Relations binaires et applications

Dans ce chapitre, E, F, G, I et J sont des ensembles tout à fait quelconques, pas spécialement des ensembles de nombres ou d'objets connus.

1 Relations binaires

1.1 Premières définitions, exemples

Les relations binaires sont partout, nous passons notre temps à comparer des objets les uns avec les autres selon tel ou tel aspect : « Minou et Matou ont la même couleur de poils », « $3 \leqslant 5$ », « Truc est amoureux de Bidule », « 4 divise 12 »… Est-il donc si facile de définir la notion de relation en général ? Par exemple, quel **objet mathématique** la relation d'infériorité stricte $<$ est-elle sur l'ensemble $[\![1,3]\!]$? Ce qui est sûr, c'est que cette relation est entièrement caractérisée par les affirmations : $1<2$, $1<3$ et $2<3$. Si on sait ça, on sait tout. Formellement, nous pouvons donc définir la relation $<$ comme l'ensemble $\big\{(1,2),(1,3),(2,3)\big\}$ des couples $(x,y) \in [\![1,3]\!]^2$ pour lesquels $x<y$, car connaître cet ensemble, c'est connaître la relation $<$. Dans cette optique :

— la relation $\leqslant$ sur $[\![1,3]\!]$ est l'ensemble $\big\{(1,1),(1,2),(1,3),(2,2),(2,3),(3,3)\big\}$,

— la relation de divisibilité $|$ sur $[\![1,4]\!]$ est l'ensemble $\big\{(1,1),(1,2),(1,3),(1,4),(2,2),(2,4),(3,3),(4,4)\big\}$.

Définition **(Relation binaire)** On appelle *relation binaire sur E* toute partie de $E \times E$.

Pour une telle relation $\mathcal{R}$, la proposition $(x,y) \in \mathcal{R}$ sera notée $x \mathcal{R} y$ pour tous $x, y \in E$ et lue « x est en relation avec y par $\mathcal{R}$ ».

Attention ! Parce que le couple (x,y) n'est pas le couple (y,x), la relation $x \mathcal{R} y$ peut être vraie sans que la relation $y \mathcal{R} x$ le soit. Dieu que l'amour est cruel !

Exemple Vous connaissez déjà un certain nombre de relations binaires :

— la relation d'égalité $=$ sur E, — les relations $\leqslant$ et $<$ sur $\mathbb{R}$, — la relation d'inclusion $\subset$ sur $\mathcal{P}(E)$,

— la relation $\leqslant$ sur l'ensemble $\mathbb{R}^{\mathbb{R}}$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, définie pour toutes fonctions $f, g \in \mathbb{R}^{\mathbb{R}}$ par l'équivalence :
$$f \leqslant g \iff \forall x \in \mathbb{R},\ f(x) \leqslant g(x),$$

— la relation de divisibilité $|$ sur $\mathbb{Z}$, définie pour tous $a, b \in \mathbb{Z}$ par l'équivalence : $a \mid b \iff \exists k \in \mathbb{Z},\ b = ak$,

— pour tout $\alpha \in \mathbb{R}$, la relation $\equiv\ [\alpha]$ de congruence modulo α sur $\mathbb{R}$, définie pour tous $x, y \in \mathbb{R}$ par l'équivalence :
$$x \equiv y\ [\alpha] \iff \exists k \in \mathbb{Z},\ x = y + k\alpha.$$

Définition **(Propriétés des relations binaires)** Soit $\mathcal{R}$ une relation binaire sur E.

- **Réflexivité :** On dit que $\mathcal{R}$ est *réflexive* si : $\forall x \in E,\ x \mathcal{R} x$.
- **Transitivité :** On dit que $\mathcal{R}$ est *transitive* si : $\forall x, y, z \in E,\ \big(x \mathcal{R} y \text{ et } y \mathcal{R} z\big) \implies x \mathcal{R} z$.
- **Symétrie :** On dit que $\mathcal{R}$ est *symétrique* si : $\forall x, y \in E,\ x \mathcal{R} y \implies y \mathcal{R} x$.
- **Antisymétrie :** On dit que $\mathcal{R}$ est *antisymétrique* si : $\forall x, y \in E,\ \big(x \mathcal{R} y \text{ et } y \mathcal{R} x\big) \implies x = y$.

Exemple

- La relation d'égalité $=$ sur E est réflexive, transitive, symétrique et antisymétrique.
- Les relations $\leqslant$ sur $\mathbb{R}$ et $\mathbb{R}^{\mathbb{R}}$ sont réflexives, transitives et antisymétriques. Elles ne sont pas symétriques car par exemple $1 \leqslant 2$ mais $2 \not\leqslant 1$, et de même $(x \longmapsto 1) \leqslant (x \longmapsto 2)$ mais $(x \longmapsto 2) \not\leqslant (x \longmapsto 1)$.
- La relation $<$ sur $\mathbb{R}$ est transitive et antisymétrique, mais elle n'est ni réflexive, ni symétrique.
- La relation « avoir le même signe » sur $\mathbb{R}^*$ est réflexive, transitive et symétrique.
- La relation d'inclusion $\subset$ sur $\mathcal{P}(E)$ est réflexive, transitive et antisymétrique.

Exemple La relation de divisibilité $|$ est réflexive et transitive sur $\mathbb{Z}$ — donc aussi sur $\mathbb{N}$. En revanche, elle est antisymétrique sur $\mathbb{N}$, mais pas sur $\mathbb{Z}$ car par exemple $-2 \mid 2$ et $2 \mid -2$ alors que $-2 \neq 2$. Plus précisément, pour tous $a, b \in \mathbb{Z}$:

$$a \mid b \text{ et } b \mid a \iff |a| = |b| \iff a = b \text{ ou } a = -b.$$

Démonstration

- **Réflexivité :** Pour tout $a \in \mathbb{Z}$, $a \mid a$ car $a = a \times 1$ avec $1 \in \mathbb{Z}$.
- **Transitivité :** Soient $a, b, c \in \mathbb{Z}$. Si $a \mid b$ et $b \mid c$, alors $b = ak$ et $c = bl$ pour certains $k, l \in \mathbb{Z}$, donc $c = a(kl)$ avec $kl \in \mathbb{Z}$, donc $a \mid c$.
- **Antisymétrie :** Soient $a, b \in \mathbb{Z}$. Montrons que $a \mid b$ et $b \mid a$ si et seulement si $|a| = |b|$. Il en découlera que la relation $|$ est antisymétrique sur $\mathbb{N}$, mais pas sur $\mathbb{Z}$.

  Si $|a| = |b|$, alors $a = b$ ou $a = -b$, donc en effet, $a \mid b$ et $b \mid a$. Réciproquement, faisons l'hypothèse que $a \mid b$ et $b \mid a$. Ainsi, $b = ak$ et $a = bl$ pour certains $k, l \in \mathbb{Z}$, donc $b = b(kl)$.

  — Si $b = 0$, alors $a = bl = 0$, donc $|a| = 0 = |b|$.

  — Si au contraire $b \neq 0$, alors $kl = 1$, donc soit $k = l = 1$, soit $k = l = -1$ car k et l sont des entiers. Ainsi $a = \pm b$, i.e. $|a| = |b|$.

  Dans les deux cas, $|a| = |b|$.

1.2 Relations d'ordre

Définition (Relation d'ordre, relation d'ordre totale) On appelle (*relation d'*)*ordre sur* E toute relation binaire sur E à la fois réflexive, transitive et antisymétrique. Les relations d'ordre sont généralement notées $\leqslant$, $\preccurlyeq$, $\lesssim$, $\precsim$…

On dit qu'une relation d'ordre $\preccurlyeq$ sur E est *totale* si : $\forall x, y \in E, \quad x \mathrel{\mathscr{R}} y \quad \text{ou} \quad y \mathrel{\mathscr{R}} x$. Dans le cas contraire, on dit que $\preccurlyeq$ est *partielle*.

Les exemples qui suivent suggèrent que les relations d'ordre traduisent mathématiquement certaines formes de ce qu'on appelle hiérarchie en français courant. La relation $x \preccurlyeq y$ est généralement lue « x est plus petit que y », mais rien ne s'oppose à ce qu'on la lise « x est plus grand que y », il faut juste choisir une fois pour toutes. Être plus grand en termes de vieillesse, c'est être plus petit en termes de jeunesse.

Exemple

- La relation $\leqslant$ est une relation d'ordre totale sur $\mathbb{R}$.
- La relation $\leqslant$ est une relation d'ordre partielle sur l'ensemble $\mathbb{R}^{\mathbb{R}}$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Par exemple, les fonctions cosinus et sinus ne sont pas comparables : $\sin \not\leqslant \cos$ et $\cos \not\leqslant \sin$.
- La relation d'inclusion $\subset$ est une relation d'ordre sur $\mathscr{P}(E)$, partielle dès que E contient au moins deux éléments. En effet, pour tous a et b éléments DISTINCTS de E : $\{a\} \not\subset \{b\}$ et $\{b\} \not\subset \{a\}$.
- La relation de divisibilité $|$ n'est pas une relation d'ordre sur $\mathbb{Z}$, mais c'en est une sur $\mathbb{N}$, partielle car 2 et 3 ne sont comparables : $2 \nmid 3$ et $3 \nmid 2$.

Toute relation d'ordre hiérarchise les éléments qu'elle compare, mais qu'attendons-nous intuitivement d'une hiérarchie ?

— Essentiellement la transitivité, c'est ce qui compte le plus. Si A est plus grand que B et B plus grand que C, alors A est plus grand que C. Toute entorse à la transitivité contredit l'idée d'une hiérarchie.

— La réflexivité est imposée dans la définition des relations d'ordre, mais elle aurait pu ne pas l'être. L'exiger revient simplement à privilégier les relations d'infériorité au sens large aux relations d'infériorité stricte.

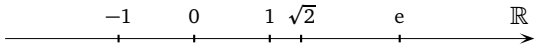
— L'antisymétrie est un autre choix conventionnel. La relation « être plus âgé au sens large » est transitive et réflexive sur l'ensemble des êtres humains, mais pas antisymétrique car deux individus peuvent être nés au même instant. Bien que non antisymétrique, cette relation a pour nous le parfum d'une hiérarchie.

En résumé, les relations d'ordre sont des exemples de hiérarchies, mais elles ne formalisent pas toutes les hiérarchies concevables.

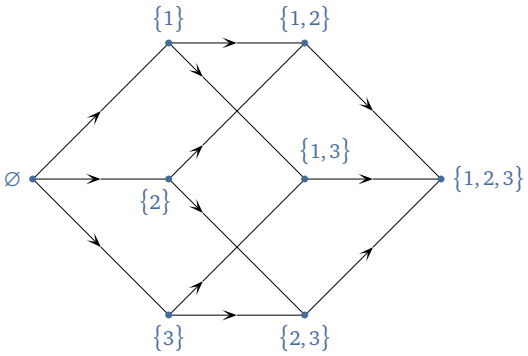
Observons à présent que les relations d'ordre excluent les boucles. Une boucle de la forme $x_1 \preccurlyeq x_2 \preccurlyeq x_3 \preccurlyeq \dots \preccurlyeq x_n \preccurlyeq x_1$ entre des éléments distincts avec $n \geqslant 2$ est inconcevable car la transitivité et l'antisymétrie forcent l'égalité $x_1 = x_2 = \dots = x_n$.

