Mathématiques et méthodes numériques

Sommaire

Limites et fonctions continues

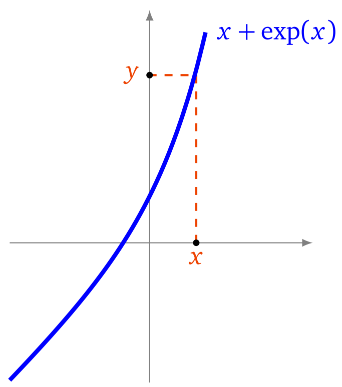
Chapitre 1

Objectif

Les équations en une variable x qu'on sait résoudre explicitement, c'est-à-dire en donnant une formule pour la solution, sont très particulières : par exemple les équations du premier degré $ax+b=0$, celles du second degré $ax^2+bx+c=0$.
Mais pour la plupart des équations, il n'est pas possible de donner une formule pour la ou les solutions. En fait il n'est même pas évident de déterminer seulement le nombre de solutions, ni même s'il en existe. Considérons par exemple l'équation extrêmement simple :

$$x+\exp x=0$$

Il n'y a pas de formule explicite (utilisant des sommes, des produits, des fonctions usuelles) pour trouver la solution x.
Dans ce chapitre nous allons voir que grâce à l'étude de la fonction $f(x)=x+\exp x$, il est possible d'obtenir beaucoup d'informations sur l'ensemble des solutions de l'équation $x+\exp x=0$, et même de l'équation plus générale $x+\exp x=y$ (où $y\in\mathbb{R}$ est fixé).



Nous serons capables de prouver que pour chaque $y\in\mathbb{R}$ l'équation « $x+\exp x=y$ » admet une solution x, que cette solution est unique, et nous saurons dire comment varie x en fonction de y. Le point clé de cette résolution est l'étude de la fonction f et en particulier de sa continuité. Même s'il n'est pas possible de trouver l'expression exacte de la solution x en fonction de y, nous allons mettre en place les outils théoriques qui permettent d'en trouver une solution approchée.

1. Notions de fonction

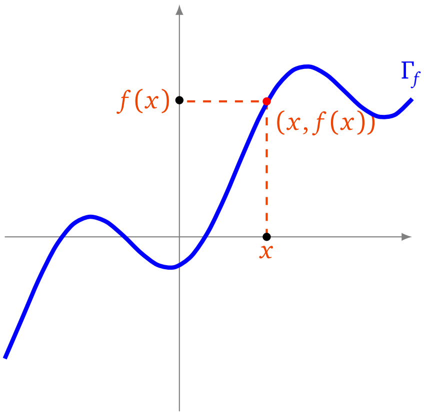
1.1. Définitions

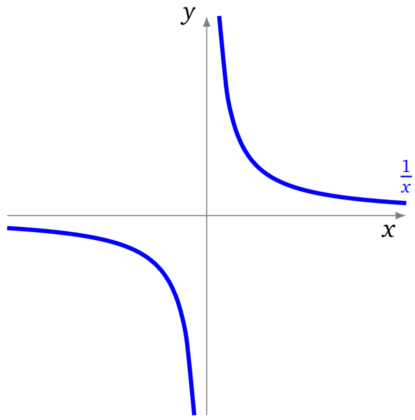
Définition 1.
Une *fonction* d'une variable réelle à valeurs réelles est une application $f : U \to \mathbb{R}$, où U est une partie de $\mathbb{R}$. En général, U est un intervalle ou une réunion d'intervalles. On appelle U le *domaine de définition* de la fonction f.

Exemple 1.
La fonction inverse :

$$\begin{array}{rcl} f :]-\infty, 0[\cup]0, +\infty[& \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{1}{x}. \end{array}$$

Le *graphe* d'une fonction $f : U \to \mathbb{R}$ est la partie Γ_f de $\mathbb{R}^2$ définie par $\Gamma_f = \{(x, f(x)) \mid x \in U\}$.
Le graphe d'une fonction (à gauche), l'exemple du graphe de $x \mapsto \frac{1}{x}$ (à droite).



