition. Soient $f : E \longrightarrow F$ et $g : F \longrightarrow G$. On appelle composée de f par g la fonction notée $g \circ f$ définie par :

$$g \circ f : \begin{cases} E & \longrightarrow & G \\ x & \longmapsto & g(f(x)) \end{cases}$$



Remarque : Dans l'exemple ci-dessus, $g \circ f$ est constante alors que ni f, ni g, n'est constante. Ainsi, le patatoïde ci-dessus permet de donner un contre-exemple à l'affirmation (fausse) : « si $g \circ f$ est constante, alors soit f, soit g, est constante ». Comme on le voit, les patatoïdes permettent d'exhiber des contre-exemples simples mais, attention, ils ont leurs limites :

- ils ne permettent pas de démontrer quoi que ce soit. Rappelons qu'un contre-exemple suffit pour prouver qu'une affirmation est fausse, mais qu'un exemple ne suffit pas à montrer qu'une affirmation est vraie.
- ils peuvent induire en erreur. En effet, certaines propriétés sont fausses uniquement sur des ensembles infinis : si on fait un diagramme sagittal pour se donner une idée, on pourra être persuadé qu'une propriété est vraie et s'échiner à tenter de la prouver alors qu'elle est fausse !

Exemple : On peut composer la fonction f_2 de l'introduction par la fonction f_3 puisque l'espace d'arrivée de f_2 est l'espace de départ de f_3. L'application $f_3 \circ f_2$ va de $D(\mathbb{R}, \mathbb{R})$ dans $\mathbb{R}^{\mathbb{N}}$ et associe, à toute fonction φ dérivable sur $\mathbb{R}$, la suite $\left(\varphi(0) \times (\varphi'(0))^n\right)_{n \in \mathbb{N}}$.

Attention, on rappelle qu'un élément de E a par définition une unique image. Ainsi, une seule flèche doit partir d'un élément de E. Cependant, un élément de F peut avoir plusieurs antécédents : plusieurs flèches peuvent arriver au même point !

Nous reviendrons sur la notation F^E dans la partie III.

Plus généralement, on peut définir $g \circ f$ quand f est définie sur un ensemble D_f, quand g est définie sur un ensemble D_g (non vides) et quand $f(D_f)$ est inclus dans D_g. Si cette dernière condition n'est pas vérifiée, $g \circ f$ n'est pas définie puisque, si $x \in D_f$, $f(x)$ n'appartient pas forcément à D_g et donc $g(f(x))$ n'existe pas forcément.

Si cela ne nous satisfait pas comme contre-exemple, on peut donner un exemple plus explicite, toujours en s'inspirant du diagramme ci-dessus : il suffit de prendre $E = [\![1 ; 3]\!]$, $F = [\![1 ; 4]\!]$ et $G = [\![1 ; 2]\!]$, ainsi que $f : E \longrightarrow F$ qui, à tout $n \in [\![1 ; 3]\!]$, associe n, et enfin, $g : F \longrightarrow G$ qui à $1, 2, 3$ associe 1 et à 4 associe 2. Mais bon, on ne fait rien d'autre que rajouter $1, 2, 3, 4$ à côté des points idoines...

La composition n'est pas commutative (alors qu'on montre facilement qu'elle est associative), c'est-à-dire qu'en général, $f \circ g \neq g \circ f$!

- D'une part, $g \circ f$ peut être définie sans que $f \circ g$ le soit.
- D'autre part, même si $g \circ f$ et $f \circ g$ sont toutes les deux définies (par exemple si f et g vont toutes les deux de E dans E), $f \circ g$ et $g \circ f$ ne sont pas forcément égales.

Exemple : Soient $f : x \longmapsto e^x$ et $g : x \longmapsto x^2$ (toutes les deux de $\mathbb{R}$ dans $\mathbb{R}$: comme dit ci-dessus, dans ce genre de cas, l'ensemble de départ et l'ensemble d'arrivée sont sous-entendus). Alors $f \circ g : x \longmapsto e^{x^2}$ et $g \circ f : x \longmapsto (e^x)^2 = e^{2x}$. Or, $g \circ f(1) = e^2 \neq e^1 = f \circ g(1)$ donc $g \circ f \neq f \circ g$.

Attention, dire : « Soit $x \in \mathbb{R}$. $e^{2x} \neq e^{x^2}$ donc $f \circ g \neq g \circ f$ » est faux ! Cela sous-entend que ces deux quantités sont toujours différentes, alors qu'elles coïncident parfois (en 0 et en 2 pour être précis). Rappelons que pour prouver que deux fonctions sont distinctes, il suffit de prouver qu'elles diffèrent **en un point**. Il suffit donc d'exhiber **un point explicite** en lequel elles diffèrent : dire que $f \circ g(1) \neq g \circ f(1)$ suffit donc pour prouver que $f \circ g \neq g \circ f$.

Définition. L'application $f : \left\{ \begin{array}{ccc} E & \longrightarrow & E \\ x & \longmapsto & x \end{array} \right.$ est appelée identité de E et est notée Id_E.

Proposition. Soit $f : E \longrightarrow F$. Alors $f \circ \mathrm{Id}_E = \mathrm{Id}_F \circ f = f$.

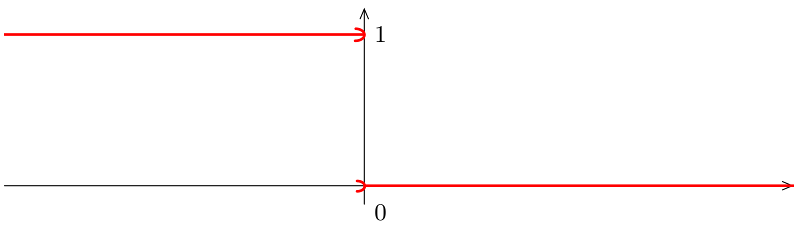
Démonstration. $\rightsquigarrow$ Exercice.

Remarque : On dit que l'identité est un élément neutre pour la composition, cf. chapitre 18.

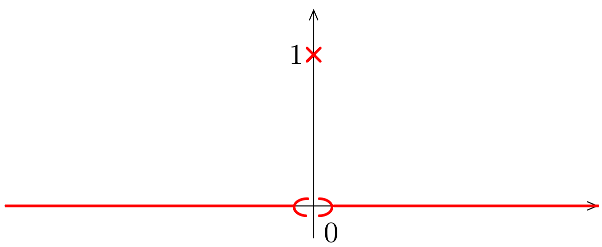
Définition. Soit A une partie de E. On appelle fonction indicatrice de A la fonction

$$\mathbb{1}_A : \left\{ \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \left\{ \begin{array}{ll} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{array} \right. \end{array} \right.$$

Exemples : Ci-dessous le graphe de $\mathbb{1}_{\mathbb{R}_-}$:



Ci-dessous le graphe de $\mathbb{1}_{\{0\}}$:



Proposition. Soient A et B des parties de E. $A = B \iff \mathbb{1}_A = \mathbb{1}_B$.

La fonction indicatrice caractérise donc la partie. On l'appelle donc parfois pour cette raison la fonction caractéristique de A, que l'on note χ_A.

Démonstration. $\rightsquigarrow$ Exercice.

Définition. Soit $f : E \to F$. Soit A une partie de E. On appelle restriction de f à A la fonction notée $f_{|A}$ suivante :

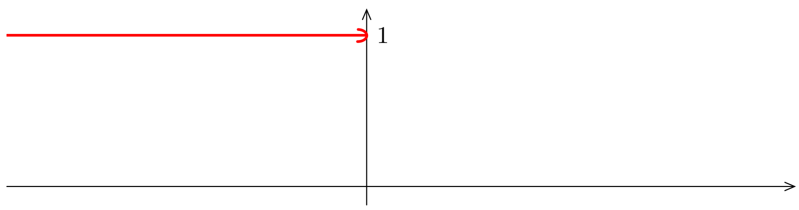
$$f_{|A} : \left\{ \begin{array}{ccc} A & \to & F \\ x & \mapsto & f(x) \end{array} \right.$$

Remarque : C'est juste la même fonction, définie sur un ensemble plus petit. L'avantage est qu'en se restreignant à un ensemble plus petit, la fonction peut vérifier des conditions ou des propriétés qu'elles ne vérifiait pas quand elle était définie sur E tout entier.

Exemples :

- Si f est la fonction inverse définie sur $\mathbb{R}^*$, sa restriction à $\mathbb{R}_+^*$ est décroissante (alors que f ne l'est pas).
- La restriction de $\mathbb{1}_{\mathbb{R}_-}$ à $\mathbb{R}_-$ est continue car constante égale à 1 :

De la même façon, si $n \in \mathbb{Z}$, la restriction de la partie entière à $[\,n\,;n+1\,[$ est une fonction continue.

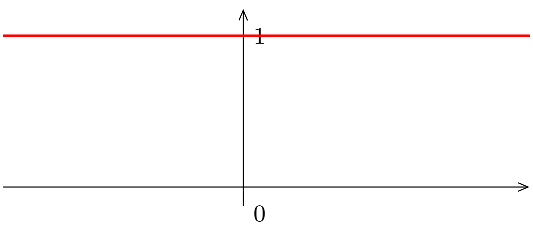


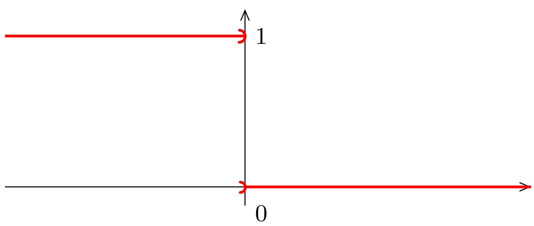
- La restriction d'une fonction à un ensemble plus petit peut être injective ou bijective (cf. paragraphe IV) alors que la fonction originelle ne l'est pas : voir des exemples dans le chapitre 5. On dit alors (le cas échéant) que f induit une injection ou une bijection d'un ensemble dans un autre. Bref, on en reparle.

Définition. Si g est une restriction de f, on dit que f est un prolongement de g.

Remarque : En d'autres termes, f est un prolongement de g si f est définie sur un ensemble plus grand et si $f = g$ là où g est définie. **Prolonger** une fonction g consiste donc à la définir en des points en lesquels elle n'était pas définie auparavant. Il n'y a bien sûr pas unicité du prolongement : par exemple, la fonction constante égale à 1 sur $\mathbb{R}_-$ (graphe en rouge ci-dessus) peut être prolongée en la fonction constante égale à 1 sur $\mathbb{R}$ (graphe à gauche ci-dessous) ou en $\mathbb{1}_{\mathbb{R}_-}$ (graphe à droite).

Il est donc idiot de parler de prolongement d'une fonction en un point en lequel elle est déjà définie... Gardez ça en tête quand vous appliquerez le théorème de la limite de la dérivée (cf. chapitre 14).





Cette année, nous nous contenterons de prolongements par continuité : cf. chapitres 2 et 13.

III Retour aux ensembles.

III.1 Produit cartésien.

III.1.a Notion de famille.

Définition. Soit I un ensemble non vide.

- Une famille $(x_i)_{i\in I}$ indexée par I est une application x définie sur I.
- L'ensemble I est appelé l'ensemble des indices de la famille $(x_i)_{i\in I}$.
- Si J est une partie non vide de $(x_i)_{i\in I}$, alors la famille $(x_j)_{j\in J}$ est appelée une sous-famille de $(x_i)_{i\in I}$.
- Si $(x_j)_{j\in J}$ est une sous-famille d'une famille $(x_i)_{i\in I}$, on dit que $(x_i)_{i\in I}$ est une

sur-famille de $(x_j)_{j\in J}$.

- Si I et J sont deux ensembles non vides disjoints (voir ci-dessous) et si $(x_i)_{i\in I}$ et $(x_j)_{j\in J}$ sont deux familles, la famille $(x_i)_{i\in I\cup J}$ est appelée la concaténation des familles $(x_i)_{i\in I}$ et $(x_j)_{j\in J}$.

Nous verrons les unions dans le paragraphe III. Pour l'instant, on se contente de l'interprétation avec les mains : la concaténation est la « mise bout à bout » des deux familles.

Remarques :

- On ne précise pas l'ensemble d'arrivée car il est quelconque : on parlera aussi bien de famille de réels (l'ensemble d'arrivée sera alors $\mathbb{R}$) que de famille de parties de E (l'ensemble d'arrivée sera alors $\mathscr{P}(E)$). Ainsi, on appelle famille indexée par I **toute** fonction définie sur I.
- Dans cette situation, si $i\in I$, on utilise plutôt la notation indicielle x_i plutôt que la notation $x(i)$ pour désigner l'image de i par l'application x.
- Dans le cas particulier où I peut être muni d'un ordre (par exemple dans le cas où I est un ensemble de la forme $[\![1\,;n]\!]$ avec $n\in\mathbb{N}^*$, ou $I=\mathbb{N}$), il est naturel de ranger les éléments de cette famille dans l'ordre de leur indice. Par exemple, si $n\in\mathbb{N}^*$ et $I=[\![1\,;n]\!]$, on notera une famille d'éléments de I sous la forme $(x_1,\dots,x_n)$. Le dessin ci-dessous permet de bien visualiser une famille en tant que fonction :

$$\begin{tikzcd} 1 \arrow[d] & 2 \arrow[d] & \cdots & n \arrow[d] \\ x_1 & x_2 & \cdots & x_n \end{tikzcd}$$

- Une suite d'éléments de E est une famille indexée par $\mathbb{N}$, ou éventuellement par un ensemble de la forme $\{n\in\mathbb{N}\,|\,n\geqslant n_0\}$. Là aussi, on préfère utiliser la notation x_n plutôt que $x(n)$ pour expliciter l'image de n par l'application x. Là aussi, le dessin ci-dessous permet de bien visualiser une suite en tant que fonction :

$$\begin{tikzcd} 0 \arrow[d] & 1 \arrow[d] & 2 \arrow[d] & 3 \arrow[d] & 4 \arrow[d] & \cdots \\ x_0 & x_1 & x_2 & x_3 & x_4 & \cdots \end{tikzcd}$$

- La famille $(0,0,7)$ est une sous-famille de $(-1,0,0,1,7)$, et donc la famille $(-1,0,0,1,7)$ est une sur-famille de la famille $(0,0,7)$.
- La concaténation des familles $(1,2,3)$ et $(3,4,5)$ est la famille $(1,2,3,3,4,5)$.
- Une famille à deux éléments (i.e. indexée par $[\![1\,;2]\!]$) est appelée un couple, une famille à 3 éléments un triplet etc. Si $n\geqslant 1$, une famille à n éléments (i.e. indexée par $[\![1\,;n]\!]$) est appelée un n-uplet. Un couple est noté (x_1,x_2), un triplet (x_1,x_2,x_3), un n-uplet $(x_1,\dots,x_n)$ etc.