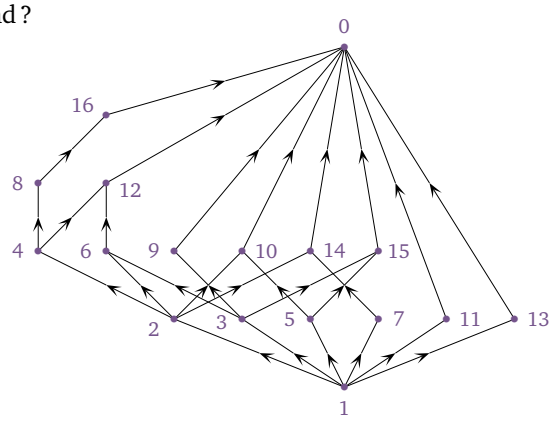
Sans boucles, les relations d'ordre ont comme une orientation naturelle. De même que les fleuves et les rivières coulent en direction de la mer sans retour en arrière, on va toujours de l'avant quand on parcourt une relation d'ordre, on ne tourne jamais en rond et c'est ça qui nous fait dire que certains éléments sont plus petits/grands que d'autres.

Ainsi, c'est parce que la relation d'ordre $\leqslant$ sur $\mathbb{R}$ est totale que nous nous représentons $\mathbb{R}$ comme une ligne unique et non pas comme un réseau fluvial compliqué doté d'une foule d'affluents.



On a représenté ci-dessous la relation $\subset$ sur l'ensemble des parties de $\{1,2,3\}$ à gauche et la relation de divisibilité $|$ sur $[\![0,16]\!]$ à droite. Le fait que ces relations ne soient pas totales se visualise bien, il ne suffit pas d'une fibre pour représenter ces relations, il en faut plusieurs. La relation $|$ donne à $\mathbb{N}$ une structure géométrique nettement plus compliquée que la relation totale $\leqslant$. Cette complexité est à la fois le charme et la difficulté de l'arithmétique. Aviez-vous remarqué que 1 est le plus petit des entiers naturels au sens de la divisibilité et que 0 en est le plus grand ?





Définition-théorème **(Relation stricte associée à une relation d'ordre)** Soit $\preccurlyeq$ une relation d'ordre sur E. La relation $\prec$ sur E définie pour tous $x, y \in E$ par l'équivalence : $x \prec y \iff x \preccurlyeq y$ et $x \neq y$ est transitive et antisymétrique. On l'appelle la *relation stricte associée à* $\preccurlyeq$.

Démonstration

- **Transitivité :** Soient $x, y, z \in E$. Supposons $x \prec y$ et $y \prec z$. En particulier $x \preccurlyeq y$ et $y \preccurlyeq z$, donc $x \preccurlyeq z$ par transitivité de $\preccurlyeq$. L'égalité $x = z$ est-elle possible ? Le cas échéant, $x \preccurlyeq y$ et $y \preccurlyeq x$, donc $x = y$ par antisymétrie de $\preccurlyeq$ — contradiction. Bref, $x \preccurlyeq z$ et $x \neq z$, i.e. $x \prec z$.
- **Antisymétrie :** Soient $x, y \in E$. Si $x \prec y$ et $y \prec x$, alors $x \preccurlyeq y$ et $y \preccurlyeq x$, donc $x = y$ par antisymétrie de $\preccurlyeq$.

Exemple Naturellement, la relation usuelle $<$ sur $\mathbb{R}$ est la relation stricte de la relation $\leqslant$.

1.3 Relations d'équivalence

Définition **(Relation d'équivalence)** On appelle *relation d'équivalence sur* E toute relation binaire sur E à la fois réflexive, transitive et symétrique. Les relations d'équivalence sont généralement notées $\sim$, $\simeq$, $\approx$, $\equiv$…

Exemple La relation « avoir le même signe » sur $\mathbb{R}^*$ est une relation d'équivalence, de même que la relation d'égalité $=$ sur E pour tout ensemble E.

Exemple Pour tout $\alpha \in \mathbb{R}$, la relation $\equiv\ [\alpha]$ de congruence modulo α sur $\mathbb{R}$ est une relation d'équivalence.
De même, pour tout $n \in \mathbb{N}$, la relation $\equiv\ [n]$ de congruence modulo n sur $\mathbb{Z}$ est une relation d'équivalence.

Démonstration

- **Réflexivité :** Pour tout $x \in \mathbb{R}$, $x = x + 0 \times \alpha$ avec $0 \in \mathbb{Z}$, donc $x \equiv x\ [\alpha]$.
- **Transitivité :** Soient $x, y, z \in \mathbb{R}$. Si $x \equiv y\ [\alpha]$ et $y \equiv z\ [\alpha]$, alors $x = y + k\alpha$ et $y = z + l\alpha$ pour certains $k, l \in \mathbb{Z}$, donc $x = z + (k+l)\alpha$ avec $k + l \in \mathbb{Z}$, donc $x \equiv z\ [\alpha]$.
- **Symétrie :** Soient $x, y \in \mathbb{R}$. Si $x \equiv y\ [\alpha]$, alors $x = y + k\alpha$ pour un certain $k \in \mathbb{Z}$, donc $y = x + (-k)\alpha$ avec $-k \in \mathbb{Z}$, donc $y \equiv x\ [\alpha]$.

Définition-théorème (Classes d'équivalence, ensemble quotient, ensemble de représentants) Soit $\sim$ une relation d'équivalence sur E.

- **Classes d'équivalence :** Pour tout $x \in E$, l'ensemble $\mathrm{cl}(x) = \{y \in E \mid \quad y \sim x\}$ est appelé la *classe d'équivalence de x (pour $\sim$)*. Les classes d'équivalence de $\sim$ forment une partition de E et pour tous $x, y \in E$:
$$\mathrm{cl}(x) = \mathrm{cl}(y) \quad \Longleftrightarrow \quad x \sim y.$$
- **Ensemble quotient :** L'ensemble des classes d'équivalences de $\sim$ est appelé l'*ensemble quotient de E par $\sim$* et souvent noté $E/\sim$.
- **Ensemble de représentants des classes d'équivalence :** On appelle *ensemble de représentants des classes d'équivalence de $\sim$* toute partie R de E qui contient un et un seul élément de chaque classe d'équivalence de $\sim$.

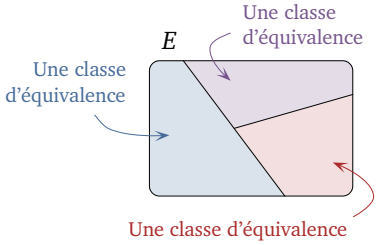
  Le cas échéant : $E = \underbrace{\bigcup_{x\in E} \mathrm{cl}(x) = \bigsqcup_{x\in R} \mathrm{cl}(x)}$ et $E/\sim = \underbrace{\{\mathrm{cl}(x) \mid \quad x \in E\} = \{\mathrm{cl}(x) \mid \quad x \in R\}}$.

  La réunion de droite est disjointe par définition de R, mais pas celle de gauche.

  Par définition de R, chaque classe apparaît ici UNE ET UNE SEULE FOIS sans redondance.

L'ensemble quotient $E/\sim$ de E par $\sim$ est un ENSEMBLE D'ENSEMBLES, en l'occurrence un ensemble de parties de E.

Ce que ce théorème raconte, c'est que la relation d'équivalence $\sim$ peut être représentée comme une carte au sens géographique. Chaque classe d'équivalence est comme un pays à l'intérieur de E dont les éléments sont caractérisés par une nationalité. Le monde E se trouve ainsi partitionné en pays et se donner un ensemble de représentants des classes d'équivalence de $\sim$, c'est ni plus ni moins se donner un ambassadeur et un seul par pays.



Démonstration

- Soient $x, y \in E$. Montrons que : $\mathrm{cl}(x) = \mathrm{cl}(y) \iff \mathrm{cl}(x) \cap \mathrm{cl}(y) \neq \varnothing \iff x \sim y$.
  - Si $\mathrm{cl}(x) = \mathrm{cl}(y)$, l'ensemble $\mathrm{cl}(x) \cap \mathrm{cl}(y)$ contient x, donc est non vide.
  - Si $\mathrm{cl}(x) \cap \mathrm{cl}(y) \neq \varnothing$, alors en notant z un élément de cette intersection, $z \sim x$ et $z \sim y$, donc $x \sim y$ par symétrie et transitivité.
  - Supposons que $x \sim y$ et montrons que $\mathrm{cl}(x) = \mathrm{cl}(y)$. Par symétrie, il suffit de montrer que $\mathrm{cl}(x) \subset \mathrm{cl}(y)$. Soit $z \in \mathrm{cl}(x)$. Ainsi $z \sim x$, or par hypothèse $x \sim y$, donc $z \sim y$ par transitivité, i.e. $z \in \mathrm{cl}(y)$.
- Montrons maintenant que les classes d'équivalence de $\sim$ forment une partition de E.
  - Pour tout $x \in E$, $x \sim x$ par réflexivité, donc $\mathrm{cl}(x)$ contient x, donc $\mathrm{cl}(x)$ est non vide. Rappelons que cette condition de non-vacuité différencie les partitions des recouvrements disjoints.
  - Ensuite, $E \subset \bigcup_{x\in E} \mathrm{cl}(x)$ car $x \in \mathrm{cl}(x)$ pour tout $x \in E$. L'inclusion réciproque étant évidente, $E = \bigcup_{x\in E} \mathrm{cl}(x)$.
  - Pour finir, soient $x, y \in E$. Si $x \sim y$, alors $\mathrm{cl}(x) = \mathrm{cl}(y)$ comme on l'a vu, et sinon $\mathrm{cl}(x) \cap \mathrm{cl}(y) = \varnothing$.

Exemple Notons $\mathscr{S}$ la relation d'équivalence « avoir le même signe » sur $\mathbb{R}^*$. Tout réel non nul a soit le même signe que 1, soit le même signe que -1, et bien sûr 1 et -1 n'ont pas le même signe, donc $\mathscr{S}$ possède deux classes d'équivalence, à savoir $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$, et $\{1,-1\}$ en est un ensemble de représentants. Pour finir, $E/\mathscr{S} = \{\mathbb{R}_+^*, \mathbb{R}_-^*\}$.

Exemple Soit $n \in \mathbb{N}^*$. On s'intéresse dans cet exemple à la relation $\equiv [n]$ de congruence modulo n sur $\mathbb{Z}$.

- Pour tout $x \in \mathbb{Z}$, la classe d'équivalence de x pour $\equiv [n]$ est l'ensemble :
$$\{y \in \mathbb{Z} \mid \quad y \equiv x\ [n]\} = \{y \in \mathbb{Z} \mid \quad \exists k \in \mathbb{Z}, \quad y = x + kn\} = \{x + kn \mid \quad k \in \mathbb{Z}\} \qquad \text{noté } x + n\mathbb{Z}.$$
  Concrètement : $x + n\mathbb{Z} = \{\ldots\ ,\ x-n\ ,\ x\ ,\ x+n\ ,\ x+2n\ ,\ x+3n\ ,\ \ldots\}$, donc :
$$\mathbb{Z}/\equiv [n] = \{x + n\mathbb{Z} \mid \quad x \in \mathbb{Z}\}.$$
  Cette description de $\mathbb{Z}/\equiv [n]$ souffre hélas d'un gros défaut, chaque classe y est répétée un grand nombre de fois. Par exemple : $1+n\mathbb{Z} = (n+1)+n\mathbb{Z} = (2n+1)+n\mathbb{Z}$. Tâchons de déterminer un ensemble de représentants des classes d'équivalence de $\equiv [n]$ pour supprimer ces redondances.
- Or d'après le théorème de la division euclidienne : $\forall x \in \mathbb{Z}, \quad \exists!\,(q,r) \in \mathbb{Z}\times\mathbb{Z}, \quad x = nq + r$ et $r \in [\![0, n-1]\!]$, ce qu'on peut aussi écrire ainsi : $\forall x \in \mathbb{Z}, \quad \exists!\, r \in [\![0, n-1]\!], \quad x \equiv r\ [n]$. Le théorème de la division euclidienne affirme donc que tout entier relatif est congru modulo n à un et un seul élément de $[\![0, n-1]\!]$. Par conséquent, $[\![0, n-1]\!]$ est un ensemble de représentants des classes d'équivalence de $\equiv [n]$, donc $\mathbb{Z}/\equiv [n] = \{n\mathbb{Z}, n\mathbb{Z}+1, \ldots, n\mathbb{Z}+n-1\}$.