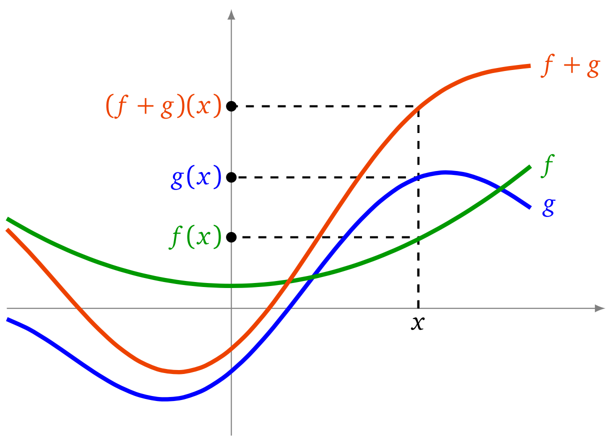


1.2. Opérations sur les fonctions

Soient $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}$ deux fonctions définies sur une même partie U de $\mathbb{R}$. On peut alors définir les fonctions suivantes :

- la *somme* de f et g est la fonction $f + g : U \to \mathbb{R}$ définie par $(f + g)(x) = f(x) + g(x)$ pour tout $x \in U$;
- le *produit* de f et g est la fonction $f \times g : U \to \mathbb{R}$ définie par $(f \times g)(x) = f(x) \times g(x)$ pour tout $x \in U$;
- la *multiplication par un scalaire* $\lambda \in \mathbb{R}$ de f est la fonction $\lambda \cdot f : U \to \mathbb{R}$ définie par $(\lambda \cdot f)(x) = \lambda \cdot f(x)$ pour tout $x \in U$.

Comment tracer le graphe d'une somme de fonction ?

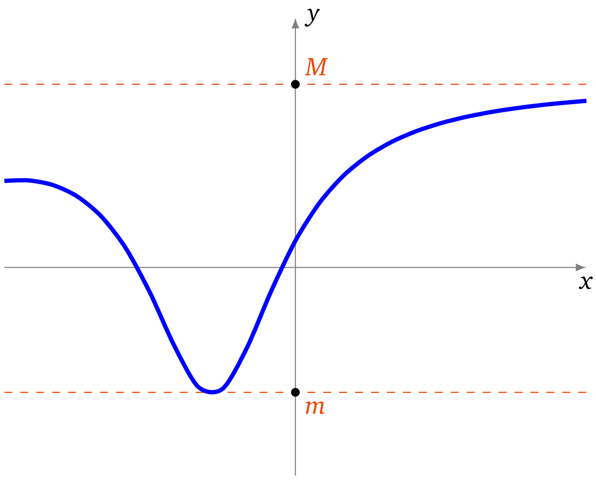


1.3. Fonctions majorées, minorées, bornées

Définition 2.
Soient $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}$ deux fonctions. Alors :
- $f \geqslant g$ si $\forall x \in U \quad f(x) \geqslant g(x)$;
- $f \geqslant 0$ si $\forall x \in U \quad f(x) \geqslant 0$;
- $f > 0$ si $\forall x \in U \quad f(x) > 0$;
- f est dite *constante* sur U si $\exists a \in \mathbb{R} \quad \forall x \in U \; f(x) = a$;
- f est dite *nulle* sur U si $\forall x \in U \quad f(x) = 0$.

Définition 3.
Soit $f : U \to \mathbb{R}$ une fonction. On dit que :
- f est *majorée* sur U si $\exists M \in \mathbb{R} \quad \forall x \in U \; f(x) \leqslant M$;
- f est *minorée* sur U si $\exists m \in \mathbb{R} \quad \forall x \in U \; f(x) \geqslant m$;
- f est *bornée* sur U si f est à la fois majorée et minorée sur U, c'est-à-dire si $\exists M \in \mathbb{R} \; \forall x \in U \quad |f(x)| \leqslant M$.