Un couple est aussi noté (x,y) ou (a,b) ou... et un triplet est aussi noté (x,y,z) ou ... Il suffit de mettre le bon nombre d'éléments **dans le bon ordre** entre parenthèses.

- Si $(x_i)_{i\in I}$ est une famille et si $i\in I$, on dit que la famille contient l'élément x_i. Par abus de langage (les familles ne sont pas des ensembles, voir certaines différences ci-dessous), on dit parfois que x_i est un élément de la famille ou appartient à la famille. De même, si $(x_j)_{j\in J}$ est une sous-famille de $(x_i)_{i\in I}$, on dit qu'elle est contenue dans la famille $(x_i)_{i\in I}$, et par abus de langage on dit parfois qu'elle est incluse dedans. On s'autorisera donc parfois également à parler d'appartenance à une famille ou de parler d'intersection ou d'union de familles (au lieu d'ensembles), ces abus de langage se comprenant facilement. Par exemple, on dira parfois que la famille $(0,0,7)$ est incluse dans la famille $(0,0,7,8,9)$ ou que l'union de $(0,0,7)$ et de $(8,9)$ est la famille $(0,0,7,8,9)$, mais cela restera marginal (on préférera parler de sous-famille et de concaténation, cf. chapitres 28 et 30).

- Une famille $(x_1, \dots, x_n)$ est une famille à n éléments, même si $x_1, \dots, x_n$ ne sont pas distincts ! Par exemple, la famille $(0, 0, 7)$ est une famille à trois éléments (alors que l'ensemble $\{0; 0; 7\} = \{0; 7\}$ est un ensemble à deux éléments). On parlera parfois également de famille de cardinal n, même si on parlera plus souvent de famille à n éléments.
- Deux fonctions étant égales si et seulement si elles coïncident en tout point de leur ensemble de définition, deux familles $(x_i)_{i \in I}$ et $(y_i)_{i \in I}$ indexées par le même ensemble I sont égales si et seulement si $x_i = y_i$ pour tout $i \in I$. En particulier, deux familles à n éléments $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ sont égales si et seulement si $x_1 = y_1, \dots, x_n = y_n$.
- Contrairement aux ensembles, il y a une notion d'ordre dans une famille, et une famille peut contenir plusieurs fois le même élément (même si on parle d'une famille à trois éléments, cela ne signifie pas que les trois éléments soient distincts). Par exemple, les deux familles à trois éléments $(0, 0, 7)$ et $(0, 7, 0)$ sont distinctes car correspondent aux deux fonctions ci-dessous :

$$\begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 0 & 0 & 7 \end{array} \qquad\qquad \begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 0 & 7 & 0 \end{array}$$

> En pratique, comme dit dans la définition, on se contente d'écrire $(0, 0, 7)$ et $(0, 7, 0)$, la fonction sous-jacente est sous-entendue (et ne sera jamais explicitée en pratique, voir ci-contre). Avec les mains : se donner un triplet de réels revient à se donner trois réels dans un certain ordre.

- En pratique, on utilise très peu la définition ci-dessus, c'est-à-dire qu'on utilise très peu le fait qu'une famille est une fonction. On se contente d'une approche intuitive : se donner une famille d'éléments de E indexée par un ensemble I revient à se donner des éléments de E, un par élément i de I, c'est-à-dire qu'à chaque élément i de I on associe un élément x_i de E. Par exemple, se donner un couple d'éléments de E revient à se donner un premier élément de E puis un deuxième élément de E dans cet ordre. Ne jamais oublier que l'ordre compte ! Pour s'en souvenir, il suffit de voir que, dans le dessin ci-dessous, le couple $(1, 2)$ et le couple $(2, 1)$ ne sont pas représentés par le même point.

> À chaque élément de I correspond un et un seul élément de E : c'est la définition d'une fonction !

Remarque : Tous les éléments d'une famille ne sont pas forcément des éléments du même ensemble. On peut par exemple définir un couple, i.e. une famille à deux éléments, (r, θ) avec $r \in \mathbb{R}_+^*$ et $\theta \in \left]-\pi\,;\pi\right]$. Pour noter les ensembles de familles, on introduit la notion de produit cartésien d'ensembles.

III.1.b Produit cartésien de 2 ensembles.

Définition. On appelle produit cartésien de E et F, et on note $E \times F$, l'ensemble des couples (e, f) d'éléments de E et F. En d'autres termes :

$$E \times F = \{(x, y) \,|\, x \in E \text{ et } y \in F\}.$$

Si $E = F$, on le note E^2 au lieu de $E \times E$.

> On rappelle que, dans ce chapitre, E et F sont deux ensembles non vides.

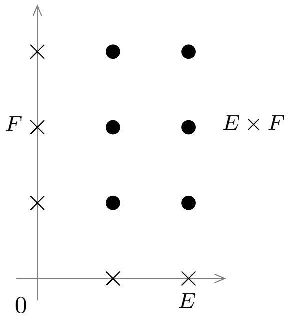
Exemples :

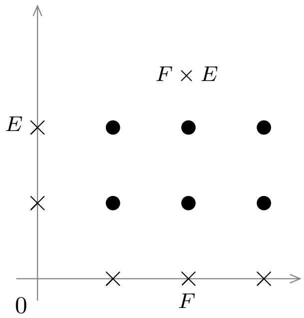
- Si $E = \{1; 2\}$ et $F = \{1; 2; 3\}$, alors :
  - ⋆ $E \times F = \big\{(1, 1); (1, 2); (1, 3); (2, 1); (2, 2); (2, 3)\big\}$.
  - ⋆ $F \times E = \big\{(1, 1); (1, 2); (2, 1); (2, 2); (3, 1); (3, 2)\big\}$.

> Comme on le voit ci-contre, $E \times F \neq F \times E$ en général !

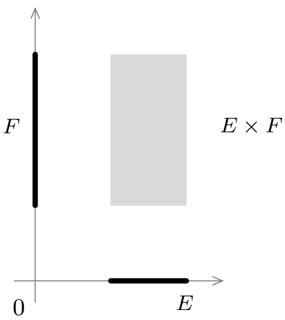
- Le produit cartésien $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ est l'ensemble des couples de réels. On l'identifie parfois au plan (en identifiant un point à ses cordonnées dans un repère orthonormé).
- $\llbracket 1\,;6 \rrbracket^2$ est l'univers que l'on associera à l'expérience aléatoire consistant à lancer deux dés à 6 faces (consécutivement ou en même temps).

Remarque : Si E et F sont des parties de $\mathbb{R}$, on peut voir le produit $E \times F$ comme l'ensemble de tous les points d'abscisse appartenant à E et d'ordonnée appartenant à F. Ci-dessous, à gauche, nous avons représenté $E \times F$, et à droite, $F \times E$:





Autre exemple : si $E = [1\,;2]$ et $F = [1\,;3]$ (attention à ne pas les confondre avec $\{1;2\}$ et $\{1;2;3\}$), alors on peut voir $E \times F$ comme le rectangle du plan ci-dessous :



On peut généraliser cette notion : si $E = [0\,;1]$ et si $F = [0\,;1]^2$ est le carré du plan formé par les points $(0,0),(0,1),(1,1)$ et $(1,0)$, alors $E \times F$ est le cube $[0\,;1]^3$ formé par les huit points dont les trois coordonnées sont 0 ou 1 (voir le dessin ci-contre). Au delà, c'est toujours possible, mais cela devient difficile à visualiser !

Remarque : Comme on peut le voir sur les dessins ci-dessus (en entourant la première colonne pour le dessin de gauche par exemple), pour chaque élément x de E, on dispose d'une copie complète de F (l'ensemble $\{x\} \times F$, c'est-à-dire l'ensemble $\{(x,y)\,|\,y \in F\}$, dont on peut montrer qu'il est en bijection, voir ci-dessous, avec F).

III.1.c Produit cartésien de n ensembles.

Définition. Soit n un entier naturel supérieur à 2. Soient $E_1, \dots, E_n$ des ensembles non vides. On appelle produit cartésien de $E_1, \dots, E_n$, et on note $E_1 \times \cdots \times E_n$, l'ensemble des n-uplets d'éléments de $E_1, \dots, E_n$. En d'autres termes :

$$E_1 \times \cdots \times E_n = \{(x_1, \dots, x_n)\,|\,\forall i \in [\![1\,;\,n]\!], x_i \in E_i\}$$

Si $E_1 = E_2 = \cdots = E_n = E$, on le note E^n au lieu de $E_1 \times \cdots \times E_n$.

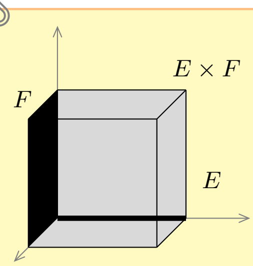
Exemple : Si $n \in \mathbb{N}^*$, on note donc $\mathbb{R}^n$ l'ensemble des n-uplets de réels. En particulier, $\mathbb{R}^3$ est l'ensemble des triplets de réels, qu'on identifiera parfois à l'espace, de la même façon qu'on a identifié $\mathbb{R}^2$ au plan.

Remarque : La notion de produit cartésien permet aussi un raccourci de notations. Par exemple on pourra écrire « $\forall (x,y,z) \in E^2 \times F$ » au lieu de « $\forall x \in E, \forall y \in E, \forall z \in F$ » . Idem avec les « $\exists$ ».

III.1.d Cas général (HP) et cohérence de la notation F^E.

Définition. Soit I un ensemble non vide et soit $(E_i)_{i \in I}$ une famille d'ensembles. On appelle produit cartésien des E_i, et on note $\displaystyle\prod_{i \in I} E_i$, l'ensemble des familles $(x_i)_{i \in I}$ telles que, pour tout $i \in I$, $x_i \in E_i$. En d'autres termes :

La représentation graphique ci-contre peut être utile pour bien comprendre que $E \times F$ n'est pas un ensemble du même type que E ou F : par exemple, si E et F sont des parties de $\mathbb{R}$, alors $E \times F$ est une partie de $\mathbb{R}^2$, c'est-à-dire une partie du plan. Contrairement aux unions, intersections etc. vues dans la suite, le produit cartésien est une opération « externe » : une intersection de parties de E est une partie de E, mais E^2 n'est pas une partie de E (par exemple, un produit d'intervalles est un rectangle du plan) ! De plus, cette représentation aide également à comprendre qu'en général, $E \times F \neq F \times E$, et qu'un couple (a,b) n'est pas égal en général au couple (b,a).



Ne pas confondre $(x_1, \dots, x_n)$ avec l'ensemble $\{x_1; \dots; x_n\}$. Par exemple le triplet $(x_1, x_2, x_3) = (0,0,7)$ n'est pas l'ensemble $\{x_1; x_2; x_3\} = \{0;7\}$. Encore une fois, dans une famille, l'ordre compte et on peut avoir plusieurs fois le même élément, ce qui n'est pas le cas avec les ensembles.

$$\prod_{i \in I} E_i = \{(x_i)_{i \in I} \,|\, \forall i \in I, x_i \in E_i\}$$

Si tous les E_i sont égaux à E, on note le produit E^I.

Cette définition est une généralisation des précédentes. En effet, si $I = [\![1\,;\,n]\!]$, alors $\displaystyle\prod_{i \in I} E_i = \prod_{i=1}^{n} E_i = E_1 \times \cdots \times E_n$.

Remarque : Par définition, E^I est donc l'ensemble des fonctions de I dans E. On a déjà vu que l'ensemble des fonctions de E dans F est noté F^E, donc cette notation n'est pas une nouveauté, mais expliquons pourquoi ce n'est pas qu'une notation, pourquoi elle se comprend très bien intuitivement.

Un élément de $E_1 \times \cdots \times E_n$ est la donnée d'un élément de $E_1, \dots,$ d'un élément de E_n. Quand on fait du dénombrement, cela s'appelle le principe multiplicatif (cf. chapitre 17) : j'ai deux sacs, l'un avec 10 livres, et l'autre avec 5 livres, je prend un livre dans chaque sac, j'ai 10×5 choix possibles, 10 choix pour le premier sac, 5 choix pour le deuxième. Ici, c'est la même chose : avec les mains (et de façon trèèèèèèès peu rigoureuse, les E_i sont des ensembles, pas des nombres, et dans la suite, E peut être infini par exemple), on a $E_1 \times \cdots \times E_n$ choix pour une famille $(x_1, \dots, x_n)$: E_1 choix pour x_1, …, E_n choix pour x_n

Pour une fonction de E dans F, c'est la même chose : on choisit un élément de F pour chaque élément de E, c'est-à-dire (encore une fois, ce n'est pas rigoureux) qu'on a $\underbrace{F \times F \times \cdots \times F}_{E \text{ fois}}$ façons de définir une fonction de E dans F, pour chaque élément de E (on fait donc l'opération E fois) on choisit un élément de F (on choisit un élément du sac).

Cela se voit très bien avec le premier dessin de $E \times F$: 3 choix pour la première image, 3 fois pour la deuxième, donc $9 = 3^2$ choix en tout, ce qui est cohérent avec la notation F^E puisque F a $\text{card}(F) = 3$ et $\text{card}(E) = 2$. Autre exemple, pour définir une suite, on choisit un élément de $\mathbb{R}$ pour x_0, un élément de $\mathbb{R}$ pour x_1 et ainsi de suite : on a $\underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{\mathbb{N} \text{ fois}}$ choix possibles pour définir une suite, ce qui légitime la notation $\mathbb{R}^{\mathbb{N}}$. Bref, on note F^E l'ensemble des fonctions de E dans F, et cette notation se comprend bien !

Ce n'est pas du tout rigoureux ! Par exemple, pour une suite, $\mathbb{N}$ est infini !

III.1.e Graphe d'une fonction.

Définition. Soit $f : E \to F$ une fonction. On appelle graphe de f la partie G de $E \times F$ suivante :

$$G = \{(x, y) \in E \times F \,|\, y = f(x)\}$$

Quand on fait de la théorie des ensembles pure et dure, on **définit** une fonction par son graphe. Mais cela nécessite que $E \times F$ ait été défini auparavant, donc cela nécessite de donner une autre définition d'un couple... Bref, la théorie des ensembles, c'est trop bien !

Remarque : En d'autres termes, si $(x, y) \in E \times F$, alors : $(x, y) \in G \iff y = f(x)$.

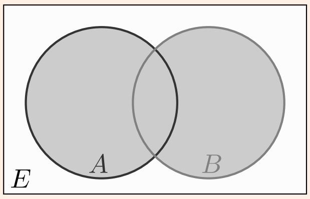
Remarque : Dans le cas où E est une partie de $\mathbb{R}$ et où $F = \mathbb{R}$, G est donc la partie de $\mathbb{R}^2$ (qu'on a identifié au plan) formée par les points dont l'ordonnée y est l'image par f de l'abscisse x.