Exemple Soit $\alpha > 0$. On s'intéresse dans cet exemple à la relation $\equiv[\alpha]$ de congruence modulo α sur $\mathbb{R}$.

- Pour tout $x \in \mathbb{R}$, la classe d'équivalence de x pour $\equiv[\alpha]$ est l'ensemble :
$$\Big\{y \in \mathbb{R} \mid \quad y \equiv x\ [\alpha]\Big\} = \Big\{y \in \mathbb{R} \mid \quad \exists k \in \mathbb{Z}, \quad y = x + k\alpha\Big\} = \Big\{x + k\alpha \mid \quad k \in \mathbb{Z}\Big\} \qquad \text{noté } x + \alpha\mathbb{Z}.$$
Par conséquent : $\mathbb{R}/_{\equiv[\alpha]} = \big\{x + \alpha\mathbb{Z} \mid \quad x \in \mathbb{R}\big\}$, mais de nouveau, que de répétitions dans cette description ! Par exemple : $\alpha\mathbb{Z} = -\alpha + \alpha\mathbb{Z} = \alpha + \alpha\mathbb{Z}$. Tâchons de déterminer un ensemble de représentants des classes d'équivalence de $\equiv[\alpha]$ pour supprimer ces redondances.
- Or par définition de la partie entière : $\forall x \in \mathbb{R}, \quad \exists!(k,\varepsilon) \in \mathbb{Z} \times [0,1[, \quad x = k + \varepsilon$, et quitte à remplacer x par $\frac{x}{\alpha}$: $\forall x \in \mathbb{R}, \quad \exists!(k,\varepsilon) \in \mathbb{Z} \times [0,1[, \quad \frac{x}{\alpha} = k + \varepsilon$, donc après multiplication par α : $\forall x \in \mathbb{R}, \quad \exists!\, \varepsilon \in [0,\alpha[, \quad x \equiv \varepsilon\ [\alpha]$. Tout réel est ainsi congru modulo α à un et un seul élément de $[0,\alpha[$, donc $[0,\alpha[$ est un ensemble de représentants des classes d'équivalence de $\equiv[\alpha]$, donc $\mathbb{R}/_{\equiv[\alpha]} = \Big\{x + \alpha\mathbb{Z} \mid \quad x \in [0,\alpha[\Big\}$.

Finalement, toute relation d'équivalence peut être exprimée en français sous la forme « Avoir le même machin » où le « machin » est une sorte d'instrument de mesure utilisé pour comparer les éléments. Par exemple, « avoir le même signe », « avoir le même reste de division euclidienne par n », « avoir la même partie fractionnaire », etc.

2 Applications

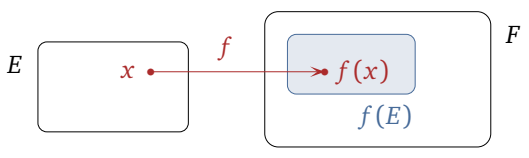
2.1 Premières définitions, exemples

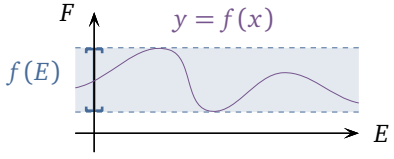
Qu'est-ce qu'une fonction ? On se contente généralement de dire ce qu'une fonction FAIT pour éviter d'avoir à dire quel objet mathématique elle EST : « Une fonction associe à tout élément d'un ensemble un unique élément d'un autre ensemble. » Ceci hélas n'est pas une définition. Intuitivement, une fonction est une courbe, un graphe. Par exemple, la fonction $x \longmapsto x^2$ peut être vue comme l'ensemble des points du plan de coordonnées $\big(x, x^2\big)$, x décrivant $\mathbb{R}$. On vous a peut-être expliqué qu'il ne faut pas confondre une fonction et sa courbe représentative, mais avec la définition qui suit au contraire, plus anglo-saxonne que française, toute fonction EST son graphe.

Définition (Application/fonction, image et antécédents)

- **Application/fonction :** On appelle *application* (ou *fonction*) *de E dans F* toute partie f de $E \times F$ pour laquelle :
$$\forall x \in E, \quad \exists!\, y \in F, \quad (x,y) \in f.$$
La présence du pseudo-quantificateur $\exists!$ permet de noter $f(x)$ l'unique $y \in F$ de la proposition ci-dessus. La proposition $(x,y) \in f$ sera dès lors toujours notée $y = f(x)$.
- **Image/antécédents :** Pour tous $x \in E$ et $y \in F$, 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*.
- **Ensemble de définition, ensemble d'arrivée, image :**
  - E est appelé l'*ensemble de définition* (ou *de départ*) *de f* et F un *ensemble d'arrivée de f*.
  - L'ensemble $f(E) = \Big\{y \in F \mid \quad \exists x \in E, \quad y = f(x)\Big\} = \Big\{f(x) \mid \quad x \in E\Big\}$ est appelé l'*image de f*. Ses éléments sont appelés les *valeurs de f*.
- **Expression « à valeurs dans... » :** Soit B une partie de F. On dit que f est *à valeurs dans B* si toute valeur de f est élément de B, i.e. si : $\forall x \in E, \quad f(x) \in B$, ou encore si l'image de f est incluse dans B : $f(E) \subset B$.

On représente classiquement les applications de deux façons — soit au moyen de « patates » (figure de gauche), soit au moyen d'un graphe (figure de droite), et ce même si les ensembles E et F ne sont pas des parties de $\mathbb{R}$. À droite, pour représenter l'image $f(E)$ de f, on projette simplement le graphe de f sur l'axe des ordonnées.





Attention ! En général, l'image de f est plus petite que F !

Exemple La fonction carré $x \longmapsto x^2$ est définie sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}$, mais son image est (seulement) $\mathbb{R}_+$. La fonction cosinus est définie sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}$, mais son image est (seulement) $[-1,1]$.

Exemple Soit A une partie de E. L'application $X \overset{f}{\longmapsto} X \cap A$ définie sur $\mathscr{P}(E)$ est à valeurs dans $\mathscr{P}(E)$, mais son image est (seulement) $\mathscr{P}(A)$.

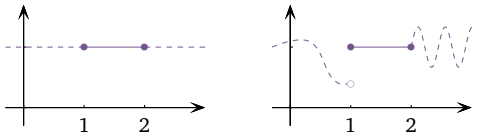
Démonstration Pour tout $X \in \mathscr{P}(E)$, $f(X) = X \cup A$ est une partie de A, donc $f\big(\mathscr{P}(E)\big) \subset \mathscr{P}(A)$. Inversement, pour tout $Y \in \mathscr{P}(A)$: $\quad Y = Y \cap A = f(Y) \in \mathscr{P}(E), \quad$ donc $\mathscr{P}(A) \subset f\big(\mathscr{P}(E)\big)$.

Définition (Restriction et prolongements) Soit A une partie de E.

- **Restriction :** Soient $f : E \longrightarrow F$ une application. On appelle *restriction de f à A*, notée $f_{|A}$, l'application définie sur A SEULEMENT par la relation $f_{|A}(x) = f(x)$ pour tout $x \in A$.
- **Prolongements :** Soit $f : A \longrightarrow F$ une application. On appelle *prolongement de f à E* toute application g de E dans F pour laquelle $g_{|A} = f$, i.e. pour laquelle $f(x) = g(x)$ pour tout $x \in A$.

Restreindre/prolonger une application, c'est diminuer/augmenter la taille de son ensemble de définition.

Attention ! Toute application possède beaucoup de prolongements, on parle toujours d'UN prolongement et non « du » prolongement. Les figures ci-contre représentent deux prolongements de la fonction $x \longmapsto 1$ définie sur $[1,2]$.



Définition (Ensembles d'applications) L'ensemble des applications de E dans F est noté F^E ou $\mathscr{F}(E,F)$.

Exemple L'ensemble $\{0,1\}^{\{1,2\}}$ des applications de $\{1,2\}$ dans $\{0,1\}$ est de cardinal 4 et ses éléments sont les applications :

$$\begin{cases} 1 \longmapsto 0 \\ 2 \longmapsto 0, \end{cases} \quad \begin{cases} 1 \longmapsto 0 \\ 2 \longmapsto 1, \end{cases} \quad \begin{cases} 1 \longmapsto 1 \\ 2 \longmapsto 0 \end{cases} \quad \text{et} \quad \begin{cases} 1 \longmapsto 1 \\ 2 \longmapsto 1. \end{cases}$$

Définition (Famille) On appelle *famille (d'éléments) de E indexée par I* toute application de I dans E. Cela dit, au lieu d'être notée comme des applications, les familles sont généralement notées sous la forme $(x_i)_{i\in I}$.

L'ensemble des familles de E indexée par I est ainsi l'ensemble E^I.

Une famille $(x_1, \dots, x_n)$ d'éléments de E n'est rien de plus que l'application $i \longmapsto x_i$ de $[\![1,n]\!]$ dans E qui associe à chaque position i l'élément x_i qui lui correspond dans $(x_1, \dots, x_n)$.

Exemple Une suite réelle $(u_n)_{n\in\mathbb{N}}$ n'est rien de plus qu'une fonction $n \longmapsto u_n$ de $\mathbb{N}$ dans $\mathbb{R}$. L'ensemble des suites réelles est donc l'ensemble $\mathbb{R}^{\mathbb{N}}$.

Exemple Une famille d'éléments de E indexée par l'ensemble vide est une application de $\varnothing$ dans E et ça existe ! Par définition d'une application, l'ensemble vide une application de $\varnothing$ dans E et c'est la seule, appelée la *famille vide de E*.

Définition-théorème (Identité, composition)

- **Identité :** On appelle (*application*) *identité de E*, notée Id_E, l'application « qui ne fait rien » $\begin{cases} E & \longrightarrow & E \\ x & \longmapsto & x. \end{cases}$
- **Composition :** Soient $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications. On appelle *composée de f suivie de g* l'application $g \circ f$ définie par la relation $g \circ f(x) = g\big(f(x)\big)$ pour tout $x \in E$.

  La composition est *associative*, autrement dit pour toutes applications $f : E \longrightarrow F$, $g : F \longrightarrow G$ et $h : G \longrightarrow H$:

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

  En outre, pour toute application $f : E \longrightarrow F$: $\quad \mathrm{Id}_F \circ f = f \circ \mathrm{Id}_E = f \quad$ (*neutralité* de l'identité).

Pour donner un sens à $g\big(f(x)\big)$ pour tout $x \in E$, on doit absolument garantir que $f(x) \in F$, autrement dit que f est à valeurs dans F. La bonne définition de $g \circ f$ requiert que l'image de f soit incluse dans l'ensemble de définition de g.