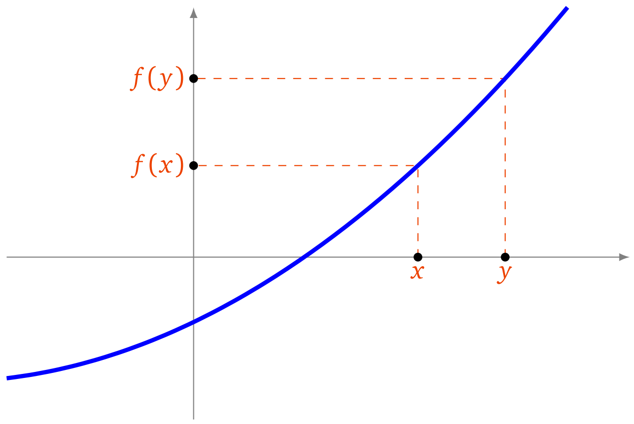
Voici le graphe d'une fonction bornée (minorée par m et majorée par M).



1.4. Fonctions croissantes, décroissantes

Définition 4.
Soit $f : U \to \mathbb{R}$ une fonction. On dit que :
- f est *croissante* sur U si $\boxed{\forall x, y \in U \quad x \leqslant y \implies f(x) \leqslant f(y)}$
- f est *strictement croissante* sur U si $\forall x, y \in U \quad x < y \implies f(x) < f(y)$
- f est *décroissante* sur U si $\forall x, y \in U \quad x \leqslant y \implies f(x) \geqslant f(y)$
- f est *strictement décroissante* sur U si $\forall x, y \in U \quad x < y \implies f(x) > f(y)$
- f est *monotone* (resp. *strictement monotone*) sur U si f est croissante ou décroissante (resp. strictement croissante ou strictement décroissante) sur U.

Un exemple de fonction croissante (et même strictement croissante) :



Exemple 2.

- La fonction racine carrée $\begin{cases} [0,+\infty[\longrightarrow \mathbb{R} \\ x \longmapsto \sqrt{x} \end{cases}$ est strictement croissante.
- Les fonctions exponentielle $\exp : \mathbb{R} \to \mathbb{R}$ et logarithme $\ln :]0,+\infty[\to \mathbb{R}$ sont strictement croissantes.
- La fonction valeur absolue $\begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto |x| \end{cases}$ n'est ni croissante, ni décroissante. Par contre, la fonction $\begin{cases} [0,+\infty[\longrightarrow \mathbb{R} \\ x \longmapsto |x| \end{cases}$ est strictement croissante.

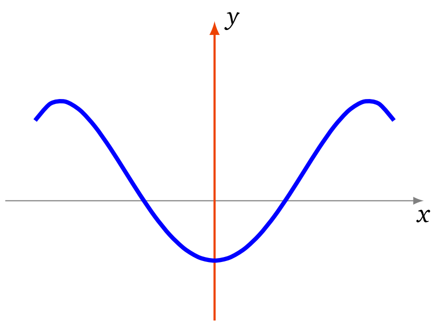
1.5. Parité et périodicité

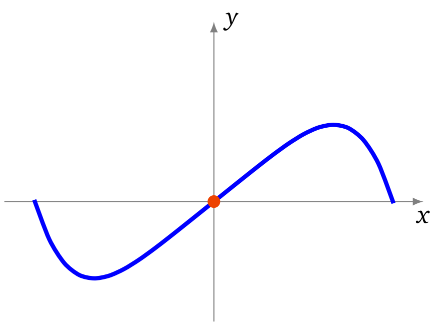
Définition 5.
Soit I un intervalle de $\mathbb{R}$ symétrique par rapport à 0 (c'est-à-dire de la forme $]-a,a[$ ou $[-a,a]$ ou $\mathbb{R}$).
Soit $f : I \to \mathbb{R}$ une fonction définie sur cet intervalle. On dit que :

- f est ***paire*** si $\forall x \in I \quad f(-x) = f(x)$,
- f est ***impaire*** si $\forall x \in I \quad f(-x) = -f(x)$.