III.2 Opérations sur les parties d'un ensemble

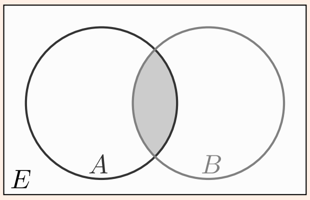
Dans toute la suite, on rappelle que E désigne un ensemble non vide.

III.2.a Union et intersection

Définition. Soient A et B deux parties de E. On définit les parties de E :



$A \cup B = \{ x \in E \,|\, x \in A \text{ ou } x \in B \}$, l'union (ou réunion) de A et B.



$A \cap B = \{ x \in E \,|\, x \in A \text{ et } x \in B \}$, l'intersection de A et B.

En d'autres termes, un élément appartient à l'intersection quand il appartient à tous les ensembles, et un élément appartient à l'union quand il appartient à au moins un des ensembles. Ce sera la même chose avec des intersections de plus de deux ensembles (voir ci-dessous).

Remarque : En d'autres termes, si $x \in E$:

$$x \in A \cup B \iff (x \in A) \text{ ou } (x \in B), \qquad x \in A \cap B \iff (x \in A) \text{ et } (x \in B).$$

Exemple : $[\,0\,;2\,] \cap [\,1\,;3\,] = [\,1\,;2\,]$ et $[\,0\,;2\,] \cup [\,1\,;3\,] = [\,0\,;3\,]$.

Définition. Deux parties A et B de E sont dites disjointes si $A \cap B = \varnothing$. Dans ce cas, on dit que $A \cup B$ est une union disjointe.

On trouve parfois, le cas échéant, les notations suivantes pour désigner l'union disjointe de A et B (c'est-à-dire pour désigner l'union de A et de B et pour signifier en même temps qu'elle est disjointe) : $A \sqcup B$, $A \amalg B$ ou $A \uplus B$. Ces notations n'étant pas universelles (la preuve : il y en a plusieurs !), on ne les utilisera pas, on se fatiguera à écrire en toutes lettres que l'union est disjointe...

Remarque : Attention de ne pas confondre distincts et disjoints :

- A et B sont distincts (c'est-à-dire $A \neq B$) si et seulement s'ils n'ont pas les mêmes éléments, c'est-à-dire s'il existe un élément dans un ensemble qui n'est pas dans l'autre. Avec des quantificateurs :

$$(\exists x \in A,\, x \notin B) \text{ ou } (\exists x \in B,\, x \notin A).$$

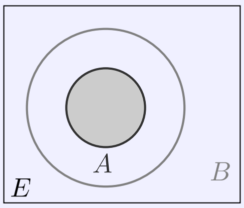
- A et B sont disjoints si et seulement si leur intersection est vide, c'est-à-dire s'ils n'ont aucun élément en commun. Avec des quantificateurs :

$$(\forall x \in A,\, x \notin B) \text{ et } (\forall x \in B,\, x \notin A).$$

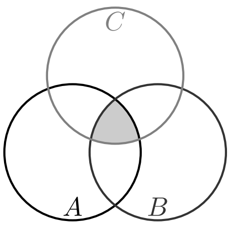
Deux ensembles disjoints sont évidemment distincts (enfin, sauf s'ils sont tous les deux vides...), mais la réciproque est fausse : $[\,0\,;2\,]$ et $[\,1\,;3\,]$ sont distincts mais pas disjoints.

Proposition. Soient A, B et C des parties de E.

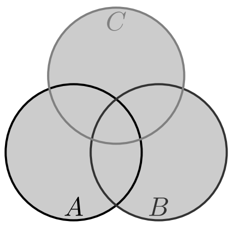
- (commutativité) $A \cup B = B \cup A$ et $A \cap B = B \cap A$,
- (associativité) $A \cup (B \cup C) = (A \cup B) \cup C$ et $A \cap (B \cap C) = (A \cap B) \cap C$,
- (distributivité) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ et $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
- $(A \cap B) \subset A \subset (A \cup B)$ et $(A \cap B) \subset B \subset (A \cup B)$.
- $A \cap A = A$ et $A \cup A = A$,
- $\varnothing \cap A = \varnothing$ et $\varnothing \cup A = A$,
- $A \cap B = A \iff A \subset B$ (voir dessin ci-contre)
- $E \cap A = A$ et $E \cup A = E$.



$\rightsquigarrow$ Exercice.



$A \cap B \cap C$



$A \cup B \cup C$

La distributivité pose souvent problème : comment ne pas se tromper ? Pour développer (de gauche vers droite) c'est facile, on peut utiliser des flèches au brouillon. Mais pour factoriser (de droite vers gauche) ? En faisant l'analogie avec deux lois, dont l'une est distributive par rapport à l'autre : la multiplication et l'addition. Par exemple (cf. exercice 9), si on veut donner une expression plus simple de

$(A \cup B \cup C) \cap (A \cup B \cup \overline{C})$

...

Remarque : L'associativité de la réunion (resp. de l'intersection) nous permet d'écrire : $A\cup B\cup C$ (respectivement $A\cap B\cap C$) au lieu de $A\cup(B\cup C)$ (respectivement $A\cap(B\cap C)$). En effet, on peut définir de façon analogue l'union ou l'intersection de plus de deux ensembles :

... alors on écrit l'analogue avec les lois + et $\times$ au brouillon :

$$(a\times b\times c)+(a\times b\times \tilde{c})$$

ce qu'on factorise en $a\times((b\times c)+(b\times\tilde{c}))$. Finalement, l'ensemble ci-dessus est égal, par distributivité de l'union sur l'intersection, à

$$A\cup((B\cup C)\cap(B\cup\overline{C}))$$

Attention, ce n'est qu'une analogie !

Définition. Soit $(A_i)_{i\in I}$ une famille de parties de E (i.e. une famille d'éléments de $\mathscr{P}(E)$) indexée par un ensemble non vide I. On définit :

- $\displaystyle\bigcup_{i\in I} A_i=\{\,x\in E\,|\,\exists i\in I,\,x\in A_i\,\}$, l'union de la famille $(A_i)_{i\in I}$.
- $\displaystyle\bigcap_{i\in I} A_i=\{\,x\in E\,|\,\forall i\in I,\,x\in A_i\,\}$, l'intersection de la famille $(A_i)_{i\in I}$.

Remarques :

- (Écriture avec des quantificateurs) Soit $x\in E$.

$$x\in\bigcup_{i\in I}A_i \iff \exists i\in I, x\in A_i \qquad\text{et}\qquad x\in\bigcap_{i\in I}A_i\iff\forall i\in I,x\in A_i$$

  ⚠ Là aussi, un élément appartient à l'intersection quand il appartient à tous les ensembles, et un élément appartient à l'union quand il appartient à au moins un des ensembles.

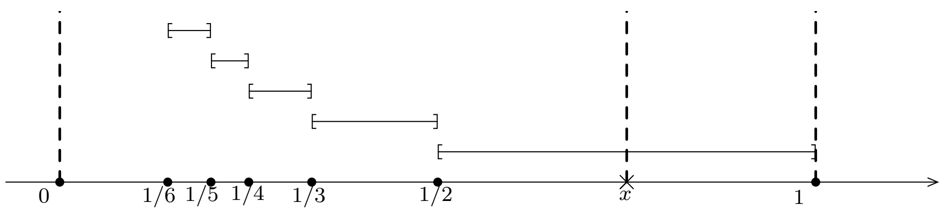
- Si $I=[\![\,p\,;\,n\,]\!]$ (l'ensemble des entiers compris entre p et n, cf. chapitre) avec p et n des entiers naturels tels que $p\leqslant n$, alors on note $\displaystyle\bigcup_{i=p}^{n}$ au lieu de $\displaystyle\bigcup_{i\in I}$ (de même pour l'intersection). De même, si $I=\mathbb{N}^*$, on note $\displaystyle\bigcup_{i=1}^{+\infty}$ au lieu de $\displaystyle\bigcup_{i\in I}$.

On peut évidemment généraliser à $\mathbb{N}$ (on a alors l'union de 0 à $+\infty$), à $\{n\in\mathbb{N}\,|\,n\geqslant 2\}$ (on a alors l'union de 2 à $+\infty$), etc.

- On dit que l'union $\displaystyle\bigcup_{i\in I}A_i$ est disjointe si les A_i sont deux à deux disjoints, c'est-à-dire si : $\forall(i,j)\in I^2, i\neq j\Rightarrow A_i\cap A_j=\varnothing$.

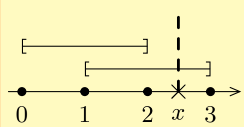
Exemples :

- Nous verrons dans le chapitre 5 que :
  - ⋆ tan est définie sur $\mathbb{R}\backslash\left(\dfrac{\pi}{2}+\pi\mathbb{Z}\right)=\displaystyle\bigcup_{k\in\mathbb{Z}}\left]-\dfrac{\pi}{2}+k\pi\,;\dfrac{\pi}{2}+k\pi\right[$.
  - ⋆ $\ln\circ\sin$ est définie sur $\{x\in\mathbb{R}\,|\,\sin(x)>0\}=\displaystyle\bigcup_{k\in\mathbb{Z}}\,]\,2k\pi\,;(2k+1)\pi\,[$.
- Montrons que $\displaystyle\bigcup_{k=1}^{+\infty}\left[\frac{1}{k+1}\,;\frac{1}{k}\right]=]\,0\,;1\,]$ et que $\displaystyle\bigcap_{k=1}^{+\infty}\left[\frac{1}{k+1}\,;\frac{1}{k}\right]=\varnothing$. Cela se voit très bien sur le dessin suivant :



  Nous avons représenté les intervalles de la forme $\left[\dfrac{1}{k+1}\,;\dfrac{1}{k}\right]$ les uns au-dessus des autres pour des raisons de lisibilité. Rappelons (il est fondamental de comprendre ce point) qu'un élément x est dans l'union quand il appartient à **au moins** l'un des ensembles, et dans l'intersection quand il est dans **tous** les ensembles. On voit sur le dessin (mais nous allons le montrer rigoureusement ci-dessous) qu'il n'y a aucun réel x qui soit dans tous les intervalles, et que tout réel dans $]\,0\,;1\,]$ (0 est ouvert car les intervalles se rapprochent de 0 sans jamais l'atteindre) est dans au moins un intervalle.

Faisons un dessin analogue pour illustrer le premier exemple de ce paragraphe, à savoir l'union et l'intersection de $[\,0\,;2\,]$ et $[\,1\,;3\,]$:



- $\star$ Tout d'abord, $\left[\frac{1}{4}\,;\frac{1}{3}\right] \cap \left[\frac{1}{2}\,;1\right] = \varnothing$: il n'existe aucun réel qui soit dans ces deux intervalles et donc, a fortiori, il n'existe aucun réel qui soit dans tous les intervalles, c'est-à-dire que l'intersection est vide.
- $\star$ Montrons à présent que l'union (que l'on note U) est égale à $]\,0\,;1\,]$ par double inclusion. Soit $x \in U$. Il existe $k \in \mathbb{N}^*$ tel que $x \in \left[\frac{1}{k+1}\,;\frac{1}{k}\right]$. Or, $0 < \frac{1}{k+1} \leqslant \frac{1}{k} \leqslant 1$ et donc $x \in \,]\,0\,;1\,[$. Ainsi $U \subset \,]\,0\,;1\,]$.
- $\star$ Soit $x \in \,]\,0\,;1\,]$. Alors $\frac{1}{x} \geqslant 1$ et donc $\ell = \left\lfloor \frac{1}{x} \right\rfloor \in \mathbb{N}^*$. Comme $\ell \leqslant \frac{1}{x} < \ell + 1$, nous avons $\frac{1}{\ell+1} < x \leqslant \frac{1}{\ell}$. Par conséquent $x \in \left[\frac{1}{\ell+1}\,;\frac{1}{\ell}\right]$: x appartient à l'un des ensembles donc il appartient à leur union, c'est-à-dire que $x \in U$. D'où l'inclusion réciproque, d'où l'égalité.

- De façon très générale, $E = \displaystyle\bigcup_{x\in E} \{x\}$, et cette union est disjointe.

Avec les mains : un ensemble est l'union disjointe de ses éléments.

Proposition (distributivité). Soient B une partie de E et $(A_i)_{i\in I}$ une famille de parties de E. Nous avons :

- $\left(\displaystyle\bigcup_{i\in I} A_i\right) \cap B = \displaystyle\bigcup_{i\in I} (A_i \cap B)$.
  On dit que l'intersection est distributive par rapport à l'union.
- $\left(\displaystyle\bigcap_{i\in I} A_i\right) \cup B = \displaystyle\bigcap_{i\in I} (A_i \cup B)$.
  On dit que l'union est distributive par rapport à l'intersection.

Démonstration. Montrons le premier point (l'autre est analogue). Plutôt que de raisonner par double inclusion, raisonnons pour une fois par équivalences. Soit $x \in E$.

$$\begin{aligned} x \in \left(\bigcup_{i\in I} A_i\right) \cap B & \iff (\exists i \in I,\, x \in A_i) \text{ et } (x \in B) \\ & \iff \exists i \in I,\, (x \in A_i \text{ et } x \in B) \\ & \iff \exists i \in I,\, x \in A_i \cap B \\ & \iff x \in \bigcup_{i\in I} (A_i \cap B). \end{aligned}$$

Un élément appartient au premier ensemble si et seulement s'il appartient au deuxième : il en découle que ces deux ensembles ont les mêmes éléments, donc qu'ils sont égaux. $\square$

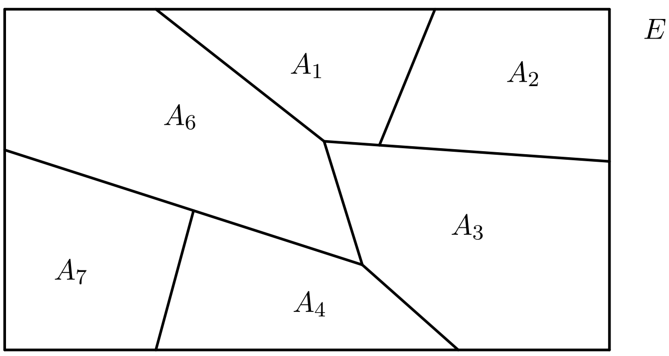
Comme expliqué dans le chapitre 1, raisonner par équivalences est plus rapide mais n'est pas toujours possible, ce n'est qu'avec de l'expérience qu'on apprend à repérer les cas où c'est possible. Dans le doute, raisonner par double inclusion.

Définition. Soit $(A_i)_{i\in I}$ une famille de parties **non vides** de E indexée par un ensemble non vide I. On dit que $(A_i)_{i\in I}$ est une partition de E si :

- Les A_i sont deux à deux disjoints.
- $\displaystyle\bigcup_{i\in I} A_i = E$.

En d'autres termes, une partition de E est une famille de parties non vides de E dont l'union disjointe est égale à E.