✖ **Attention !** En général, la composition n'est possible que dans un seul sens, et quand elle est possible dans les deux, f et g n'ont aucune raison de commuter. Par exemple, si f est la fonction $x \longmapsto x^2$ et g la fonction $x \longmapsto x^2+1$, $g \circ f$ est la fonction $x \longmapsto \left(x^2\right)^2+1 = x^4+1$ et $f \circ g$ la fonction $x \longmapsto \left(x^2+1\right)^2 = x^4 + \boxed{+2x^2} + 1$, donc $g \circ f \neq f \circ g$.

Exemple La fonction $x \longmapsto \sqrt{1-x^2}$ est définie sur $[-1,1]$.

Démonstration La fonction $x \longmapsto 1-x^2$ est définie sur $\mathbb{R}$ et la fonction $\sqrt{\cdot}$ l'est sur $\mathbb{R}_+$, mais quand x décrit $\mathbb{R}$, $1-x^2$ n'appartient pas forcément à $\mathbb{R}_+$. Pour quels $x \in \mathbb{R}$ est-il vrai que $1-x^2 \geqslant 0$? Réponse : $x \in [-1,1]$ car les racines du polynôme sont -1 et 1 et son coefficient dominant est strictement négatif.

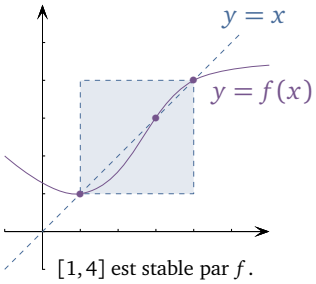
Dans la définition qui suit, les ensembles de départ et d'arrivée de f coïncident.

Définition (Itérées, point fixe, partie stable) Soient $f : E \longrightarrow E$ une application.

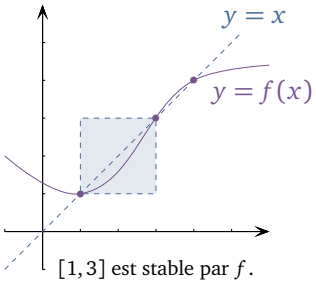
- **Itérées :** Soit $f : E \longrightarrow E$ une application. On pose $f^0 = \mathrm{Id}_E$ et $f^n = \underbrace{f \circ \ldots \circ f}_{n \text{ termes}}$ pour tout $n \in \mathbb{N}$. Les applications ainsi définies sont appelées les *itérées de* f.
- **Point fixe :** On appelle *point fixe de* f tout élément $x \in E$ pour lequel $f(x) = x$.
- **Partie stable :** Soit A une partie de E. On dit que A est *stable par* f si $f(A) \subset A$, i.e. si : $\quad \forall x \in A, \quad f(x) \in A$.

Exemple

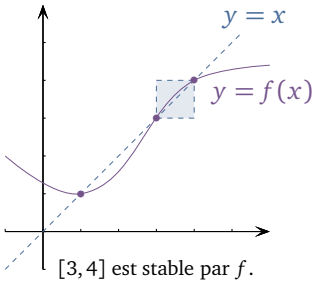
- L'intervalle $\mathbb{R}_+^*$ est stable par la fonction inverse $x \longmapsto \dfrac{1}{x}$ car pour tout $x > 0$: $\quad \dfrac{1}{x} > 0$.
- L'intervalle $[0,1]$ est stable par $x \longmapsto \sqrt{1-x}$ car pour tout $x \in [0,1]$, : $\quad 0 \leqslant 1-x \leqslant 1, \quad$ donc $0 \leqslant \sqrt{1-x} \leqslant 1$ par croissance de la fonction $\sqrt{\cdot}$.
- Sur les figures ci-dessous, les intervalles $[1,4]$, $[1,3]$, $[3,4]$ et $[4,+\infty[$ sont stables par f, mais $]-\infty,1]$ ne l'est pas.



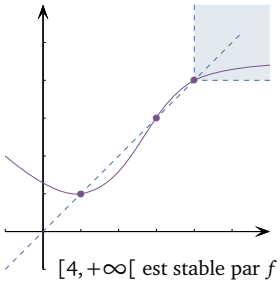
$[1,4]$ est stable par f.



$[1,3]$ est stable par f.



$[3,4]$ est stable par f.


$[4,+\infty[$ est stable par f.

Dernier exemple de ce paragraphe, celui des *indicatrices*. Connaître une partie A de E, c'est savoir pour chaque élément de E s'il appartient à A ou non. Si on représente l'appartenance par l'entier 1 et la non-appartenance par l'entier 0, connaître A revient donc à connaître l'application $\mathbb{1}_A$ de E dans A définie ci-dessous.

Définition-théorème (Indicatrice d'une partie) Soit A une partie de E. Pour tout $x \in E$, on pose :

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \in E \setminus A. \end{cases}$$

La fonction $\mathbb{1}_A$ ainsi définie de E dans $\{0,1\}$ est appelée l'*indicatrice de* A (*sur* E).

Il est facile de récupérer A quand on connaît $\mathbb{1}_A$: $\quad A = \left\{x \in E \mid \quad \mathbb{1}_A(x) = 1\right\}$.

(i) **Inclusion :** $\quad A \subset B \iff \mathbb{1}_A \leqslant \mathbb{1}_B$. $\qquad$ **Égalité :** $\quad A = B \iff \mathbb{1}_A = \mathbb{1}_B$.

(ii) **Complémentaire :** $\quad \mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A$. $\qquad$ **Intersection :** $\quad \mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B$. $\qquad$ **Réunion :** $\quad \mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B$.

Démonstration

(ii) **Inclusion :**

- Faisons l'hypothèse que $A \subset B$ et montrons que $\mathbb{1}_A \leqslant \mathbb{1}_B$, i.e. que $\mathbb{1}_A(x) \leqslant \mathbb{1}_B(x)$ pour tout $x \in E$. Soit $x \in E$. Si $x \in A$, alors $x \in B$ car $A \subset B$, donc $\mathbb{1}_A(x) = 1 \leqslant 1 = \mathbb{1}_B(x)$. Et si $x \in E \setminus A$, alors $\mathbb{1}_A(x) = 0 \leqslant \mathbb{1}_B(x)$ quelle que soit la valeur de $\mathbb{1}_B(x)$. Dans les deux cas, $\mathbb{1}_A(x) \leqslant \mathbb{1}_B(x)$.
- Réciproquement, faisons l'hypothèse que $\mathbb{1}_A \leqslant \mathbb{1}_B$ et montrons que $A \subset B$. Pour tout $x \in A$, $\mathbb{1}_A(x) = 1$ et par ailleurs $\mathbb{1}_A(x) \leqslant \mathbb{1}_B(x)$, donc $\mathbb{1}_B(x) = 1$, i.e. $x \in B$.

Égalité : $A=B \iff \big(A\subset B \text{ et } B\subset A\big) \iff \big(\mathbb{1}_A\leqslant\mathbb{1}_B \text{ et } \mathbb{1}_B\leqslant\mathbb{1}_A\big) \iff \mathbb{1}_A=\mathbb{1}_B$.

(ii) **Complémentaire :** Pour tout $x\in A$: $\mathbb{1}_{\overline{A}}(x)=0=1-1=1-\mathbb{1}_A(x)$, et pour tout $x\in E\setminus A$: $\mathbb{1}_{\overline{A}}(x)=1=1-0=1-\mathbb{1}_A(x)$. Conclusion : $\mathbb{1}_{\overline{A}}=1-\mathbb{1}_A$.

Intersection : Pour tout $x\in A\cap B$: $\mathbb{1}_{A\cap B}(x)=1=1\times 1=\mathbb{1}_A(x)\mathbb{1}_B(x)$, et pour tout $x\in E\setminus(A\cap B)$, $x\notin A$ ou $x\notin B$, donc $\mathbb{1}_A(x)=0$ ou $\mathbb{1}_B(x)=0$, donc $\mathbb{1}_{A\cap B}(x)=0=\mathbb{1}_A(x)\mathbb{1}_B(x)$. Conclusion : $\mathbb{1}_{A\cap B}=\mathbb{1}_A\mathbb{1}_B$.

Réunion : $\mathbb{1}_{A\cup B}=1-\mathbb{1}_{\overline{A\cup B}}=1-\mathbb{1}_{\overline{A}\cap\overline{B}}=1-\mathbb{1}_{\overline{A}}\mathbb{1}_{\overline{B}}=1-(1-\mathbb{1}_A)(1-\mathbb{1}_B)=\mathbb{1}_A+\mathbb{1}_B-\mathbb{1}_A\mathbb{1}_B$.

2.2 Surjections

Définition (Surjection) Soit $f : E\longrightarrow F$ une application. On dit que f est *surjective de E sur F* ou que c'est une *surjection de E sur F* si :

$$\forall y\in F,\quad \exists x\in E,\quad y=f(x),$$ ce qui revient à dire que l'image de f est égale à F : $f(E)=F$,

ou encore que tout élément de F possède **au moins** un antécédent dans E par f.

Dire que f est surjective de E sur F, c'est dire que tout élément de F s'écrit « f de quelqu'un ».

L'application f est bien sûr à valeurs dans son image $f(E)$ et tout élément de $f(E)$ possède un antécédent par f, donc...

Toute application est surjective de son ensemble de définition **sur son image.**

Attention, on ne dit pas que f est surjective de E « dans » F mais qu'elle l'est de E **sur** F, car en cas de surjectivité, f permet à E de « couvrir » entièrement F. Cette idée de couverture justifie l'emploi de la préposition « sur ».

Exemple La fonction carré $x\longmapsto x^2$ est surjective de $\mathbb{R}$ sur (son image) $\mathbb{R}_+$. La fonction cosinus est surjective de $\mathbb{R}$ sur (son image) $[-1,1]$.

Exemple L'application $(x,y)\overset{f}{\longmapsto}x+y$ est surjective de $\mathbb{R}^2$ sur $\mathbb{R}$ car tout réel possède un antécédent par f. Par exemple, $z=f(z,0)$ pour tout $z\in\mathbb{R}$, mais z a plein d'antécédents ici et on pourrait aussi dire que $z=f(z-1,1)$.

Exemple L'application $X\overset{g}{\longmapsto}X\cup\{0\}$ n'est pas surjective de $\mathscr{P}(\mathbb{N})$ sur $\mathscr{P}(\mathbb{N})$ car par exemple, $\varnothing$ n'a pas d'antécédent par g. Il n'existe en effet pas de partie X de $\mathbb{N}$ pour laquelle $\varnothing=g(X)=X\cup\{0\}$.

Théorème (Surjectivité et composition) Soient $f : E\longrightarrow F$ et $g : F\longrightarrow G$ deux applications.

(i) Si f et g sont surjectives, $g\circ f$ l'est aussi. (ii) Si $g\circ f$ est surjective, g l'est aussi.

L'assertion (ii) raconte juste que si tout élément de G s'écrit « $g\circ f$ de quelqu'un », alors tout élément de G s'écrit aussi « g de quelqu'un (d'autre) ».

Démonstration

(i) Soit $y\in G$. Montrons que y possède un antécédent par $g\circ f$. Or $y=g(t)$ pour un certain $t\in F$ par surjectivité de g, puis $t=f(x)$ pour un certain $x\in E$ par surjectivité de f, donc $y=g(t)=g\circ f(x)$.

(ii) Soit $y\in G$. Montrons que y possède un antécédent par g. Or $y=g\circ f(t)$ pour un certain $t\in E$ par surjectivité de $g\circ f$, donc $y=g(x)$ si on pose $x=f(t)\in F$.

2.3 Injections

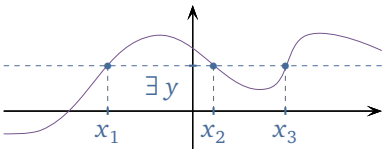
Définition (Injection) Soit $f : E\longrightarrow F$ une application. On dit que f est *injective sur E* ou que c'est une *injection sur E* si :

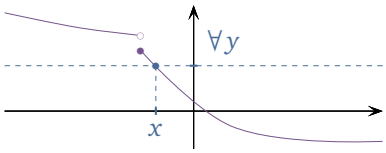
$$\forall x,x'\in E,\quad f(x)=f(x')\implies x=x',$$

ce qui revient à dire que tout élément de F possède **au plus** un antécédent dans E par f.

Plus précisément, si f est injective, les éléments de son image $f(E)$ possèdent tous exactement un antécédent par f et les éléments de $F\setminus f(E)$ n'en possèdent aucun.

f N'est PAS injective, CERTAINS y ont plusieurs antécédents.

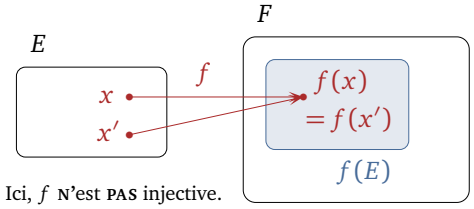




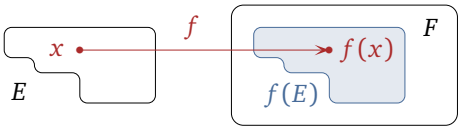
f est injective, AUCUN y n'a plusieurs antécédents, certains même n'en ont pas.

D'un point de vue calculatoire, une application injective est une application que l'on peut « simplifier » en cours de calcul. Dès que $f(x) = f(x')$, alors $x = x'$ après « simplification ».

On comprend également bien l'injectivité en contraposant sa définition. L'application f est injective lorsqu'elle donne toujours des valeurs différentes à des points différents — si $x \neq x'$, alors $f(x) \neq f(x')$.



Ici, f N'est PAS injective.



On peut aussi dire que comme f distingue à l'arrivée les éléments qui le sont au départ, l'image de f est comme une copie de E à l'intérieur de F.

Pour finir, se demander si f est injective, c'est se poser la question suivante :

> Quand je connais $f(x)$, puis-je récupérer x ?
> Si la réponse est oui pour tout x, alors f est injective.

Exemple La fonction $x \longmapsto \dfrac{x}{x-1}$ est injective sur $\mathbb{R} \setminus \{1\}$.

Démonstration Soient $x, x' \in \mathbb{R} \setminus \{1\}$. Si $\dfrac{x}{x-1} = \dfrac{x'}{x'-1}$, alors $x(x'-1) = x'(x-1)$, donc $x = x'$.

Exemple La fonction carré $x \longmapsto x^2$ n'est pas injective sur $\mathbb{R}$ car par exemple $(-1)^2 = 1^2$. Elle est en revanche injective sur $\mathbb{R}_+$ car pour tous $x, x' \in \mathbb{R}_+$, si $x^2 = x'^2$, alors $x = x'$ ou $x = -x'$, donc $x = x'$ par positivité.

Exemple L'application $(x,y) \overset{f}{\longmapsto} x + y$ n'est pas injective sur $\mathbb{R}^2$ car par exemple $f(0,1) = 1 = f(1,0)$, mais l'application $(x,y) \overset{g}{\longmapsto} (x+y, x-y)$ l'est.

Démonstration Soient $(x,y), (x',y') \in \mathbb{R}^2$. Si $g(x,y) = g(x',y')$, alors $x + y = x' + y'$ et $x - y = x' - y'$, donc $x = x'$ par demi-somme et $y = y'$ par demi-différence, et enfin $(x,y) = (x',y')$.

Exemple L'application $X \overset{h}{\longmapsto} X \cup \{0\}$ n'est pas injective sur $\mathscr{P}(\mathbb{N})$ car par exemple $h(\varnothing) = \{0\} = h(\{0\})$.

Théorème (Injectivité et composition) Soient $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications.

(i) Si f et g sont injectives, $g \circ f$ l'est aussi.

(ii) Si $g \circ f$ est injective, f l'est aussi.

Attention ! Dans l'assertion (ii), g n'a aucune raison d'être injective. Pensez par exemple aux fonctions $x \overset{f}{\longmapsto} \mathrm{e}^x$ et $x \overset{g}{\longmapsto} x^2$ de $\mathbb{R}$ dans $\mathbb{R}$. La fonction $x \overset{g\circ f}{\longmapsto} \mathrm{e}^{2x}$ est injective, mais g ne l'est pas.

Démonstration

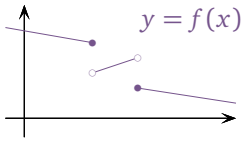
(i) Soient $x, x' \in E$. Si $g \circ f(x) = g \circ f(x')$, alors $f(x) = f(x')$ par injectivité de g, puis $x = x'$ par injectivité de f.

(ii) Soient $x, x' \in E$. Si $f(x) = f(x')$, alors $g\big(f(x)\big) = g\big(f(x')\big)$ après composition par g, donc $x = x'$ par injectivité de $g \circ f$.

Théorème (Injectivité et stricte monotonie) Pour une fois, E est une partie de $\mathbb{R}$.

Soit $f : E \longrightarrow \mathbb{R}$ une fonction. Si f est strictement monotone, f est injective.

Attention ! La réciproque est fausse en général ! Sur la figure ci-contre, f est injective, mais pas du tout monotone. Et oui, elle n'est pas continue. Nous verrons plus tard qu'une fonction injective continue sur un intervalle y est toujours strictement monotone.



Démonstration Supposons f strictement croissante et montrons que f est injective.
Soient $x, x' \in E$ deux réels pour lesquels $f(x) = f(x')$. Si $x < x'$, alors $f(x) < f(x')$ par stricte croissance, et si $x > x'$, alors de même $f(x) > f(x')$. Bref, $x = x'$.

Si f est strictement décroissante, alors $-f$ est strictement croissante, donc injective comme on vient de le voir, donc $f = (x \longmapsto -x) \circ (-f)$ aussi par composition. ∎

2.4 Bijections, réciproques

Définition **(Réciproque)** Soit $f : E \longrightarrow F$ une application. On appelle *réciproque de f sur F* toute application $g : F \longrightarrow E$ pour laquelle $g \circ f = \mathrm{Id}_E$ et $f \circ g = \mathrm{Id}_F$.

En termes simples, g défait le travail effectué par f — et vice versa. Ce que l'une tricote, l'autre le détricote.

Exemple Les fonctions $\left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & x^2 \end{array} \right.$ et $\left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & \sqrt{x} \end{array} \right.$ sont réciproques l'une de l'autre car pour tout $x \geqslant 0$:
$$\left(\sqrt{x}\right)^2 = x \quad \text{et} \quad \sqrt{x^2} = |x| = x.$$

Définition-théorème **(Bijection)** Soit $f : E \longrightarrow F$ une application. Les assertions suivantes sont équivalentes :

(i) Tout élément de F possède UN ET UN SEUL antécédent dans E par f : $\quad \forall y \in F, \quad \exists ! \, x \in E, \quad y = f(x)$.

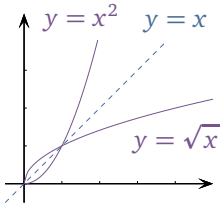
(ii) f est injective sur E et surjective de E sur F.

(iii) f possède une réciproque sur F.

Le cas échéant, on dit que f est *bijective de E SUR F* ou que c'est une *bijection de E SUR F*.

En outre, f ne possède alors qu'une seule réciproque, notée f^{-1}. Pour tous $x \in E$ et $y \in F$:
$$y = f(x) \quad \Longleftrightarrow \quad x = f^{-1}(y).$$
Dans le cas d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$, cette équivalence signifie géométriquement que le graphe de f et celui de f^{-1} sont symétriques l'un de l'autre par rapport à la droite d'équation $y = x$.

Comme pour la surjectivité, on ne dit pas que f est bijective de E « dans » F mais qu'elle l'est de E SUR F.



La symétrie des graphes de f et f^{-1} par rapport à la droite d'équation $y = x$ se visualise aisément sur les graphes des fonctions $x \longmapsto x^2$ et $x \longmapsto \sqrt{x}$ sur $\mathbb{R}_+$. Un point de coordonnées (x, y) appartient au graphe de la fonction carré sur $\mathbb{R}_+$ si et seulement si $y = x^2$, si et seulement si $x = \sqrt{y}$, si et seulement si le point de coordonnées (y, x) appartient au graphe de la fonction $\sqrt{\cdot}$.

Démonstration

(iii) $\Longrightarrow$ (ii) Si f possède une réciproque g sur F, alors $g \circ f = \mathrm{Id}_E$ est injective sur E, donc f l'est sur E, et $f \circ g = \mathrm{Id}_F$ est surjective de F sur F, donc f l'est de E sur F.

(ii) $\Longrightarrow$ (i) Si f est injective sur E et surjective de E sur F, tout élément de F possède au moins un antécédent par f par surjectivité et au plus un par injectivité, donc exactement un antécédent par f.

(i) $\Longrightarrow$ (iii) Si tout élément y de F possède un et un seul antécédent x dans E par f, nous pouvons sans ambiguïté noter $g(y)$ cet unique x et définir ainsi une application g de F dans E.

— Pour tout $x \in E$, x est l'unique antécédent de $f(x)$ par f, donc $g\big(f(x)\big) = x$. Ainsi $g \circ f = \mathrm{Id}_E$.

— Pour tout $y \in F$, $g(y)$ est l'unique antécédent de y par f, donc $f\big(g(y)\big) = y$. Ainsi $f \circ g = \mathrm{Id}_F$.

Conclusion : g est une réciproque de f sur F. ∎

Exemple L'application Id_E est bijective de E sur E de réciproque elle-même car $\mathrm{Id}_E \circ \mathrm{Id}_E = \mathrm{Id}_E$.

Exemple Soient $a \in \mathbb{R}^*$ et $b \in \mathbb{R}$. La fonction $x \longmapsto ax + b$ est bijective de $\mathbb{R}$ sur $\mathbb{R}$ de réciproque $x \longmapsto \dfrac{x - b}{a}$.

Démonstration Il suffit de montrer que les fonctions $x \overset{f}{\longmapsto} ax + b$ et $x \overset{g}{\longmapsto} \dfrac{x - b}{a}$ de $\mathbb{R}$ dans $\mathbb{R}$ sont réciproques l'une de l'autre. Or pour tout $x \in \mathbb{R}$: $\quad g \circ f(x) = \dfrac{(ax + b) - b}{a} = x \quad$ et $\quad f \circ g(x) = a \times \dfrac{x - b}{a} + b = x$.

Exemple Soit $f : E \longrightarrow E$ une *involution de E*, i.e. une application pour laquelle $f \circ f = \mathrm{Id}_E$. Alors f est une bijection et $f^{-1} = f$.

Exemple Soient $\sim$ une relation d'équivalence sur E et R un ensemble de représentants des classes d'équivalence de $\sim$. L'application $x \overset{\gamma}{\longmapsto} \mathrm{cl}(x)$ est bijective de R sur $E/\sim$.

Démonstration D'abord, γ est à valeurs dans $E/\sim$. Ensuite, toute classe d'équivalence C de $\sim$ contient un et un seul élément $\rho(C)$ de R. Montrons simplement que γ et ρ sont réciproques l'une de l'autre.

— Pour tout $x \in R$, $R \cap \mathrm{cl}(x)$ est un singleton contenant x et $\rho\big(\mathrm{cl}(x)\big)$, donc $\rho\big(\mathrm{cl}(x)\big) = x$, et ainsi $\rho \circ \gamma = \mathrm{Id}_R$.