Le côté non vide est indispensable. Si on s'autorise à prendre des parties vides, on parle plutôt de partage (ou de système complet d'événements, quand on fait des probabilités, cf. chapitre 26).



Exemples :

- $\mathbb{R}_+^*$, $\mathbb{R}_-^*$ et $\{0\}$ forment une partition de $\mathbb{R}$.
- L'ensemble des nombres pairs et l'ensemble des nombres impairs forment une partition de $\mathbb{Z}$. Plus généralement, si $n \geqslant 2$, si on note $\overline{0}$ l'ensemble des entiers congrus à 0 modulo n, $\overline{1}$ l'ensemble des entiers congrus à 1 modulo n, ..., $\overline{n-1}$ l'ensemble des entiers congrus à $n-1$ modulo n, alors $\overline{0}, \overline{1}, \dots, \overline{n-1}$ est une partition de $\mathbb{Z}$.

Cette partition est la partition associée à la relation d'équivalence « être congru modulo n », cf. chapitre 16. Cela permettra de définir $\mathbb{Z}/n\mathbb{Z}$, cf. chapitres 16 et 18.

Définition. Soit $(A_i)_{i\in I}$ une famille de parties non vides de E indexée par un ensemble non vide I. Soit F une partie de E. On dit que $(A_i)_{i\in I}$ est un recouvrement de F si $F \subset \displaystyle\bigcup_{i\in I} A_i$. On dit que c'est un recouvrement disjoint si les A_i sont de plus deux à deux disjoints.

Exemples :

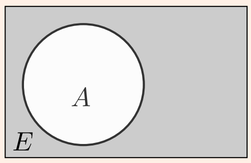
- La famille $(\llbracket 0\,;\,n\rrbracket)_{n\in\mathbb{N}}$ est un recouvrement (non disjoint !) de $\mathbb{N}$.
- $\left]-\dfrac{1}{2}\,;\dfrac{3}{2}\right[$ et $\left]\dfrac{3}{2}\,;\dfrac{7}{2}\right[$ forment un recouvrement disjoint de $[\,0\,;1\,]\cup[\,2\,;3\,]$.

Remarque : Si $(A_i)_{i\in I}$ est un recouvrement disjoint de F, alors $(A_i\cap F)_{i\in I}$ est une partition de F.

III.2.b Complémentaire

Définition. Soit A une partie de E.

L'ensemble $\overline{A} = \{x \in E \,|\, x \notin A\}$ est une partie de E appelée complémentaire de A dans E.



On trouve également les notations A^c ou $E\backslash A$ (voir ci-dessous). En probabilités, on utilisera surtout la notation $\overline{A}$, mais en topologie (cf. deuxième année), $\overline{A}$ désignera l'adhérence de A, et donc on privilégiera la notation A^c.

Remarque : Pour tout $x\in E$, nous avons donc

$$x \in \overline{A} \iff x \notin A \iff \text{non}(x\in A).$$

Proposition. Soient A et B des parties de E. On a

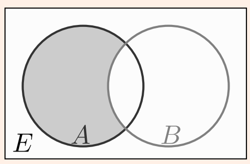
$$\overline{\varnothing} = E, \qquad \overline{E} = \varnothing, \qquad \overline{\overline{A}} = A, \qquad A\subset B \iff \overline{B}\subset\overline{A}.$$

⇝ Exercice.

Définition (différence). Soient A et B deux parties de E. L'ensemble

$$A\backslash B = A\cap\overline{B} = \{\,x\in E\,|\,x\in A \text{ et } x\notin B\,\}$$

est une partie de E appelée différence de A et B (et on note aussi parfois $A-B$). On l'appelle aussi « A privé de B ».



III.2.c Liens entre les différentes opérations : lois de Morgan

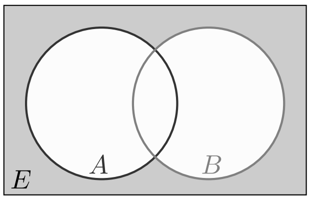
Proposition (lois de Morgan). Si A et B sont des parties de E, alors

$$\overline{A\cap B} = \overline{A}\cup\overline{B} \qquad \text{et} \qquad \overline{A\cup B} = \overline{A}\cap\overline{B}.$$

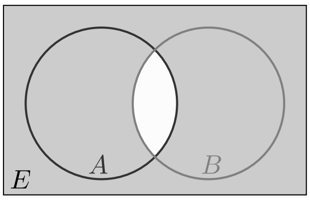
Plus généralement, si $(A_i)_{i\in I}$ est une famille de parties de E indexée par un ensemble non vide I, alors

$$\overline{\bigcap_{i\in I} A_i} = \bigcup_{i\in I} \overline{A_i} \qquad \text{et} \qquad \overline{\bigcup_{i\in I} A_i} = \bigcap_{i\in I} \overline{A_i}.$$

Nous avons déjà vu des lois de Morgan dans le chapitre 0. Ce sont en fait les mêmes ! Il y a en effet un lien très fort entre les opérations ensemblistes et les connecteurs logiques : le « et » logique est l'analogue de l'intersection, le « ou » est l'analogue de l'union, le « non » est l'analogue du passage au complémentaire, l'implication logique est l'analogue de l'inclusion, et l'équivalence logique est l'analogue de l'égalité (ensembliste). On voit alors que la loi de Morgan non(A et B) $\Longleftrightarrow$ (non(A)) ou (non(B)) est « la même » que la loi de Morgan $\overline{A\cap B} = \overline{A}\cup\overline{B}$.



$\overline{A\cup B} = \overline{A}\cap\overline{B}$



$\overline{A\cap B} = \overline{A}\cup\overline{B}$

Démonstration. Montrons la seconde égalité par équivalences (la première est laissée en exo) :

$$\begin{aligned} x \in \overline{\bigcup_{i\in I} A_i} &\Longleftrightarrow x \notin \bigcup_{i\in I} A_i \\ &\Longleftrightarrow \text{non}\big(\exists i \in I,\, x \in A_i\big) \\ &\Longleftrightarrow \forall i \in I,\, x \notin A_i \\ &\Longleftrightarrow x \in \bigcap_{i\in I} \overline{A_i} \end{aligned}$$

□

III.2.d La différence symétrique (HP).

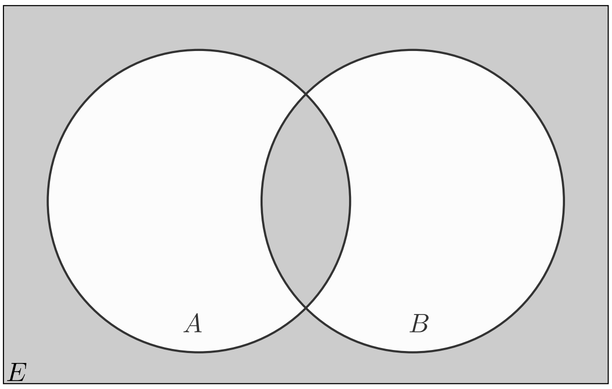
Définition. Soient A et B deux parties de E. L'ensemble :

$$A\Delta B = (A\cup B)\backslash(A\cap B)$$

est une partie de E appelée différence symétrique de A et B.

Remarque : En d'autres termes, $A\Delta B$ est l'ensemble des éléments de E qui appartiennent à A ou à B mais pas aux deux en même temps.

C'est donc le « ou » exclusif, le XOR en informatique.



$A\Delta B$

Proposition. Soient A, B, C trois parties de E.

- $A\Delta B = \big(A\cap\overline{B}\big)\cup\big(\overline{A}\cap B\big)$.
- $A\Delta A = \varnothing$.
- (commutativité) $A\Delta B = B\Delta A$.
- (associativité) $A\Delta(B\Delta C) = (A\Delta B)\Delta C$.

Démonstration.

- On a successivement

$$\begin{aligned} A\Delta B &= (A\cup B)\backslash(A\cap B) \\ &= (A\cup B)\cap\overline{(A\cap B)} \\ &= (A\cup B)\cap\left(\overline{A}\cup\overline{B}\right) \\ &= \left[(A\cup B)\cap\overline{A}\right]\cup\left[(A\cup B)\cap\overline{B}\right] \\ &= \left[(A\cap\overline{A})\cup(B\cap\overline{A})\right]\cup\left[(A\cap\overline{B})\cup(B\cap\overline{B})\right] \\ &= \left[\varnothing\cup(B\cap\overline{A})\right]\cup\left[(A\cap\overline{B})\cup\varnothing\right] \end{aligned}$$

Loi de Morgan.

L'intersection est distributive par rapport à l'union.

ce qui permet de conclure (car $B\cap\overline{A}=\overline{A}\cap B$).

- $A\cup A=A\cap A=A$ donc $A\Delta A=A\backslash A=\varnothing$.
- Découle de la commutativité de l'union et de l'intersection.
- **Première méthode : avec des tables de vérité.** Soit $x\in E$. Montrons avec des tables de vérité que les propositions « $x\in A\Delta(B\Delta C)$ » et « $x\in(A\Delta B)\Delta C$ » sont équivalentes.

$x\in A$	$x\in B$	$x\in C$	$x\in A\Delta B$	$x\in(A\Delta B)\Delta C$	$x\in B\Delta C$	$x\in A\Delta(B\Delta C)$
F	F	F	F	F	F	F
F	F	V	F	V	V	V
F	V	F	V	V	V	V
F	V	V	V	F	F	F
V	F	F	V	V	F	V
V	F	V	V	F	V	F
V	V	F	F	F	V	F
V	V	V	F	V	F	V

Finalement : $x\in(A\Delta B)\Delta C \iff A\Delta(B\Delta C)$. En d'autres termes, $(A\Delta B)\Delta C$ et $A\Delta(B\Delta C)$ ont les mêmes éléments donc sont égaux.

Deuxième méthode : par double inclusion. Soit $x\in A\Delta(B\Delta C)$. Différencions les cas.

⋆ Premier cas : $x\in A$. Alors $x\notin B\Delta C$

— Premier sous-cas : $x\in B\cap C$. Alors $x\in B$, donc $x\notin A\Delta B$, et $x\in C$, donc $x\in(A\Delta B)\Delta C$.

— Deuxième sous-cas : $x\notin B\cup C$. Alors $x\notin B$, donc $x\in A\Delta B$, et $x\notin C$ donc $x\in(A\Delta B)\Delta C$.

⋆ Deuxième cas : $x\notin A$, si bien que $x\in B\Delta C$.

— Premier sous-cas : $x\in B$. Alors $x\in A\Delta B$, et $x\notin C$, donc $x\in(A\Delta B)\Delta C$.

— Deuxième sous-cas : $x\in C$ et $x\notin B$ donc $x\notin A\Delta B$ et $x\in(A\Delta B)\Delta C$.

Rappelons qu'un élément est dans la différence symétrique quand il est dans l'un mais pas dans l'autre. Par conséquent, puisque $x\in A\Delta(B\Delta C)$ et $x\in A$, il en découle que $x\notin B\Delta C$, si bien que $x\notin B\cup C$ ou $B\in B\cap C$.

Dans tous les cas, $x\in(A\Delta B)\Delta C$. En d'autres termes, $A\Delta(B\Delta C)\subset(A\Delta B)\Delta C$. Par symétrie des rôles (A,B,C jouent le même rôle), $C\Delta(B\Delta A)\subset(C\Delta B)\Delta A$ et la différence symétrique est commutative, donc $(A\Delta B)\Delta C\subset A\Delta(B\Delta C)$. En d'autres termes, on a l'inclusion réciproque, d'où l'égalité.

Nous verrons une troisième méthode avec les fonctions indicatrices dans l'exercice 41.

IV Injections, surjections, bijections

IV.1 Définitions

Définition. Soit $f : E \longrightarrow F$. On dit que

- f est injective ou est une injection si tout élément de F admet **au plus** un antécédent (par f).
- f est surjective ou est une surjection si tout élément de F admet **au moins** un antécédent (par f).
- f est bijective ou est une bijection si f est à la fois injective et surjective, c'est-à-dire si tout élément de F admet **un unique** antécédent (par f).

Proposition. La fonction $f : E \longrightarrow F$ est injective si et seulement si tout couple d'éléments distincts de E fournit deux images distinctes par f, i.e.

$$\forall (x_1, x_2) \in E^2, \qquad x_1 \neq x_2 \quad \Longrightarrow \quad f(x_1) \neq f(x_2).$$

En d'autres termes, une fonction est injective lorsqu'elle ne prend pas deux fois la même valeur. La négation est : « $\exists (x_1, x_2) \in E^2, x_1 \neq x_2$ et $f(x_1) = f(x_2)$ ». Ainsi, pour montrer qu'une application n'est pas injective, il suffit d'exhiber deux éléments distincts ayant la même image.

Démonstration. Supposons f injective. Soient x_1 et x_2 deux éléments distincts de E. Si $f(x_1) = f(x_2)$, alors cet élément de F a au moins deux antécédents (x_1 et x_2), ce qui est absurde car f est injective. Réciproquement, supposons que tout couple d'éléments distincts de E donne deux images distinctes par f, et prouvons que f est injective. Soit $y \in F$. Par hypothèse, y ne peut pas avoir deux antécédents distincts donc f est injective.$\square$

Du résultat ci-dessus et de la définition on déduit immédiatement le théorème suivant :

Théorème (écriture avec des quantificateurs). Soit $f : E \longrightarrow F$.

- f est injective si et seulement si : $\forall (x_1, x_2) \in E^2$, $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$ ou, par contraposée : $\forall (x_1, x_2) \in E^2$, $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$.
- f est surjective si et seulement si : $\forall y \in F$, $\exists x \in E$, $f(x) = y$.
- f est bijective si et seulement si : $\forall y \in F$, $\exists ! x \in E$, $f(x) = y$.

IV.2 Rédaction type.

Pour montrer qu'une application $f : E \longrightarrow F$ est injective :

- On fixe $(x_1, x_2) \in E^2$ (E étant l'ensemble de départ de f : on adapte bien sûr si la fonction est définie sur un autre ensemble), on suppose $x_1 \neq x_2$ et on montre que $f(x_1) \neq f(x_2)$. On attend donc une rédaction du type :

  Soit $(x_1, x_2) \in E^2$. Supposons que $x_1 \neq x_2$.
  $\vdots$
  Alors $f(x_1) \neq f(x_2)$. f est injective.

- Autre méthode : on fixe $(x_1, x_2) \in E^2$, on suppose $f(x_1) = f(x_2)$ et on montre que $x_1 = x_2$. On attend donc une rédaction du type :

  Soit $(x_1, x_2) \in E^2$. Supposons que $f(x_1) = f(x_2)$.
  $\vdots$
  Alors $x_1 = x_2$: f est injective.