— Pour toute classe $C \in E/\sim$, $\rho(C)$ appartient à C, donc $\mathrm{cl}\big(\rho(C)\big) = C$, autrement dit $\gamma \circ \rho(C) = C$, et ainsi $\gamma \circ \rho = \mathrm{Id}_{E/\sim}$.

Définition (Permutation, groupe symétrique) On appelle *permutation de E* toute bijection de E sur E.

L'ensemble des permutations de E est noté S_E et appelé le *groupe symétrique de E*. Si $E = [\![1,n]\!]$ pour un certain $n \in \mathbb{N}^*$, S_E est plutôt noté S_n.

Exemple Le groupe symétrique S_2 est de cardinal 2, il contient seulement les applications $\mathrm{Id}_{[\![1,2]\!]}$ et $\begin{cases} 1 \longmapsto 2 \\ 2 \longmapsto 1. \end{cases}$ Le groupe symétrique S_3 est quant à lui de cardinal 6 et ses éléments sont les applications :

$$\mathrm{Id}_{[\![1,3]\!]},\quad \begin{cases} 1 \longmapsto 2 \\ 2 \longmapsto 1 \\ 3 \longmapsto 3, \end{cases} \quad \begin{cases} 1 \longmapsto 3 \\ 2 \longmapsto 2 \\ 3 \longmapsto 1, \end{cases} \quad \begin{cases} 1 \longmapsto 1 \\ 2 \longmapsto 3 \\ 3 \longmapsto 2, \end{cases} \quad \begin{cases} 1 \longmapsto 2 \\ 2 \longmapsto 3 \\ 3 \longmapsto 1 \end{cases} \quad \text{et} \quad \begin{cases} 1 \longmapsto 3 \\ 2 \longmapsto 1 \\ 3 \longmapsto 2. \end{cases}$$

Théorème (Bijectivité, réciproque et composition) Soient $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications.

(i) Si f est bijective de E sur F, f^{-1} est bijective de F sur E et $\big(f^{-1}\big)^{-1} = f$.

(ii) Si f et g sont bijectives, $g \circ f$ l'est aussi et $\big(g \circ f\big)^{-1} = f^{-1} \circ g^{-1}$.

Attention ! Gare à l'ordre ! C'est bien $\big(g \circ f\big)^{-1} = f^{-1} \circ g^{-1}$ et non pas $\big(g \circ f\big)^{-1} = g^{-1} \circ f^{-1}$. Si vous cachez un trésor dans un coffre (f), puis ce coffre sous terre (g), et si ensuite vous voulez récupérer votre trésor (défaire $g \circ f$), vous devez d'abord déterrer le coffre (g^{-1}), puis l'ouvrir (f^{-1}) — i.e. appliquer la composée $f^{-1} \circ g^{-1}$.

Démonstration

(i) Les égalités $f^{-1} \circ f = \mathrm{Id}_E$ et $f \circ f^{-1} = \mathrm{Id}_F$ — qui expriment la bijectivité de f — expriment pour la même raison la bijectivité de f^{-1} et cela montre bien que $\big(f^{-1}\big)^{-1} = f$.

(ii) Pour commencer : $\big(f^{-1} \circ g^{-1}\big) \circ \big(g \circ f\big) = f^{-1} \circ \big(g^{-1} \circ g\big) \circ f = f^{-1} \circ \mathrm{Id}_F \circ f = f^{-1} \circ f = \mathrm{Id}_E$ et de même : $\big(g \circ f\big) \circ \big(f^{-1} \circ g^{-1}\big) = \mathrm{Id}_G$, donc $g \circ f$ est bijective de réciproque $f^{-1} \circ g^{-1}$. ∎

L'exemple qui suit met en avant une idée simple mais importante — les bijections possèdent une réciproque, donc on peut facilement les « défaire ».

Exemple Soient E un ensemble et $f : E \longrightarrow E$ et $g : E \longrightarrow E$ deux applications. Si $f \circ g \circ f$ est bijective de E sur E, alors f et g le sont aussi.

Démonstration Pour commencer, f est injective car $(f \circ g) \circ f$ l'est et surjective car $f \circ (g \circ f)$ l'est, donc f est bijective. Elle possède ainsi une réciproque f^{-1} que nous pouvons exploiter pour « défaire » f. Or tout simplement $g = f^{-1} \circ (f \circ g \circ f) \circ f^{-1}$ avec f^{-1} et $f \circ g \circ f$ bijectives, donc g l'est aussi par composition.

Et à présent, comment montre-t-on concrètement qu'une application est bijective ? Suivez le guide !

Priorité	Ce qu'on fait	Ce qu'on obtient
1	Si on connaît spontanément une expression explicite de f^{-1}, on appelle g la fonction en question et on vérifie simplement que $g \circ f = \mathrm{Id}_E$ et $f \circ g = \mathrm{Id}_F$.	Bijectivité + Réciproque
2	Si on ne connaît pas spontanément f^{-1}, on peut essayer d'en trouver une expression explicite via l'équivalence : $y = f(x) \iff x = f^{-1}(y)$.	Bijectivité + Réciproque
3	Si on ne se sent pas capable de trouver une expression explicite de f^{-1}, on montre en deux temps que f est à la fois injective et surjective.	Bijectivité

Exemple La fonction $x \overset{f}{\longmapsto} \dfrac{x-1}{x+1}$ est bijective de $\mathbb{R}\setminus\{-1\}$ sur $\mathbb{R}\setminus\{1\}$.

Démonstration Ici, à défaut d'avoir directement une expression de réciproque en tête, on sent qu'on est capable d'en dénicher une par le calcul. Pour tous $x \in \mathbb{R}\setminus\{-1\}$ et $y \in \mathbb{R}$:

$$y = f(x) \iff y = \frac{x-1}{x+1} \iff xy + y = x - 1 \iff x(1-y) = 1 + y.$$

On a ensuite envie de diviser par $1-y$, mais si $y = 1$, l'équation s'écrit $0 = 2$ et n'a pas de solution. En d'autres termes, 1 n'a pas d'antécédent par f, donc l'image de f est incluse dans $\mathbb{R}\setminus\{1\}$.

Pour finir, sous l'hypothèse que $y \neq 1$: $y = f(x) \iff x = \dfrac{1+y}{1-y}$. Il en découle d'abord que l'image de f est exactement $\mathbb{R}\setminus\{1\}$ car on a réussi à trouver un antécédent de tout élément de cet ensemble, mais on vient surtout de montrer que f est bijective de $\mathbb{R}\setminus\{-1\}$ sur $\mathbb{R}\setminus\{1\}$ de réciproque $y \longmapsto \dfrac{1+y}{1-y}$.

Exemple L'application $(x,y) \overset{g}{\longmapsto} (x+y, xy)$ de $\mathbb{R}^2$ dans $\mathbb{R}^2$ n'est pas injective car par exemple $g(0,1) = (1,0) = g(1,0)$, mais elle est bijective de $\big\{(x,y) \in \mathbb{R}^2 \mid x \leqslant y\big\}$ sur $\big\{(x,y) \in \mathbb{R}^2 \mid x^2 - 4y \geqslant 0\big\}$.

Démonstration Ici aussi, à défaut d'avoir une expression de réciproque en tête, on va réussir à en trouver une. La question est au fond la suivante — si je connais $a = x+y$ et $b = xy$, suis-je capable de récupérer x et y ? La réponse est essentiellement oui car nous savons résoudre les équations du second degré. Observons en effet que $(X-x)(X-y) = X^2 - (x+y)X + xy$. Par identification polynomiale, il en découle que pour tous $(x,y),(a,b) \in \mathbb{R}^2$:

$$(a,b) = g(x,y) \iff \begin{cases} x+y = a \\ xy = b \end{cases} \iff (X-x)(X-y) = X^2 - aX + b$$
$$\iff x \text{ et } y \text{ sont les deux racines de } X^2 - aX + b \text{ (éventuellement égales).}$$

Le couple (a,b) possède ainsi un antécédent (x,y) par g si et seulement si le polynôme $X^2 - aX + b$ possède deux racines RÉELLES, i.e. si et seulement si son discriminant $a^2 - 4b$ est positif (ou nul). On en déduit l'image de g : $g\big(\mathbb{R}^2\big) = \big\{(a,b) \in \mathbb{R}^2 \mid a^2 - 4b \geqslant 0\big\}$. Sous l'hypothèse que $a^2 - 4b \geqslant 0$ et $a^2 - 4b \geqslant 0$:

$$(a,b) = g(x,y) \iff (x,y) = \left(\frac{a+\sqrt{a^2-b}}{2}, \frac{a-\sqrt{a^2-b}}{2}\right) \text{ ou } (x,y) = \left(\frac{a-\sqrt{a^2-b}}{2}, \frac{a+\sqrt{a^2-b}}{2}\right).$$

Le couple (a,b) possède ainsi exactement un antécédent par g si $a^2 - 4b = 0$ et deux si $a^2 - 4b > 0$.

Posons pour finir $E = \big\{(x,y) \in \mathbb{R}^2 \mid x \leqslant y\big\}$. Sous l'hypothèse additionnelle que $(x,y) \in E$:

$$(a,b) = g(x,y) \iff (x,y) = \left(\frac{a-\sqrt{a^2-b}}{2}, \frac{a+\sqrt{a^2-b}}{2}\right),$$

donc $g_{|E}$ est bijective de E sur $\big\{(a,b) \in \mathbb{R}^2 \mid a^2 - 4b \geqslant 0\big\}$ de réciproque $(a,b) \longmapsto \left(\dfrac{a-\sqrt{a^2-b}}{2}, \dfrac{a+\sqrt{a^2-b}}{2}\right)$.

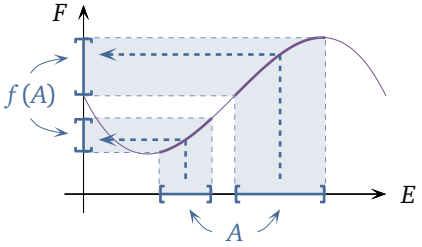
2.5 Images directes et réciproques

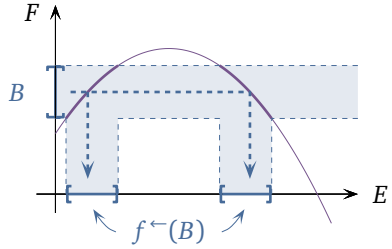
Définition (Image directe/réciproque d'une partie par une application) Soit $f : E \longrightarrow F$ une application.

- **Image directe :** Pour toute partie A de E, on appelle *image (directe) de A par f* l'ensemble des images par f des éléments de A : $f(A) = \big\{y \in F \mid \exists a \in A,\ y = f(a)\big\} = \big\{f(a) \mid a \in A\big\}$.
- **Image réciproque :** Pour toute partie B de F, on appelle *image réciproque de B par f* l'ensemble des éléments de E dont l'image par f appartient à B : $f^{\leftarrow}(B) = \big\{x \in E \mid f(x) \in B\big\}$. La notation $f^{\leftarrow}(B)$ est provisoire.

  L'équivalence suivante est vitale. Pour tout $x \in E$: $x \in f^{\leftarrow}(B) \iff f(x) \in B$.

Pour représenter $f(A)$, on projette sur l'axe des ordonnées la portion du graphe de f qui se situe au-dessus de A. Pour représenter $f^{\leftarrow}(B)$, on projette sur l'axe des abscisses la portion du graphe de f située dans le tube horizontal défini par B.