Pour montrer qu'une application $f : E \longrightarrow F$ est surjective : On fixe $y \in F$ (F étant l'ensemble d'arrivée de f : on adapte bien sûr si la fonction est à valeurs dans un autre ensemble) et on trouve un élément $x \in E$ tel que $f(x) = y$. On attend donc une rédaction du type :

> Soit $y \in F$.
>
> $\vdots$
>
> $y = f(\dots)$.
>
> Alors $\dots$ est un antécédent de y : f est surjective.

Pour montrer qu'une application f est bijective :

- On montre qu'elle est injective puis surjective. De plus, l'antécédent de y trouvé lors de l'étape de surjectivité est alors unique et est égal à $f^{-1}(y)$ (cf. paragraphe VI).
- Autre méthode : on fixe $y \in F$ et $x \in E$ et on résout l'équation (d'inconnue $x \in E$) $y = f(x)$. Si on trouve une unique solution, alors f est bijective et $x = f^{-1}(y)$. L'inconvénient est qu'on ne sait pas en général résoudre cette équation. On n'utilisera donc cette méthode que lorsqu'il est demandé d'expliciter f^{-1}.
- Si f est définie sur un intervalle de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$, on peut chercher à appliquer le théorème de la bijection (cf. paragraphe VII).

Pour montrer qu'une application $f : E \longrightarrow F$ n'est pas injective, on exhibe deux éléments de E distincts **explicites** qui ont la même image.

Pour montrer qu'une application $f : E \longrightarrow F$ n'est pas surjective, on exhibe un élément de F **explicite** qui n'est pas atteint.

Pour montrer qu'une application n'est pas bijective : on prouve qu'elle n'est pas injective OU qu'elle n'est pas surjective.

> En effet, la négation de « f est injective et f est surjective » est « f n'est pas injective ou f n'est pas surjective. »

IV.3 Erreurs à ne PAS commettre (même en hommage à vos illustres prédécesseurs)

- **Erreur :** « $x = y \Rightarrow f(x) = f(y)$ donc f est injective ».

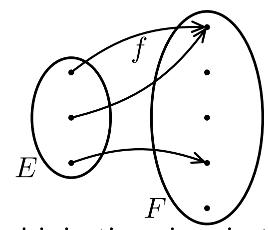
  Explication : Cette implication est vraie pour toute fonction f ! Deux éléments égaux ont TOUJOURS la même image, cela n'a rien à voir avec une quelconque injectivité !
- **Erreur :** « f est surjective si toute image (ou tout élément de l'image) admet au moins un antécédent ».

  Explication : Une image (ou un élément de l'image, cf. partie III) admet toujours un antécédent, par définition. Une fonction est surjective si tout élément **de l'ensemble d'arrivée** (ce qui n'est pas du tout la même chose) admet au moins un antécédent.
- **Erreur :** « f n'est pas injective donc est surjective ».

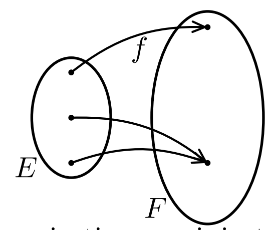
  Explication : Une fonction peut n'être ni injective ni surjective, voir le paragraphe suivant. Ça n'a rien à voir !
- **Erreur :** « f est injective quand elle admet au plus un antécédent ».

  Explication : Cette phrase n'a aucun sens. Ce sont les éléments de l'ensemble d'arrivée qui doivent admettre au plus un antécédent, cela n'a aucun sens de dire qu'une fonction a un antécédent.

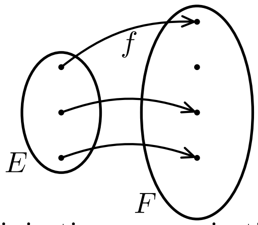
IV.4 Exemples patatoïdaux



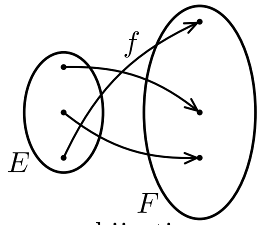
ni injective ni surjective
(oui, ça existe)



surjective non injective



injective non surjective



bijective

Encore une fois, une fonction est surjective quand tous les éléments de l'ensemble d'arrivée sont atteints (mais certains peuvent l'être deux fois), et est injective quand aucun n'est atteint deux fois (mais certains peuvent ne pas être atteints). Enfin, une application est bijective quand tous les éléments de l'ensemble d'arrivée sont atteints exactement une fois, ou admettent exactement un antécédent.

IV.5 Exemples plus explicites.

- L'application
$$f : \left\{ \begin{array}{ccc} (\mathbb{N}^*)^2 & \longrightarrow & \mathbb{R}_+^* \\ (n,p) & \longmapsto & \dfrac{n}{p} \end{array} \right.$$
n'est pas injective car $f(1,2) = 1/2 = f(2,4)$. Elle n'est pas non plus surjective car $\sqrt{2}$ est irrationnel : il n'existe donc pas d'entiers n et p tels que $\dfrac{n}{p} = \sqrt{2}$. Il n'existe pas d'entiers $(n,p) \in (\mathbb{N}^*)^2$ tels que $f(n,p) = \sqrt{2}$: $\sqrt{2}$ n'a pas d'antécédent par f.
- Notons $D(\mathbb{R},\mathbb{R})$ l'ensemble des fonctions dérivables sur $\mathbb{R}$ (cf. chapitre 2.4), et on rappelle que $\mathbb{R}^\mathbb{R}$ est l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. L'application
$$D : \left\{ \begin{array}{ccc} D(\mathbb{R},\mathbb{R}) & \longrightarrow & \mathbb{R}^\mathbb{R} \\ f & \longmapsto & f' \end{array} \right.$$
n'est pas injective car toutes les fonctions constantes ont la même image : la fonction nulle.
- L'application
$$f : \left\{ \begin{array}{ccc} \mathbb{R}_+^* \times \left]-\pi\,;\pi\right] & \longrightarrow & \mathbb{C}^* \\ (r,\theta) & \longmapsto & re^{i\theta} \end{array} \right.$$
est bijective (cf. chapitre 7).
- L'application
$$\varphi : \left\{ \begin{array}{ccc} \mathbb{R}^\mathbb{R} & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f(0) \end{array} \right.$$
n'est pas injective car $\varphi(\exp) = \varphi(\cos) = 1$. Cependant, φ est surjective. En effet, soit $y \in \mathbb{R}$. Soit f la fonction constante égale à y. Alors $f(0) = y$, d'où $\varphi(f) = y$: φ est surjective.
- La fonction $\exp : \mathbb{R} \longrightarrow \mathbb{R}$ est injective (car elle est strictement monotone, cf. paragraphe VI.1) non surjective (car 0 n'est pas atteint).
- $\exp : \mathbb{R} \longrightarrow \mathbb{R}_+^*$ est bijective.
- $\cos : \mathbb{R} \longrightarrow [-1\,;1]$ est surjective non injective.
- Introduisons
$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right. , \quad g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & x^2 \end{array} \right. , \quad h : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right. .$$
Alors f n'est ni injective ni surjective, g est surjective non injective, h est injective non surjective. Enfin l'application $\varphi : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & x^2 \end{array} \right.$ est bijective.
- L'application $\varphi : n \longmapsto n+1$ est une bijection de $\mathbb{N}$ dans $\mathbb{N}^*$:
  - ⋆ Soient $n_1 \neq n_2$ deux entiers naturels. Alors $n_1 + 1 \neq n_2 + 1$ c'est-à-dire que $\varphi(n_1) \neq \varphi(n_2)$: φ est injective.
  - ⋆ Soit $n \in \mathbb{N}^*$. Alors $n - 1 \in \mathbb{N}$ et $\varphi(n-1) = n$: $n-1$ est un antécédent de n donc φ est surjective.

D n'est pas non plus surjective mais ce n'est pas aussi simple. On peut montrer (cf. exercice 25 du chapitre 14) qu'il n'existe pas de fonction f telle que f' soit la fonction partie entière, et donc la partie entière n'a pas d'antécédent par D : D n'est par conséquent pas surjective. Montrer qu'une fonction est (ou n'est pas) surjective (ou injective) est parfois difficile !

La surjectivité de exp (en tant que fonction de $\mathbb{R}$ dans $\mathbb{R}_+^*$), de cos et de g découle du TVI (cf. chapitre 13) : le TVI peut être utile pour montrer la surjectivité d'une fonction. La fonction cos n'est pas injective car $\cos(0) = \cos(2\pi)$, f et g ne sont pas injectives car 1 et -1 ont la même image, f et h ne sont pas surjectives car -1 n'est pas atteint. Enfin, la dernière fonction φ est bijective d'après le théorème de la bijection (cf. paragraphe VII).

φ est injective et surjective donc est bijective. C'est le principe de *l'hôtel de Hilbert.*

On voit donc qu'il faut se méfier de son intuition quand on manie des ensembles infinis : on peut avoir un ensemble avec un élément de plus, mais qui est en bijection avec l'ensemble initial ! Nous montrerons dans l'exercice 44 que $\mathbb{N}$ et $\mathbb{Z}$ sont en bijection. Plus fort : nous montrerons qu'il existe une bijection de $\mathbb{N}$ dans $\mathbb{Q}$ (mais pas de $\mathbb{N}$ dans $\mathbb{R}$) au chapitre 17. Ce n'est pas parce qu'un ensemble a l'air « plus gros qu'un autre » qu'il n'y a pas de bijection entre les deux. Nous en reparlerons dans le chapitre 17.

- Si on a $E \subset F$, alors la fonction

$$i : \left\{ \begin{array}{ccc} E & \longrightarrow & F \\ x & \longmapsto & x \end{array} \right.$$

est injective : on l'appelle l'injection canonique de E dans F. Elle est surjective si et seulement si $E = F$. En particulier, Id_E est une bijection de E dans E.

- Examinons les exemples du paragraphe II :
  - ⋆ f_1 n'est pas surjective (il ne s'est rien passé en -20×10^9 car l'univers n'existait pas encore) ni injective (il s'est passé beaucoup de choses en 1789 par exemple).
  - ⋆ f_2 est surjective (si $(a,b) \in \mathbb{R}^2$, la fonction $x \longmapsto bx + a$ est un antécédent de (a,b) par f_2) mais pas injective (exp et $x \longmapsto 1 + x$ sont distinctes mais admettent $(1,1)$ pour image par f_2).
  - ⋆ f_3 n'est ni surjective (car les images des éléments de $\mathbb{R}^2$ par f_3 sont des suites géométriques et toute suite n'est pas géométrique) ni injective (car $f_3((0,1)) = f_4((0,2)) = (0)_{n \in \mathbb{N}}$).
  - ⋆ f_4
- Notons E l'ensemble des lauréats de la médaille Fields (jusqu'en 2022), P l'ensemble des prénoms, N l'ensemble des noms de famille.
  - ⋆ L'application $f : E \longrightarrow P$ qui à tout lauréat associe son prénom n'est ni injective (en effet f(Laurent Lafforgue) $= f$(Laurent Schwartz) $=$ Laurent) ni surjective (car ni Léon $\in P$ ni Marius $\in P$ n'admettent d'antécédent).
  - ⋆ L'application $g : E \longrightarrow N$ qui à tout lauréat associe son nom de famille est injective (si on se donne deux lauréats distincts, alors ils n'ont pas le même nom de famille) mais pas surjective (car Dupond $\in N$ n'admet pas d'antécédent).
  - ⋆ L'application $h : E \longrightarrow \{$janvier,février,...,décembre$\}$ qui à tout lauréat associe son mois de naissance n'est pas injective (car Laurent Schwartz et Alexandre Grothendieck sont nés en mars) mais surjective (quel que soit le mois, on peut trouver un lauréat qui est né le mois en question).

IV.6 Composition de fonctions injectives etc.

Théorème. Soient $f : E \longrightarrow F$ et $g : F \longrightarrow G$.

- Si f et g sont injectives, alors $g \circ f$ est injective.
- Si f et g sont surjectives, alors $g \circ f$ est surjective.
- Si f et g sont bijectives, alors $g \circ f$ est bijective.

En d'autres termes, une composée d'injections (respectivement surjections, bijections) est une injection (respectivement surjection, bijection).

La réciproque est fausse : si $g \circ f$ est injective, g ne l'est pas forcément (mais f l'est). De même, si $g \circ f$ est surjective, alors g est surjective, mais f ne l'est pas forcément (cf. exercice 28).

DÉMONSTRATION.
- Supposons f et g injectives, et montrons que $g \circ f$ est injective. Soit $(x_1, x_2) \in E^2$ (l'ensemble de départ de $g \circ f$) et supposons $x_1 \neq x_2$. Comme f est injective, on a $f(x_1) \neq f(x_2)$. Comme g est injective, on a $g(f(x_1)) \neq g(f(x_2))$, c'est-à-dire que $g \circ f(x_1) \neq g \circ f(x_2)$. Ainsi $g \circ f$ est injective.
- Supposons f et g surjectives, et montrons que $g \circ f$ est surjective. Soit $z \in G$ (l'ensemble d'arrivée de $g \circ f$). Comme g est surjective (de F dans G), il existe $y \in F$ tel que $z = g(y)$. Comme f est surjective (de E dans F) ; il existe $x \in E$ tel que $y = f(x)$, si bien que $z = g(f(x)) = g \circ f(x)$. Autrement dit x est un antécédent de z par $g \circ f$. Ainsi $g \circ f$ est surjective.

- Découle des deux points précédents. $\square$

Exemple : La fonction th est une bijection de $\mathbb{R}$ dans $]-1\,;1\,[$. Si f est une bijection de $]-1\,;1\,[$ dans $]\,a\,;b\,[$, alors $f\circ \text{th}$ est une bijection de $\mathbb{R}$ dans $]\,a\,;b\,[$: cf. exercice 59

Et cf. exercice 30 du chapitre 5 pour un raisonnement analogue.

V Image directe, image réciproque d'une partie

V.1 Image directe

Définition. Soit A une partie de E, soit $f : E \longrightarrow F$. On appelle image (directe) de A par f et on note $\{\, f(x)\,|\, x \in A\,\}$ ou encore $f(A)$ l'ensemble

$$\left\{\, y \in F \,\middle|\, \exists x \in A,\ y = f(x)\,\right\},$$

des valeurs prises par $f(x)$ lorsque x parcourt A.

On note parfois $f(E) = \text{Im}(f)$, qu'on appelle donc l'image de f. Mais on réservera cette écriture aux applications linéaires, cf. chapitre 29.

Remarque : Nous avons déjà évoqué la notation $\{\, f(x)\,|\, x \in A\,\}$ dans le chapitre 2. Précisons une fois pour toutes que la notation : $f(A) = \{\, f(x)\,|\, x \in A\,\}$ signifie que $f(A)$ est l'ensemble formé des $f(x)$ quand x parcourt A. En d'autres termes, $f(A)$ est l'ensemble des images des éléments de A par f. Cette notation est une façon plus concise d'écrire l'ensemble suivant :

$$f(A) = \left\{\, y \in F \,\middle|\, \exists x \in A,\ y = f(x)\,\right\},$$

c'est-à-dire l'ensemble des éléments de F admettant un antécédent qui appartient à A. Avec des quantificateurs, si y est un élément de F :

$$y \in f(A) \qquad \Longleftrightarrow \qquad \exists x \in A, \quad y = f(x).$$

On a $f(A) \subset F$. Il n'y a aucune raison que ce soit une égalité (cf. remarque plus bas).