Pour une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, chercher l'image réciproque d'un singleton $\{y\}$ par f revient à résoudre l'équation $y = f(x)$ d'inconnue x, alors que pour un intervalle $[a,b]$, cela revient à résoudre l'inéquation $a \leqslant f(x) \leqslant b$.

Exemple $\exp(\mathbb{R}_+) = [1,+\infty[$, $\exp(\mathbb{R}_-^*) =]0,1[$, $\exp^{\leftarrow}(\mathbb{R}_-) = \{x \in \mathbb{R} \mid e^x \in \mathbb{R}_-\} = \varnothing$

$\exp^{\leftarrow}(\{3\}) = \{x \in \mathbb{R} \mid e^x = 3\} = \{\ln 3\}$, $\exp^{\leftarrow}([1,2[) = \{x \in \mathbb{R} \mid e^x \in [0,1[\} = [0,\ln 2[$.

Exemple En notant f la fonction carré $x \longmapsto x^2$ sur $\mathbb{R}$: $f([1,2]) = [1,4]$, $f^{\leftarrow}([4,+\infty[) =]-\infty,-2] \cup [2,+\infty[$.

Exemple $\sin^{\leftarrow}(\{-1\}) = \{x \in \mathbb{R} \mid \sin x = -1\} = -\frac{\pi}{2} + 2\pi\mathbb{Z}$, $\sin^{\leftarrow}([0,2]) = [0,\pi] + 2\pi\mathbb{Z}$.

Exemple Soit $n \in \mathbb{N}^*$. En notant g la fonction qui associe à tout entier relatif le reste de sa division euclidienne par n : $g(\mathbb{Z}) = [\![0,n-1]\!]$ et pour tout $r \in [\![0,n-1]\!]$: $g^{\leftarrow}(\{r\}) = r + n\mathbb{Z}$.

Théorème (Bijectivité et image réciproque) Soit f une bijection de E sur F. Pour toute partie B de F :

$$f^{\leftarrow}(B) = f^{-1}(B),$$

où $f^{\leftarrow}(B)$ est l'image RÉCIPROQUE de B par f et $f^{-1}(B)$ l'image DIRECTE de B par f^{-1}.

Démonstration Pour tout $x \in E$: $x \in f^{-1}(B) \iff \exists\, b \in B,\ x = f^{-1}(b) \iff \exists\, b \in B,\ f(x) = b \iff f(x) \in B \iff x \in f^{\leftarrow}(B)$. ∎

Attention ! Grâce à ce théorème, nous noterons désormais TOUJOURS $f^{-1}(B)$ l'image réciproque de B par f et plus jamais $f^{\leftarrow}(B)$. Nous avons introduit cette notation provisoire dans le seul but de comprendre que la notation définitive $f^{-1}(B)$ contient au départ une bonne dose d'ambiguïté.

- Dans le cas où f est bijective, le théorème affirme que l'image réciproque de B par f est exactement l'image directe de B par f^{-1}. La confusion des notations $f^{-1}(B)$ et $f^{\leftarrow}(B)$ n'est donc pas gênante.
- Dans le cas où f n'est pas bijective, IL N'Y A PAS de réciproque f^{-1}, donc pas d'image directe de B par f^{-1}, donc pas de confusion possible. La notation $f^{-1}(B)$ ne pose aucun problème là non plus.

En guise de conclusion : La notation $f^{-1}(B)$ ne requiert pas la bijectivité de f !

Exemple L'application $A \overset{\varphi}{\longmapsto} \mathbb{1}_A$ est bijective de $\mathscr{P}(E)$ sur $\{0,1\}^E$.

Démonstration Pour toute partie A de E, il est facile de récupérer A quand on connaît $\mathbb{1}_A$:

$$A = \{x \in E \mid \mathbb{1}_A(x) = 1\} = \mathbb{1}_A^{-1}(\{1\}).$$

Montrons que l'application $f \overset{\psi}{\longmapsto} f^{-1}(\{1\})$ est la réciproque de φ sur $\{0,1\}^E$.

— Montrons que $\psi \circ \varphi = \mathrm{Id}_{\mathscr{P}(A)}$. Or pour tout $A \in \mathscr{P}(A)$: $\psi \circ \varphi(A) = \psi(\mathbb{1}_A) = \mathbb{1}_A^{-1}(\{1\}) = A$.

— Montrons que $\varphi \circ \psi = \mathrm{Id}_{\{0,1\}^E}$. Soit $f \in \{0,1\}^E$. Alors $\varphi \circ \psi(f) = \varphi\left(f^{-1}(\{1\})\right) = \mathbb{1}_{f^{-1}(\{1\})}$, donc pour tout $x \in E$: $\varphi \circ \psi(f)(x) = \begin{cases} 1 & \text{si } x \in f^{-1}(\{1\}), \text{ i.e. si } f(x) = 1 \\ 0 & \text{sinon, i.e. si } f(x) = 0 \end{cases} = f(x)$, donc $\varphi \circ \psi(f) = f$.

Théorème (Réunions/intersection d'images directes/réciproques) Soient $f : E \longrightarrow F$ une application, $\{A_i \mid i \in I\}$ un ensemble de parties de E et $\{B_j \mid j \in J\}$ un ensemble de parties de F.

(i) **Images réciproques :** Avec les images réciproques, tout va bien !

$$f^{-1}\left(\bigcup_{j\in J} B_j\right) = \bigcup_{j\in J} f^{-1}(B_j) \quad \text{et} \quad f^{-1}\left(\bigcap_{j\in J} B_j\right) = \bigcap_{j\in J} f^{-1}(B_j).$$

(ii) **Images directes :** $f\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f(A_i)$, mais l'égalité $f\left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I} f(A_i)$ n'est pas vraie en général. Elle l'est si f est injective.

Démonstration

(i) Pour tout $x \in E$: $\quad x \in f^{-1}\left(\bigcup_{j\in J} B_j\right) \iff f(x) \in \bigcup_{j\in J} B_j \iff \exists j \in J,\quad f(x) \in B_j$
$\iff \exists j \in J,\quad x \in f^{-1}(B_j) \iff x \in \bigcup_{j\in J} f^{-1}(B_j).$

Pour la deuxième égalité, simplement remplacer $\bigcup$ par $\bigcap$ et $\exists$ par $\forall$.

(ii) Montrons que $f\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f(A_i)$.

— Soit $y \in f\left(\bigcup_{i\in I} A_i\right)$. Alors $y = f(x)$ pour un certain $x \in \bigcup_{i\in I} A_i$, donc $x \in A_{i_0}$ pour un certain $i_0 \in I$, donc $y = f(x) \in f(A_{i_0}) \subset \bigcup_{i\in I} A_i$.

— Inversement, soit $y \in \bigcup_{i\in I} f(A_i)$. Alors $y \in f(A_{i_0})$ pour un certain $i_0 \in I$, donc $y = f(x)$ pour un certain $x \in A_{i_0}$. En particulier, $x \in \bigcup_{i\in I} A_i$ donc $y = f(x) \in f\left(\bigcup_{i\in I} A_i\right)$.

Montrons que si f est injective, alors $f\left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I} f(A_i)$.

— L'inclusion $\subset$ est en fait vraie sans injectivité. Soit $y \in f\left(\bigcap_{i\in I} A_i\right)$. Alors $y = f(x)$ pour un certain $x \in \bigcap_{i\in I} A_i$, donc $x \in A_i$ pour tout $i \in I$, puis $y = f(x) \in f(A_i)$. Comme voulu, $y \in \bigcap_{i\in I} f(A_i)$.

— Inversement, sous l'hypothèse que f est injective, soit $y \in \bigcap_{i\in I} f(A_i)$. Alors $y \in f(A_i)$ pour tout $i \in I$, donc $y = f(x_i)$ pour un certain $x_i \in A_i$. Par conséquent, $f(x_{i_1}) = y = f(x_{i_2})$ pour tous $i_1, i_2 \in I$, donc $x_{i_1} = x_{i_2}$ par injectivité de f. Les x_i sont ainsi tous égaux et nous pouvons les noter x. Finalement, $x \in \bigcap_{i\in I} A_i$ donc $y = f(x) \in f\left(\bigcap_{i\in I} A_i\right)$.

Théorème **(Partition de l'ensemble de départ par une application)** Soit $f : E \longrightarrow F$ une application. La relation binaire définie par la relation $f(x) = f(x')$ pour tous $x, x' \in E$ est une relation d'équivalence sur E et ses classes d'équivalence sont les ensembles $f^{-1}(\{y\})$, y décrivant $f(E)$. Par conséquent : $\quad E = \bigsqcup_{y\in f(E)} f^{-1}(\{y\})$.

Pour toute valeur y de f, $f^{-1}(\{y\})$ est l'ensemble des éléments de E que f envoie sur y. Le théorème raconte juste que les éléments de E peuvent être groupés par paquets en fonction de la valeur que f leur attribue.

Démonstration Notons $\equiv$ la relation binaire de l'énoncé.

- **Réflexivité :** Pour tout $x \in E$, $f(x) = f(x)$ donc $x \equiv x$.
- **Transitivité :** Pour tous $x, x', x'' \in E$, si $x \equiv x'$ et $x' \equiv x''$, alors $f(x) = f(x') = f(x'')$, donc $f(x) = f(x'')$, i.e. $x \equiv x''$.
- **Symétrie :** Pour tous $x, x' \in E$, si $x \equiv x'$, alors $f(x) = f(x')$, donc $f(x') = f(x)$, i.e. $x' \equiv x$.
- **Classes d'équivalence :** Pour tout $x \in E$, la classe d'équivalence de x pour $\equiv$ est l'ensemble
$$\{x' \in E \mid \quad f(x') = f(x)\}, \qquad \text{i.e. l'ensemble } f^{-1}(\{y\}) \text{ si on pose } y = f(x).$$

3 L'équipotence et ses paradoxes

Définition **(Équipotence)** On dit que F est *équipotent à* E s'il existe une bijection de E sur F.

Tiré du latin, « équipotent » veut dire « de même puissance », mais en quel sens ? Avec une bijection de E sur F, on peut faire correspondre parfaitement les éléments de E aux éléments de F, associer à tout élément de E un et un seul élément de F et vice versa. Dire que F est équipotent à E revient ainsi à dire que « F a exactement le même nombre d'éléments que E ». Je mets des guillemets car la notion de *nombre d'éléments* ou *cardinal* n'a pas encore été définie proprement, nous nous

sommes contentés d'une intuition jusqu'ici. En réalité, la notion d'équipotence précède la notion de cardinal du point de vue de la théorie et nous donnerons une définition rigoureuse du cardinal au prochain chapitre.

Dans le même registre, l'existence d'une injection de E dans F fait de $f(E)$ une copie de E dans F, donc « F a plus d'éléments que E ou éventuellement autant ». Enfin, l'existence d'une surjection de E sur F permet à E de couvrir F à travers f, donc « F a moins d'éléments que E ou éventuellement autant ».

Théorème **(Propriétés de la relation d'équipotence)** La relation d'équipotence est une relation d'équivalence sur la classe des ensembles.

Démonstration Pour tout ensemble E, Id_E est bijective de E sur E — réflexivité. Ensuite, la composée de deux bijections est une bijection — transitivité. Enfin, la réciproque d'une bijection est une bijection — symétrie. ∎

Je ne démontrerai pas l'important théorème que voici, mais sa preuve est tout à fait accessible.

Théorème **(Théorème de Cantor-Bernstein)** S'il existe une injection de E dans F et une injection de F dans E, E et F sont équipotents.

Bref — toujours sans grande rigueur — si E a moins d'éléments que F et F moins d'éléments que E, E et F en ont autant !

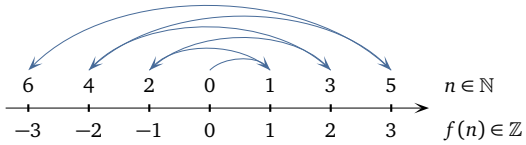
Exemple L'application $n \longmapsto n+1$ est bijective de $\mathbb{N}$ sur $\mathbb{N}^*$, donc $\mathbb{N}$ et $\mathbb{N}^*$ sont équipotents alors que l'un est inclus strictement dans l'autre !

Exemple $\mathbb{R}$ et $]0,1[$ sont équipotents. Il y a donc autant d'éléments dans $\mathbb{R}$ que dans $]0,1[$.

Démonstration On peut montrer que la fonction $x \longmapsto \dfrac{x}{|x|+1}$ est bijective de $]0,1[$ sur $\mathbb{R}$.

Exemple $\mathbb{N}$ et $\mathbb{Z}$ sont équipotents.

Démonstration L'application f représentée ci-contre est bijective de $\mathbb{N}$ sur $\mathbb{Z}$. Trouvez-en une expression explicite !



Exemple $\mathbb{N}$ et $\mathbb{N}^2$ sont équipotents.

Démonstration Montrons que l'application $(p,q) \overset{g}{\longmapsto} 2^p(2q+1)$ est bijective de $\mathbb{N}^2$ sur $\mathbb{N}^*$. Il en découlera que $\mathbb{N}^2$ et $\mathbb{N}^*$ sont équipotents, mais $\mathbb{N}$ et $\mathbb{N}^*$ le sont, donc $\mathbb{N}^2$ et $\mathbb{N}$ seront équipotents par transitivité.

- Pour l'injectivité, soient $(p,q),(p',q') \in \mathbb{N}^2$ deux couples pour lesquels $g(p,q) = g(p',q')$. Quitte à les permuter, on peut supposer $p \leqslant p'$ sans perte de généralité. Ainsi $2^{p'-p}(2q'+1) = 2q+1$, égalité dans laquelle $2^{p'-p}$, $2q+1$ et $2q'+1$ sont des entiers. Comme $2q+1$ est impair, on en déduit que $p = p'$, puis que $2q+1 = 2q'+1$, et enfin $(p,q) = (p',q')$.
- Pour la surjectivité, soit $y \in \mathbb{N}^*$. Dans la factorisation première de y, mettons de côté le nombre premier 2 et multiplions entre eux les autres diviseurs premiers. Cela permet d'écrire y sous la forme $y = 2^p(2q+1)$ pour certains $p,q \in \mathbb{N}$, et ainsi $y = g(p,q)$.

Exemple $\mathbb{N}$ et $\mathbb{Q}$ sont équipotents. Autrement dit — et c'est assez fou au premier abord — on peut numéroter les rationnels. Il y a un rationnel n°0, un rationnel n°1, un rationnel n°2, etc.

Démonstration L'application $\mathrm{Id}_{\mathbb{N}}$ est injective de $\mathbb{N}$ dans $\mathbb{Q}$, donc d'après le théorème de Cantor-Bernstein, il nous suffit d'exhiber une injection de $\mathbb{Q}$ dans $\mathbb{N}$ pour montrer que $\mathbb{N}$ et $\mathbb{Q}$ sont équipotents.

— Tout rationnel r possède une et une seule forme irréductible $\dfrac{p}{q}$ avec $(p,q) \in \mathbb{Z} \times \mathbb{N}^*$. L'application u qui associe (p,q) à r est définie de $\mathbb{Q}$ dans $\mathbb{Z} \times \mathbb{N}^*$ et injective.

— Étant donnée une bijection f de $\mathbb{Z}$ sur $\mathbb{N}$, l'application $(m,n) \overset{v}{\longmapsto} \big(f(m), n-1\big)$ est bijective de $\mathbb{Z} \times \mathbb{N}^*$ sur $\mathbb{N}^2$.

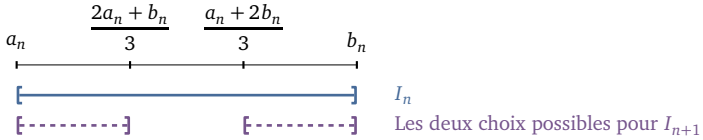
Par composition, $v \circ u$ est injective de $\mathbb{Q}$ dans $\mathbb{N}^2$. On obtient l'injection de $\mathbb{Q}$ dans $\mathbb{N}$ souhaitée en composant finalement $v \circ u$ avec une bijection de $\mathbb{N}^2$ sur $\mathbb{N}$.

Exemple $\mathbb{R}$ et $\mathbb{Q}$ ne sont pas équipotents. Il y a donc infiniment plus d'éléments dans $\mathbb{R}$ que dans $\mathbb{Q}$, donc a fortiori infiniment plus d'irrationnels que de rationnels.

Démonstration Parce que $\mathbb{N}$ et $\mathbb{Q}$ sont équipotents, montrer que $\mathbb{R}$ et $\mathbb{Q}$ ne le sont pas revient à montrer que $\mathbb{R}$ et $\mathbb{N}$ ne le sont pas non plus. Et pour montrer qu'il n'existe pas de bijection de $\mathbb{N}$ sur $\mathbb{R}$, nous allons en fait prouver qu'aucune application de $\mathbb{N}$ dans $\mathbb{R}$ ne peut être surjective, ce sera suffisant.

Soit $\varphi : \mathbb{N} \longrightarrow \mathbb{R}$ une application quelconque. Montrons que φ n'est pas surjective.

- L'un des intervalles $\left[0, \frac{1}{3}\right]$ et $\left[\frac{2}{3}, 1\right]$ ne contient pas $\varphi(0)$, disons I_0 — si les deux intervalles conviennent, on note I_0 celui de gauche par exemple. Par construction, $\varphi(0) \notin I_0$. Notons a_0 et b_0 les bornes de I_0, de sorte que $I_0 = [a_0, b_0]$. L'intervalle I_0 a pour longueur $\frac{1}{3}$.
- Ensuite, on répète. Pour tout $n \in \mathbb{N}$, une fois les intervalles $I_0, I_1, \dots, I_n$ construits, on construit l'intervalle I_{n+1} de la façon suivante. L'un des intervalles $\left[a_n, \frac{2a_n + b_n}{3}\right]$ et $\left[\frac{a_n + 2b_n}{3}, b_n\right]$ ne contient pas $\varphi(n+1)$, disons I_{n+1} — si les deux intervalles conviennent, on note I_{n+1} celui de gauche par exemple. Par construction, $\varphi(n+1) \notin I_{n+1}$. Notons a_{n+1} et b_{n+1} les bornes de I_{n+1}, de sorte que $I_{n+1} = [a_{n+1}, b_{n+1}]$. L'intervalle I_{n+1} a pour longueur $\frac{1}{3^{n+1}}$ car la longueur est divisée par 3 à chaque étape.



- Les intervalles ainsi construits sont emboîtés les uns dans les autres avec $[a_{n+1}, b_{n+1}] = I_{n+1} \subset I_n = [a_n, b_n]$ pour tout $n \in \mathbb{N}$, donc $a_0 \leqslant a_1 \leqslant a_2 \leqslant \dots \leqslant a_n \leqslant a_{n+1} \leqslant b_{n+1} \leqslant b_n \leqslant \dots \leqslant b_2 \leqslant b_1 \leqslant b_0$. Croissante et majorée par b_0, la suite $(a_n)_{n\in\mathbb{N}}$ converge d'après le théorème de la limite monotone, disons vers a_∞, et la suite $(b_n)_{n\in\mathbb{N}}$ converge de même vers un certain b_∞. Cela dit, $b_n - a_n = \frac{1}{3^{n+1}}$ pour tout $n \in \mathbb{N}$, donc par passage à la limite : $b_\infty - a_\infty = 0$, i.e. $a_\infty = b_\infty$. Les suites $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ convergent ainsi vers le même réel ℓ, et plus précisément $\ell \in [a_n, b_n] = I_n$ pour tout $n \in \mathbb{N}$ en raison des monotonies évoquées. Or n'oublions pas que pour tout $n \in \mathbb{N}$, $\varphi(n) \notin I_n$ par construction de I_n, donc $\varphi(n) \neq \ell$. Conclusion : φ ne prend pas la valeur ℓ, donc n'est pas surjective !

Exemple On peut montrer que $\mathbb{R}$, $\mathbb{C}/\mathbb{R}^2$ et $\mathbb{R}^3$ sont équipotents. Il y a donc autant de points sur une droite ou sur un plan que dans notre espace à trois dimensions.

Nous terminerons ce chapitre en beauté par un petit résultat tout bête, mais d'une portée épistémologique et historique considérable.

Théorème **(Théorème de Cantor)** Il n'existe pas de surjection de E sur $\mathscr{P}(E)$.

Démonstration Soit $\varphi : E \longrightarrow \mathscr{P}(E)$ une application. On pose $A = \left\{x \in E \mid x \notin \varphi(x)\right\}$. Comme A est une partie de E, on peut se demander si A possède ou non un antécédent par φ. Pour tout $x \in E$:

— si $x \in A$, alors $x \notin \varphi(x)$, donc $\varphi(x) \neq A$,

— si $x \notin A$, alors $x \in \varphi(x)$, donc $\varphi(x) \neq A$.

Ainsi, $A \neq \varphi(x)$ dans les deux cas, et ce pour tout $x \in E$, donc A n'a pas d'antécédent par f. Conclusion : f n'est pas surjective de E sur $\mathscr{P}(E)$.

Dans la mesure où l'application $x \longmapsto \left\{x\right\}$ est injective de E dans $\mathscr{P}(E)$, le théorème de Cantor montre au fond que E est toujours STRICTEMENT plus petit que $\mathscr{P}(E)$ en termes d'équipotence. En particulier, $\mathbb{N}$ est strictement plus petit que $\mathscr{P}(\mathbb{N})$, qui est lui-même strictement plus petit que $\mathscr{P}\big(\mathscr{P}(\mathbb{N})\big)$, lui-même strictement plus petit que $\mathscr{P}\Big(\mathscr{P}\big(\mathscr{P}(\mathbb{N})\big)\Big)$, etc. Il n'est par ailleurs pas trop dur de montrer que $\mathscr{P}(\mathbb{N})$ et $\mathbb{R}$ sont équipotents.

À la fin du 19ème siècle, le mathématicien allemand Georg Cantor (1845-1918) se demande s'il existe ou non entre $\mathbb{N}$ et $\mathscr{P}(\mathbb{N})$, c'est-à-dire entre $\mathbb{N}$ et $\mathbb{R}$, un infini de taille intermédiaire. Il n'obtient cependant aucun résultat ni dans un sens ni dans l'autre. L'énoncé selon lequel un tel infini intermédiaire n'existe pas s'appelle depuis l'*hypothèse du continu*.

En 1938, le mathématicien autrichien Kurt Gödel (1906-1978) montre que l'hypothèse du continu ne réfute pas le cadre traditionnel dit *ZFC* des mathématiques. Ce résultat est compliqué à comprendre. Gödel n'a pas montré que l'hypothèse du continu est vraie, mais que si on l'ajoute aux axiomes de la théorie ZFC comme un axiome supplémentaire, la théorie obtenue n'est ni plus ni moins contradictoire que la théorie ZFC.

En 1963, le mathématicien américain Paul Cohen (1934-2007) montre que l'hypothèse du continu n'est pas démontrable dans la théorie usuelle ZFC. L'hypothèse du continu est dès lors un énoncé *indécidable*, impossible à prouver et impossible à réfuter.