Cette dernière équivalence (l'écriture avec des quantificateurs de « $y \in f(A)$ ») doit être sue sur le bout des doigts !

Exemples :

- $\exp(\mathbb{R}_+) = \left\{ e^x \,\middle|\, x \in \mathbb{R}_+ \right\} = [\,1\,;+\infty\,[$.
- Si f est la fonction carré, alors $f([\,-1\,;2\,]) = [\,0\,;4\,]$.
- Si f est la partie entière, $f([\,-1\,;2\,]) = \{-1; 0; 1; 2\}$.
- $\ln(\mathbb{R}_+^*) = \left\{ \ln(x) \,\middle|\, x \in \mathbb{R}_+^* \right\} = \mathbb{R}$.
- $\cos(\mathbb{R}) = \{\cos(x)\,|\, x \in \mathbb{R}\} = [\,-1\,;1\,]$.

Remarque : Comme on peut le voir avec la fonction carré et la partie entière, si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, en général, $f([\,a\,;b\,]) \neq [\,f(a)\,;f(b)\,]$! Ce n'est le cas que si f est monotone (pour les bornes) et continue (pour les valeurs intermédiaires) !

Ne pas confondre $f(E)$, l'image de f, avec F, son ensemble d'arrivée ! Tous les éléments de $f(E)$ sont atteints par définition, mais ce n'est pas forcément le cas de tous les éléments de F ! Plus précisément, tous les éléments de F sont atteints si et seulement si f est surjective. Dans le même ordre d'idée, $f(E)$ est inclus dans F, mais il n'y a pas forcement égalité : il y a égalité si et seulement si f est surjective !

Remarque : Toute application $f : E \longrightarrow F$ peut être « rendue surjective » en prenant $f(E)$ comme ensemble d'arrivée, c'est-à-dire que $f : E \longrightarrow f(E)$ est surjective par définition. Par exemple $\cos : \mathbb{R} \longrightarrow \mathbb{R}$ n'est ni injective ni surjective, mais $\cos : \mathbb{R} \longrightarrow [\,-1\,;1\,]$ est surjective.

En particulier, une fonction injective est une bijection sur son image. Par exemple, $\exp : \mathbb{R} \longrightarrow \mathbb{R}$ est injective non surjective, et $\exp : \mathbb{R} \longrightarrow \mathbb{R}_+^*$ est bijective.

On peut parfois faire le raisonnement inverse, c'est-à-dire « rendre injective » une fonction en prenant un domaine de définition plus petit (c'est-à-dire en prenant une certaine restriction de cette fonction). Cela ne servirait à rien de donner un résultat général car cela dépend trop de la fonction considérée (par exemple, pour rendre une fonction constante injective, il faut se restreindre à un singleton, ce qui a un intérêt limité), il vaut mieux le faire au cas par cas. Par exemple, la fonction carré (de $\mathbb{R}$ dans $\mathbb{R}$) n'est ni injective, ni surjective, mais sa restriction à $\mathbb{R}_+$ est injective (voir ci-dessous) d'image $\mathbb{R}_+$: on dit donc que la fonction carré induit une bijection de $\mathbb{R}_+$ dans lui-même. Nous verrons dans le chapitre 5 de nombreux exemples de fonctions « devenant » injectives ou bijectives en restreignant le domaine de définition.

V.2 Image réciproque.

Définition. Soit B une partie de F, soit $f : E \to F$. On appelle image réciproque de B par f et on note $f^{-1}(B)$ l'ensemble

$$\{x \in E \,|\, f(x) \in B\}$$

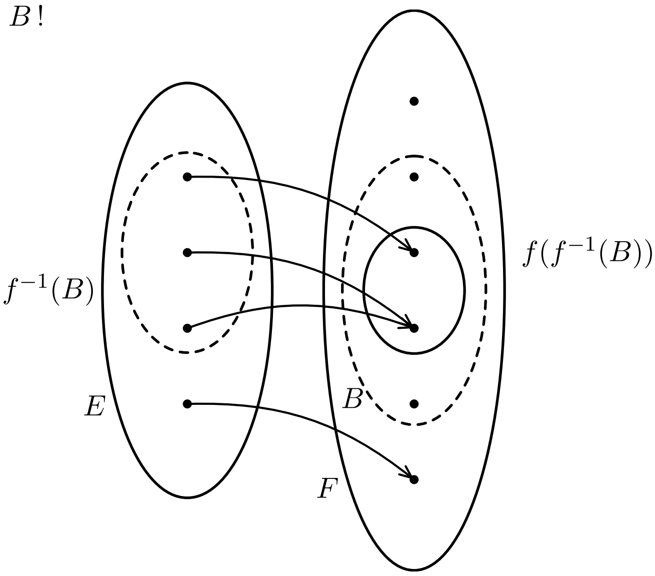
des éléments de E dont l'image appartient à B.

Remarques :

- $f^{-1}(B)$ est une partie de E (alors que $f(A)$ est une partie de F) et, si $x \in E$, alors : $x \in f^{-1}(B) \iff f(x) \in B$. Là aussi, cette équivalence doit être sue sur le bout des doigts.
- En d'autres termes, $f^{-1}(B)$ est l'ensemble des antécédents des éléments de B. Attention, certains éléments de B peuvent ne pas avoir d'antécédents. En particulier, si aucun élément de B n'admet d'antécédent, $f^{-1}(B) = \varnothing$!
- $f^{-1}(B)$ est un **ensemble** qui existe toujours (même si f n'est pas injective, même si B n'est pas inclus dans $f(E)$), contrairement à $f^{-1}(y)$ (voir paragraphe VI) qui n'est défini que lorsque f^{-1} est injective et $y \in f(E)$. On ne confondra donc pas $f^{-1}(B)$, l'image réciproque de B par f, avec $f^{-1}(y)$, l'image de y par l'application réciproque de f. Pour être sûr de ne pas se tromper : l'une prend en argument un ensemble, une partie de F, et l'autre un élément de F.
- Contrairement à ce que la notation f^{-1} pourrait laisser croire (et contrairement à ce qui est le cas pour l'application réciproque f^{-1} dans le paragraphe VI), on n'a pas forcément $f\left(f^{-1}(B)\right) = B$ pour toute partie B de F, ni $f^{-1}\left(f(A)\right) = A$ pour toute partie A de E (ce n'est le cas, respectivement, que si f est surjective et si f est injective, cf. exercice 43). Voir des contre-exemples ci-dessous.

Exemples :

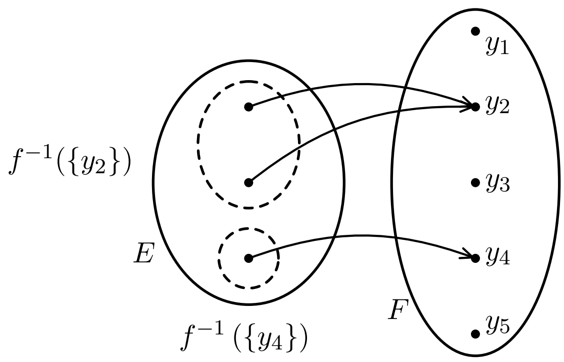
- Commençons par un exemple patatoïdal. En pointillés, à droite, B, et à gauche, $f(B)$. Nous avons aussi représenté $f\left(f^{-1}(B)\right)$ à droite (en traits pleins) : on voit que $f\left(f^{-1}(B)\right) \neq B$!



- Si f est la fonction carré et si $B = [\,1\,;4\,]$, alors $f^{-1}(B) = [\,-2\,;-1\,] \cup [\,1\,;2\,]$. Ici, on a $f\left(f^{-1}(B)\right) = B$. Cependant, si on prend $B = [\,-1\,;4\,]$, alors $f^{-1}(B) = [\,-2\,;2\,]$ et $f\left(f^{-1}(B)\right) = [\,0\,;4\,] \neq B$.
- Toujours si f est la fonction carré, $f^{-1}\left(\left[-1\,;-\dfrac{1}{2}\right]\right) = \varnothing$ et $f^{-1}(\mathbb{R}_+) = \mathbb{R}$.

Remarque : Si $f : E \to F$, alors $E = \displaystyle\bigcup_{y \in F} f^{-1}\left(\{y\}\right)$, et cette union est disjointe. Nous utiliserons régulièrement ce résultat cette année, par exemple pour définir le système complet d'événements associé à une variable aléatoire, cf. chapitre 27.

Rappelons que $f^{-1}(B)$ est l'ensemble des éléments de E dont l'image appartient à B ou, ce qui revient au même, l'ensemble des éléments des éléments de B. Par conséquent, pour construire $f^{-1}(B)$ dans l'exemple ci-contre (c'est moins facile avec des ensembles infinis mais le principe est le même), on prend tous les éléments de B puis on prend tous leurs antécédents. Pour construire $f\left(f^{-1}(B)\right)$, on prend les images de tous les éléments de $f^{-1}(B)$. Ces images appartiennent à B par définition, mais tous les éléments de B ne sont pas forcément atteints ! C'est pour cela qu'on n'a pas forcément $f\left(f^{-1}(B)\right) = B$.



Si, de plus, tous les ensembles $f^{-1}(\{y\})$ sont non vides (c'est-à-dire si f est surjective), alors ces ensembles forment une partition de E. Cependant, ce résultat est utile même quand ce n'est pas le cas.

Attention, encore une fois, on ne confondra pas $f^{-1}(\{y\})$ qui est l'image réciproque d'un **ensemble** et qui est toujours définie, qui peut être un ensemble quelconque (un singleton, un ensemble infini, l'ensemble vide etc.) et $f^{-1}(y)$ (cf. paragraphe suivant), l'image d'un **élément** de y, qui n'existe que lorsque f est injective et que $y \in f(E)$, et qui est un élément de E.

VI Application réciproque.

VI.1 Généralités.

Proposition/Définition. Soit $f : E \longrightarrow F$ injective. On pose

$$f^{-1} : \left\{ \begin{array}{ccl} f(E) & \longrightarrow & E \\ y & \longmapsto & \text{l'unique } x \in E \text{ tel que } f(x) = y \end{array} \right.$$

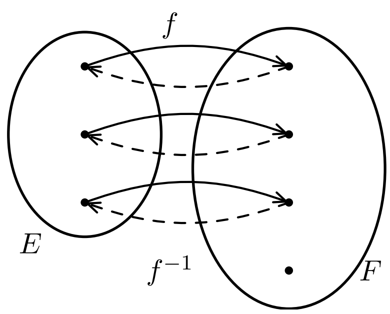
Alors f^{-1} est bien définie et est une bijection de $f(E)$ dans E : on l'appelle l'application ou la bijection réciproque de f. Enfin, $\left(f^{-1}\right)^{-1} = f$.

En d'autres termes, $f^{-1}(y)$ est l'unique antécédent de y par f. Le fait que f soit injective est indispensable : en effet, si un élément y de F admet plusieurs antécédents, lequel choisir comme valeur pour $f^{-1}(y)$?

Démonstration.
- Tout d'abord, f^{-1} est bien définie car tout élément de $f(E)$ admet un unique antécédent par f : l'existence découle du fait que tout élément de $f(E)$ admet au moins un antécédent (par définition d'une image) et l'unicité vient de l'injectivité de f.
- f^{-1} est injective car deux éléments distincts de $f(E)$ ne peuvent pas admettre un antécédent commun par f donc ne peuvent pas avoir la même image par f^{-1}.
- f^{-1} est surjective car tout élément de E admet une image. Dès lors, pour tout $x \in E$, x est un antécédent (et même le seul par injectivité de f) de $f(x)$ donc $f^{-1}(f(x)) = x$: $f(x)$ est un antécédent de x, f^{-1} est surjective donc bijective.
- Soit $x \in E$. Alors x est l'unique antécédent de $f(x)$ par f, par injectivité de f, si bien que $f^{-1}(f(x)) = x$: par injectivité de f^{-1} (voir ci-dessus), $f(x)$ est l'unique antécédent de x par f^{-1}. En d'autres termes, $\left(f^{-1}\right)^{-1}(x) = f(x)$.

Remarques :

- ⚠ f^{-1} n'est pas forcément définie sur F tout entier !



En effet, $f^{-1}(y)$ n'a un sens que si y admet un antécédent donc si y est une image, c'est-à-dire si $y \in f(E)$. f^{-1} n'est définie sur F tout entier que si $f(E) = F$ ie si f est surjective donc bijective (car elle est supposée injective) et alors f^{-1} est une bijection de F dans E, appelée la bijection réciproque de f.

- Par définition (toujours si f est injective) :

$$\forall x \in E, \quad f^{-1}(f(x)) = x \qquad \text{et} \qquad \forall y \in f(E), \quad f\left(f^{-1}(y)\right) = y$$

- Comme on l'a déjà dit, on ne confondra pas l'application réciproque f^{-1} de ce paragraphe, qui prend en argument un **élément** de F et qui n'existe que lorsque f est injective, et l'image réciproque $f^{-1}(B)$, qui prend donc en entrée une **partie** de F et qui existe toujours. On pourrait s'étonner du fait d'utiliser la même notation pour deux choses différentes, mais en fait ce choix est cohérent car les deux notations sont compatibles : si f est une bijection de E dans F et si B est une partie de F, alors $f^{-1}(B)$ désigne à la fois l'image réciproque de B par la fonction f (cf. paragraphe V.2) et l'image directe de B par la fonction f^{-1} (cf. paragraphe V.1) et ces deux ensembles sont égaux : il est donc naturel de les noter de la même façon !

Dans le cas général, f est une bijection de E dans $f(E)$ (rappelons qu'une injection est toujours une bijection sur son image.

Lemme. Soit I un intervalle (de $\mathbb{R}$) non vide, non réduit à un point. Soit $f : I \longrightarrow \mathbb{R}$ strictement monotone. Alors f est injective sur I.

Démonstration. Soit $(x_1, x_2) \in I^2$. Supposons $x_1 \neq x_2$. Alors l'un des deux est strictement supérieur à l'autre. Sans perte de généralité, supposons $x_1 < x_2$. Si f est strictement croissante, $f(x_1) < f(x_2)$ tandis que, si f est strictement décroissante, $f(x_1) > f(x_2)$. Dans tous les cas, $f(x_1) \neq f(x_2)$, ce qui est le résultat voulu.

Exemples :

- La fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right.$$

est injective (car strictement croissante), d'image $\mathbb{R}_+$. Son application réciproque est donc la fonction f^{-1} définie sur $\mathbb{R}_+$ par : $f^{-1} : y \longmapsto$ « l'unique réel x positif tel que $x^2 = y$ ». En d'autres termes, f^{-1} est la fonction racine carrée.

- Soit $f = \exp$. La fonction f est strictement croissante sur $\mathbb{R}$ donc injective et $f(\mathbb{R}) = \mathbb{R}_+^*$. Son application réciproque est donc la fonction f^{-1} définie sur $\mathbb{R}_+^*$ par : $f^{-1} : y \longmapsto$ « l'unique $x \in \mathbb{R}$ tel que $e^x = y$ ». En d'autres termes, f^{-1} est la fonction ln.

La fonction carré n'est pas injective sur $\mathbb{R}$: on ne peut pas définir son application réciproque. Cela n'a aucun sens de parler de « l'unique réel dont le carré vaut y » !

Méthode pour expliciter f^{-1} (quand f est injective) : Prendre un élément $y \in f(E)$ et résoudre l'équation (d'inconnue $x \in E$) $y = f(x)$. L'unique solution sera $f^{-1}(y)$.

Exemple : Montrons que sh est une bijection de $\mathbb{R}$ dans $\mathbb{R}$ et expliciter sa bijection réciproque.

Soient $y \in \mathbb{R}$ et $x \in \mathbb{R}$.

$$\begin{aligned} y = \operatorname{sh}(x) \quad & \iff \quad e^x - e^{-x} = 2y \\ & \iff \quad e^{2x} - 2ye^x - 1 = 0 \\ & \iff \quad e^x \text{ est solution de } z^2 - 2yz - 1 = 0. \end{aligned}$$

Cette équation (un trinôme du second degré de discriminant $\Delta = 4y^2 + 4 > 0$) admet deux solutions : $y \pm \sqrt{y^2+1}$. Or, par stricte croissance de la racine carrée, et puisque $|y| \geqslant \pm y$, il vient :

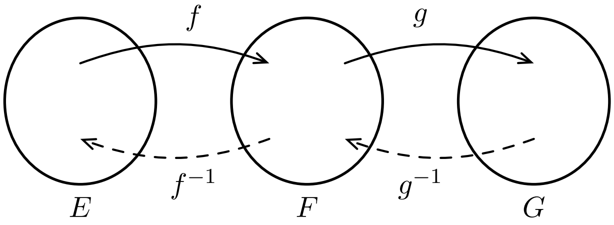
$$\sqrt{y^2+1} > \sqrt{y^2} = |y| \geqslant \pm y$$

En particulier, $y + \sqrt{y^2+1} > 0$ et $y - \sqrt{y^2+1} < 0$. Finalement, $y = \operatorname{sh}(x)$ si et seulement si $e^x = y + \sqrt{y^2+1}$ (l'autre égalité est impossible) si et seulement si $x = \ln\left(y + \sqrt{y^2+1}\right)$. En d'autres termes, sh est bien une bijection de $\mathbb{R}$ dans $\mathbb{R}$ (voir ci-contre) et sh^{-1} est définie sur $\mathbb{R}$ par $\operatorname{sh}^{-1} : y \longmapsto \ln\left(y + \sqrt{y^2+1}\right)$.

Cas particulier important : on veut montrer que $f : E \longrightarrow F$ est bijective et expliciter f^{-1}. Si on arrive à montrer que pour tout $y \in F$ l'équation $y = f(x)$ admet une unique solution x, alors on obtient à la fois la bijectivité de f (car tout élément de F admet un unique antécédent) et l'expression de f^{-1} (car $f^{-1}(y) = x$ par définition de f^{-1}). Le problème est qu'en général on ne sait pas expliciter f^{-1}. On n'appliquera donc cette méthode que si on demande l'expression explicite de f^{-1}. Sinon, on appliquera plutôt le théorème de la bijection (cf. VII).

Proposition. Soient $f : E \to F$ et $g : F \to G$ bijectives. Alors $g \circ f$ est bijective et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Remarque : On change l'ordre des fonctions ! Cela se voit d'une part si on fait un dessin représentant les deux fonctions et leur réciproque :



Cela se comprend aussi très bien : si on applique f puis g et qu'on veut revenir en arrière, on commence par appliquer g^{-1} puis f^{-1}. Par exemple, si on met un coffre dans un trou (f) puis qu'on met de la terre dans le trou (g), et si on veut ensuite récupérer le coffre (défaire $g \circ f$), on commence par enlever la terre (g^{-1}) puis sortir le coffre du trou (f^{-1}).

Ce sera la même chose quand on voudra inverser un produit dans un groupe, cf. chapitre 18.

Démonstration. On sait déjà que $g \circ f$ est bijective (cf. paragraphe IV.6). Soit $z \in G$. Soit y l'unique antécédent de z par g et soit x l'unique antécédent de y par f (ils existent et sont uniques car f et g sont bijectives). Alors $g \circ f(x) = g(y) = z$ donc x est un antécédent de z par $g \circ f$ et, par injectivité, c'est le seul, c'est-à-dire que $x = (g \circ f)^{-1}(z)$. Or, $x = f^{-1}(y)$ et $y = g^{-1}(z)$ si bien que $(g \circ f)^{-1}(z) = f^{-1} \circ g^{-1}(z)$. z étant quelconque, on a le résultat voulu.

Remarque : On généralise aisément à un nombre fini quelconque d'applications par récurrence.

En utilisant l'associativité de la composition.

VI.2 Retour aux changements d'indice

Donnons enfin le résultat théorique justifiant la validité des changements d'indice que nous faisons en pratique :

Proposition. Soit $(x_i)_{i \in I}$ une famille de réels indexée par un ensemble (fini) I et soit φ une fonction injective sur I. Alors :

$$\sum_{i \in I} x_i = \sum_{j \in \varphi(I)} x_{\varphi^{-1}(j)}$$

Remarque : Une fonction injective étant une bijection sur son image, on peut reformuler le théorème : si φ est une bijection de I dans J alors

$$\sum_{i \in I} x_i = \sum_{j \in J} x_{\varphi^{-1}(j)}$$

Le fait qu'un changement de variable doive être bijectif n'est pas surprenant : un changement de variable n'est qu'une façon différente d'écrire la même somme, et chaque terme de la deuxième écriture de la somme doit correspondre à un et un seul terme de la première.

Il n'y a rien à montrer : φ étant bijective de I dans J, les deux sommes sont toutes deux égales à $x_{i_1} + \cdots + x_{i_n}$, où $I = \{i_1 ; \dots ; i_n\}$. En pratique, comme on l'a vu, on ne définit pas φ et on raisonne de façon plus simple, mais garder en tête le fait qu'un changement de variable doit être bijectif permet d'éviter certaines erreurs, par exemple cela permet d'éviter d'écrire l'ânerie

$$\sum_{k=0}^{n} \binom{2n}{2k} = \sum_{j=0}^{2n} \binom{2n}{j}$$

La raison profonde est que la fonction $k \mapsto 2k$ n'est pas une bijection entre $[\![0 ; n]\!]$ et $[\![0 ; 2n]\!]$: par exemple 3 n'est pas atteint. Mais il est inutile d'être aussi subtil : il suffit de voir que la somme de droite contient des termes qui ne sont pas dans la somme de gauche, alors que, lorsqu'on fait un changement d'indice, les deux sommes doivent contenir les mêmes termes, puisqu'elles sont égales ! Morale de l'histoire : toujours se demander si les deux sommes contiennent les mêmes termes !

Le changement de variable $j = 2k$ correctement effectué donne

$$\sum_{k=0}^{n}\binom{2n}{2k} = \sum_{\substack{j=0\\ j \text{ pair}}}^{2n}\binom{2n}{j}$$

VI.3 Involutions

Définition. Soit $f : E \longrightarrow E$. On dit que f est une involution si $f \circ f = \mathrm{Id}_E$ c'est-à-dire si, pour tout $x \in E, f(f(x)) = x$.

Une involution f va, par définition, d'un ensemble dans lui-même, sinon $f \circ f$ n'est même pas définie.

Exemples :

- Id_E est évidemment une involution de E.
- La fonction inverse est une involution de $\mathbb{R}^*$.
- La fonction $x \longmapsto -x$ est une involution de $\mathbb{R}$.
- La fonction $x \longmapsto 1789 - x$ est une involution de $\mathbb{R}$.
- Dans le plan ou l'espace : une symétrie est une involution.
- La fonction $A \mapsto \overline{A}$, qui va de $\mathscr{P}(E)$ dans lui-même.

En clair : une fonction $f : E \longrightarrow E$ est une involution si, quand on l'applique deux fois, on revient à son point de départ.

Proposition. Soit $f : E \longrightarrow E$ une involution. Alors f est une bijection de E dans lui-même.

Démonstration.
- Soit $(x_1, x_2) \in E^2$. Supposons $f(x_1) = f(x_2)$. Alors $f(f(x_1)) = f(f(x_2))$ donc $x_1 = x_2$: f est injective.
- Soit $y \in E$. Alors $y = f(f(y))$: $f(y)$ est un antécédent de y par f donc f est surjective.

Rappelons que pour toute fonction f, si $a = b$ alors $f(a) = f(b)$, cf. paragraphe IV.3.

Remarques :

- En particulier, pour tout $y \in E$, $f(y)$ est même l'unique antécédent de y par f, c'est-à-dire que $f^{-1}(y) = f(y)$. Ainsi $f = f^{-1}$.
- Ce résultat est un cas particulier du théorème suivant :

Théorème. Soient $f : E \longrightarrow F$ et $g : F \longrightarrow E$ telles que

$$f \circ g = \mathrm{Id}_F \qquad \text{et} \qquad g \circ f = \mathrm{Id}_E$$

Alors f et g sont bijectives et $g = f^{-1}$ et $f = g^{-1}$.

Il faut impérativement les deux égalités ! Si on a simplement $f \circ g = \mathrm{Id}_F$, alors f n'est pas forcément bijective (cf. exercice 37). On pourra se contenter d'une seule égalité quand on sera dans des espaces vectoriels de dimension finie (cf. chapitre 30).

Démonstration.
- Soit $(x_1, x_2) \in E^2$. Supposons $f(x_1) = f(x_2)$. Alors $g(f(x_1)) = g(f(x_2))$ donc $x_1 = x_2$: f est injective.
- Soit $y \in F$. Alors $y = f(g(y))$: $g(y)$ est un antécédent de y par f donc f est surjective.

Finalement, f est injective et surjective donc est bijective. De plus, pour tout $y \in F$, l'unique antécédent de y est $g(y)$, si bien que $f^{-1}(y) = g(y)$. Ainsi $f^{-1} = g$. Par symétrie (f et g jouent le même rôle), g est également bijective et $g^{-1} = f$. □

VI.4 Activité : théorème de Cantor-Bernstein.

Nous verrons au chapitre 17 que si E et F sont des ensembles finis, alors :

- il existe une injection de E dans F si et seulement si $\mathrm{card}(E) \leqslant \mathrm{card}(F)$.
- il existe une surjection de E dans F si et seulement si $\mathrm{card}(E) \geqslant \mathrm{card}(F)$.
- il existe une bijection de E dans F si et seulement si $\mathrm{card}(E) = \mathrm{card}(F)$.

Ces résultats sont complètement intuitifs et se voient bien sur les patatoïdes dessinés plus haut. Cependant, cela se gâte un peu quand on manipule des ensembles infinis. En effet, si E et F sont des ensembles infinis, il n'existe pas forcément de bijection entre E et F : nous montrerons au chapitre 17, par exemple, qu'il n'existe pas de bijection de $\mathbb{N}$ dans $\mathbb{R}$, et il n'existe pas de bijection de $\mathbb{R}$ dans $\mathscr{P}(\mathbb{R})$ (cf. exercice 20). Par conséquent, pour les ensembles infinis, on prend le chemin inverse :

- on se dit que deux ensembles infinis « ont la même taille » lorsqu'il existe une bijection entre eux.
- on se dit que « E est plus petit que F » s'il existe une injection de E dans F. On dit alors que E s'injecte dans F, et on note $E \hookrightarrow F$ (par exemple, cf. exercice 20, E s'injecte dans $\mathscr{P}(E)$).
- on se dit « que E est plus gros que F » lorsqu'il existe une surjection de E dans F.

Tout cela peut être rendu rigoureux avec la notion de cardinal d'un ensemble infini. Nous en dirons un mot au chapitre 17, le but de ce paragraphe est de vous forger une intuition et de manipuler les outils vus plus haut. Retenez simplement qu'il faut faire attention, et que ce n'est pas parce que deux ensembles sont infinis « qu'ils ont le même nombre d'éléments » : il y a des infinis plus gros que d'autres ! Par exemple, puisqu'il n'y a pas de bijection de $\mathbb{N}$ dans $\mathbb{R}$ (cf. chapitre 17), alors $\mathbb{R}$ « est plus gros que $\mathbb{N}$ ». Cependant, puisqu'il existe une bijection de $\mathbb{N}$ dans $\mathbb{Z}$, alors « $\mathbb{N}$ et $\mathbb{Z}$ ont la même taille », ce qui se voit très bien avec un dessin (cf. exercice 44).

Le théorème suivant est donc complètement intuitif (même si la démonstration est loin d'être triviale) :

Théorème (Théorème de Cantor-Bernstein (HP)). Soient E et F des ensembles non vides. S'il existe une injection de E dans F et une injection de F dans E, alors il existe une bijection de E dans F.

Démonstration. On suppose donc qu'il existe $f : E \to F$ et $g : F \to E$ injectives.

- Première étape : montrer qu'une fonction $\varphi : \mathscr{P}(E) \to \mathscr{P}(E)$ croissante au sens de l'inclusion (i.e. telle que si $X \subset Y \subset E$ alors $\varphi(X) \subset \varphi(Y)$) admet un point fixe (i.e. qu'il existe $A \in \mathscr{P}(E)$ telle que $\varphi(A) = A$).

  Soit donc $\varphi : \mathscr{P}(E) \to \mathscr{P}(E)$ croissante au sens de l'inclusion. Posons

  $$A = \bigcup_{X \subset E, X \subset \varphi(X)} X$$

  Montrons que A convient i.e. que $A = \varphi(A)$. Tout d'abord, si $X \subset \varphi(X)$, alors $X \subset A$ (car X est un des éléments dont l'union forme A) donc, par croissance de φ, $\varphi(X) \subset \varphi(A)$. Mais alors, puisque $X \subset \varphi(X)$, alors $X \subset \varphi(A)$. En d'autres termes, A est une union d'ensembles tous inclus dans $\varphi(A)$ donc $A \subset \varphi(A)$.

  C'est l'analogue de l'exercice 68 du chapitre 13.

  On utilise le fait (évident : exo) que si les A_i sont tous inclus dans $\varphi(A)$ alors $\displaystyle\bigcup_{i \in I} A_i \subset \varphi(A)$.

  Par croissance de φ, $\varphi(A) \subset \varphi(\varphi(A))$: en d'autres termes, $\varphi(A)$ est un des ensembles composant l'union définissant A donc $\varphi(A) \subset A$: on a le résultat par double inclusion.

- Deuxième étape : montrer qu'il existe $A \subset E$ tel que $g(F \backslash f(A)) = \overline{A}$.

  On cherche à utiliser la question précédente : reformulons le résultat sous la forme $A = \varphi(A)$ avec φ bien choisie. On voit que le problème revient à trouver $A \subset E$ telle que $A = \overline{g(F\backslash f(A))} = E \backslash g(F\backslash f(A))$. Pour appliquer la question précédente, il suffit de définir

  Rappelons que g est à valeurs dans E donc, si A est une partie de E, $g(F\backslash f(A))$ est aussi une partie de E.

  $$\varphi : \left\{ \begin{array}{ccc} \mathscr{P}(E) & \longrightarrow & \mathscr{P}(E) \\ X & \longmapsto & E\backslash g(F\backslash f(X)) \end{array} \right.$$

  et de montrer que φ est croissante pour l'inclusion : la première étape permettra de conclure. Soient donc X et Y deux parties de E avec $X \subset Y$. Montrons que $f(X) \subset f(Y)$. Soit $y \in f(X)$: il existe $x \in X$ tel que $y = f(x)$. Or, $X \subset Y$ donc $x \in Y$ si bien que $y = f(x) \in f(Y)$: $f(X) \subset f(Y)$. Par conséquent , $F\backslash f(Y) \subset F\backslash f(X)$. L'inclusion

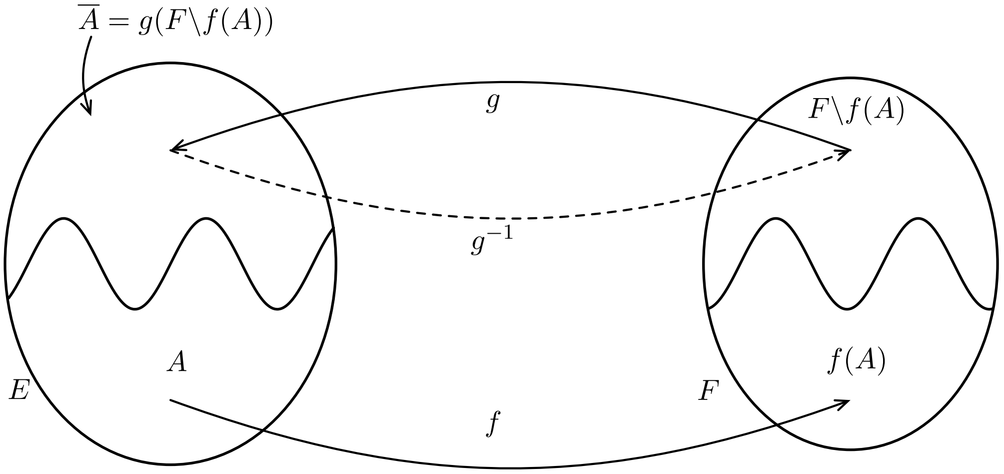
  Rappelons que le passage au complémentaire est décroissant, c'est-à-dire que si A et B sont deux parties de F et si $A \subset B$ alors $\overline{B} \subset \overline{A}$, cf. paragraphe III.

  $$g\left(F\backslash f(Y)\right) \subset g\left(F\backslash f(X)\right)$$

  en découle de la même façon que l'inclusion $f(X) \subset f(Y)$ découle de l'inclusion $X \subset Y$. Finalement (toujours par décroissance du passage au complémentaire), $\varphi(X) \subset \varphi(Y)$, ce qui permet de conclure.

- Troisième étape : conclure.

  Soit donc $A \subset E$ telle que $\overline{A} = g\left(F \backslash f(A)\right)$. Faisons un dessin :



  On cherche une bijection de E dans F. On veut une fonction de E dans F : « en bas », on prend f, et « en haut », on prend g^{-1}. Soit donc :

$$h : \begin{cases} E & \longrightarrow & F \\ x & \longmapsto & \begin{cases} f(x) & \text{si } x \in A \\ g^{-1}(x) & \text{sinon} \end{cases} \end{cases}$$

□

  h est bien définie : en effet, si $x \in \overline{A}$, alors $x \in g(F\backslash f(A))$ et en particulier $x \in g(F)$ et g est injective, donc $g^{-1}(x)$ existe bien et est bien défini. Montrons que h est bijective.

  Montrons que h est injective. Soit $(x_1, x_2) \in E^2$. Supposons $x_1 \neq x_2$.

  - ⋆ Si x_1 et x_2 appartiennent à A, alors $h(x_1) = f(x_1)$ et $h(x_2) = f(x_2)$. Or, $f(x_1) \neq f(x_2)$ car f est injective donc $h(x_1) \neq h(x_2)$.
  - ⋆ De même, g^{-1} est injective donc $h(x_1) \neq h(x_2)$ si x_1 et x_2 appartiennent à $\overline{A}$.
  - ⋆ Supposons enfin que $x_1 \in A$ et $x_2 \notin A$. Alors $h(x_1) = f(x_1) \in f(A)$ et $h(x_2) = g^{-1}(x_2) \in F\backslash f(A)$: en effet, $x_2 \in g(F\backslash f(A))$ donc x_2 est l'image d'un élément de $F\backslash f(A)$. g étant injective, l'unique antécédent de x_2 appartient à $F\backslash f(A)$, c'est-à-dire que $g^{-1}(x_2) \in F\backslash f(A)$. En particulier, $h(x_1) \neq h(x_2)$.

  Dans tous les cas, $h(x_1) \neq h(x_2)$: h est injective.

  Montrons que h est surjective. Soit $y \in F$.

  - ⋆ Si $y \in f(A)$ alors il existe $x \in A$ tel que $y = f(x) = h(x)$.
  - ⋆ Si $y \in F\backslash f(A)$, alors $x = g(y) \in g(F\backslash f(A)) = \overline{A}$ donc $g^{-1}(x) = h(x) = y$.

VII Cas particulier des fonctions de la variable réelle à valeurs dans $\mathbb{R}$.

Théorème (Théorème de la bijection (admis provisoirement)). Soit I un intervalle, soit $f : I \to \mathbb{R}$ continue strictement monotone. Alors f est une bijection de I dans $f(I)$, et f^{-1} est continue de même monotonie que f.

De plus, $f(I)$ est un intervalle. Ce théorème sera démontré dans le chapitre 13.

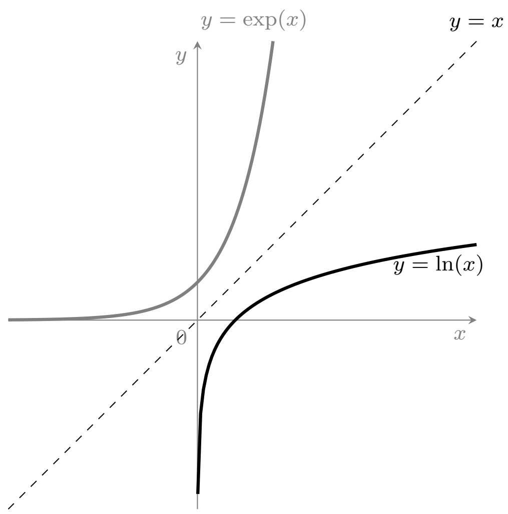
Remarques :

- Cependant, cela ne donne pas f^{-1} ni $f(I)$. En pratique (quand c'est possible : on ne cherchera donc à expliciter f^{-1} que lorsque cela sera demandé), on explicite f^{-1} comme ci-dessus et $f(I)$ grâce au tableau de variations, voir ci-dessous et l'exercice 57.

- Cela ne marche que parce que l'on dispose d'outils puissants sur $\mathbb{R}$ (la monotonie et la continuité). Cependant, sur un ensemble plus général (par exemple : sur $\mathbb{C}$!), on ne dispose pas de ces outils. On doit donc dans ce cas prouver la bijectivité éventuelle « à la main ».

Pour faire simple : interdiction d'utiliser le théorème de la bijection lorsqu'on ne se trouve pas sur un intervalle de $\mathbb{R}$.

- Le graphe de f^{-1} est le symétrique du graphe de f par rapport à la première bissectrice : ci-dessous les graphes du ln et de l'exponentielle.



- On voit avec la fonction carré de $\mathbb{R}_+$ dans $\mathbb{R}_+$ que le résultat est faux si on remplace « continue » par « dérivable » (en effet, sa f^{-1} est alors la racine carrée, non dérivable) : ce n'est pas parce que f est dérivable que f^{-1} l'est ! On se demande donc quand la réciproque d'une bijection dérivable est dérivable.

Théorème (admis provisoirement). Soit $f : I \longrightarrow \mathbb{R}$ dérivable strictement monotone. Soit $a \in I$ tel que $\underline{f'(a) \neq 0}$. Alors f^{-1} est dérivable en $b = f(a)$ et

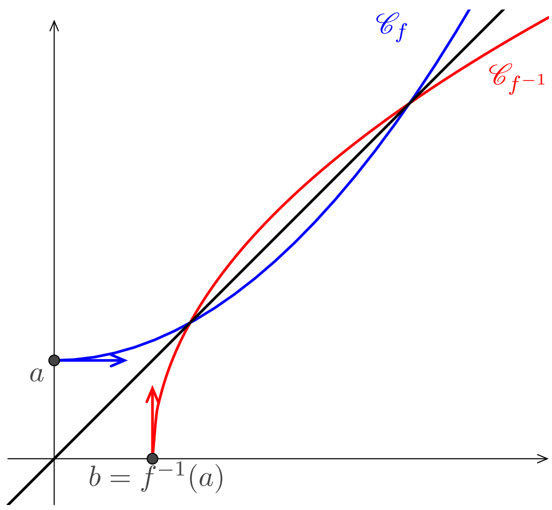
$$\left(f^{-1}\right)'(b) = \frac{1}{f'\left(f^{-1}(b)\right)}.$$

Remarque : Si $f'(a) = 0$, on peut également montrer que f^{-1} n'est pas dérivable en b.

En conclusion : ⚠ f^{-1} est dérivable en $b = f(a)$ si et seulement si $f'(a) \neq 0$. En d'autres termes : f^{-1} **est dérivable uniquement en l'image (par f) des points en lesquels f' ne s'annule pas !**

Remarque : C'est intuitif ! Rappelons que le graphe de f^{-1} est le symétrique du graphe de f par rapport à la première bissectrice. Ainsi, si $f'(a) = 0$, alors $\mathscr{C}_f$ admet une tangente horizontale en a donc $\mathscr{C}_{f^{-1}}$ admet une tangente verticale en $b = f(a)$ donc f^{-1} n'est pas dérivable en b.

Interprétation géométrique de ce théorème : sous réserve d'existence, la tangente au graphe de f^{-1} en $b = f(a)$ est le symétrique de la tangente au graphe de f en a par rapport à la première bissectrice.



Exemple : Montrer que la fonction ch induit une bijection de $\mathbb{R}_+$ dans un ensemble que l'on explicitera, et étudier la dérivabilité de ch^{-1}.

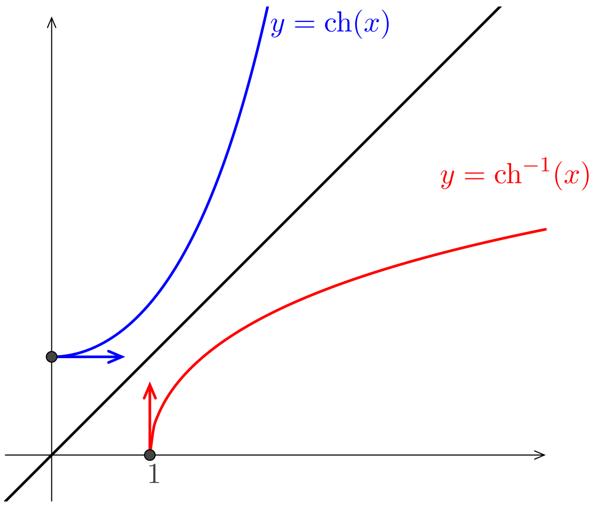
- ch est dérivable sur $\mathbb{R}_+$ car somme de fonctions dérivables. Soit $x \geqslant 0$. On a : $\text{ch}'(x) = \text{sh}(x)$. D'où le tableau de variations suivant :

x	0		$+\infty$
$\text{ch}'(x)$	0	$+$	
ch	1	$\nearrow$	$+\infty$

- ch est continue, strictement croissante, $\text{ch}(0) = 1$ et $\text{ch}(x) \xrightarrow[x \to +\infty]{} +\infty$. D'après le théorème de la bijection, ch est une bijection de $\mathbb{R}_+$ dans $[\,1\,;+\infty\,[$.

En d'autres termes, la restriction de ch à $\mathbb{R}_+$ est une bijection de $\mathbb{R}_+$ dans $[\,1\,;+\infty\,[$, ou encore : ch induit une bijection de $\mathbb{R}_+$ dans $[\,1\,;+\infty\,[$. Attention, la fonction ch tout court (i.e. définie sur $\mathbb{R}$) n'est pas injective car elle est paire !

- Étudions la dérivabilité de ch^{-1} (définie sur $[\,1\,;+\infty\,[$). Rappelons que ch^{-1} est dérivable sauf en l'image (par ch) des points où ch' s'annule. Or, ch' s'annule uniquement en 0. Dès lors, ch^{-1} est dérivable sur $[\,1\,;+\infty\,[\,\backslash\{\text{ch}(0)\} = \,]\,1\,;+\infty\,[$.



- Soit $x > 0$.

$$\left(\text{ch}^{-1}\right)(x) = \frac{1}{\text{ch}'\left(\text{ch}^{-1}(x)\right)} = \frac{1}{\text{sh}\left(\text{ch}^{-1}(x)\right)}$$

On rappelle que $\text{ch}(x)^2 - \text{sh}(x)^2 = 1$ donc $\text{sh}(x)^2 = \text{ch}(x)^2 - 1$. Or, $\text{sh}(x) \geqslant 0$ (on est sur $\mathbb{R}_+$) si bien que $\text{sh}(x) = \sqrt{\text{ch}(x)^2 - 1}$. Dès lors,

$$\left(\text{ch}^{-1}\right)'(x) = \frac{1}{\sqrt{\text{ch}\left(\text{ch}^{-1}(x)\right)^2 - 1}} = \frac{1}{\sqrt{x^2 - 1}}$$

Un raisonnement analogue à celui du paragraphe VI.1 donne : $\text{ch}^{-1}(x) = \ln\left(x + \sqrt{x^2 - 1}\right)$, ce qui nous permet également de trouver $\left(\text{ch}^{-1}\right)'$. Cependant, il n'est pas possible en général d'expliciter la réciproque d'une fonction : le théorème ci-dessus doit donc être su sur le bout de doigts.