Interprétation graphique :

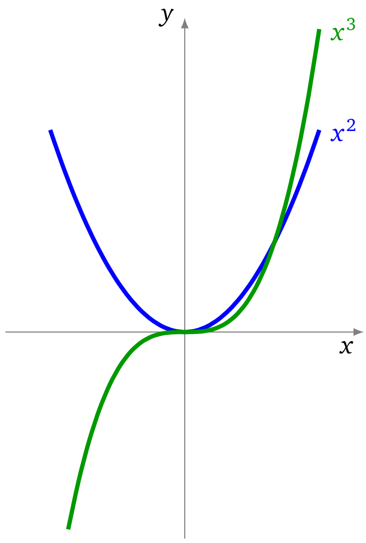
- f est paire si et seulement si son graphe est symétrique par rapport à l'axe des ordonnées (figure de gauche).
- f est impaire si et seulement si son graphe est symétrique par rapport à l'origine (figure de droite).



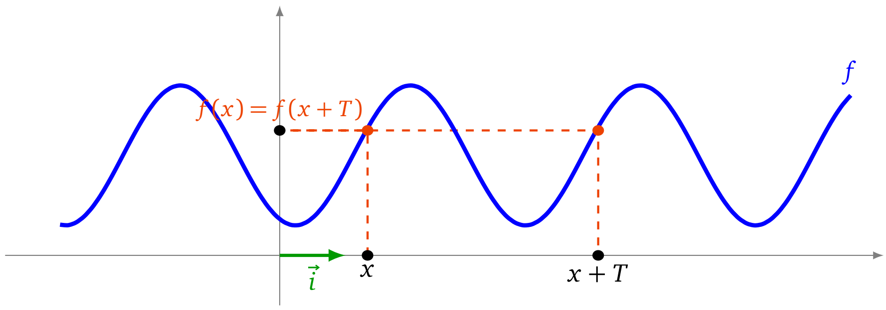


Exemple 3.

- La fonction définie sur $\mathbb{R}$ par $x \mapsto x^{2n}$ $(n \in \mathbb{N})$ est paire.
- La fonction définie sur $\mathbb{R}$ par $x \mapsto x^{2n+1}$ $(n \in \mathbb{N})$ est impaire.
- La fonction $\cos : \mathbb{R} \to \mathbb{R}$ est paire. La fonction $\sin : \mathbb{R} \to \mathbb{R}$ est impaire.

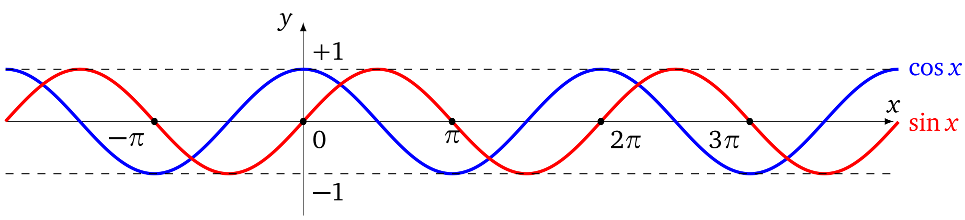


Définition 6.
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction et T un nombre réel, $T > 0$. La fonction f est dite ***périodique*** de période T si $\forall x \in \mathbb{R} \quad f(x+T) = f(x)$.



Interprétation graphique : f est périodique de période T si et seulement si son graphe est invariant par la translation de vecteur $T\vec{i}$, où $\vec{i}$ est le premier vecteur de coordonnées.

Exemple 4.
Les fonctions sinus et cosinus sont 2π-périodiques. La fonction tangente est π-périodique.



Mini-exercices.

1. Soit $U =]-\infty, 0[$ et $f : U \to \mathbb{R}$ définie par $f(x) = 1/x$. f est-elle monotone ? Et sur $U =]0, +\infty[$? Et sur $U =]-\infty, 0[\cup]0, +\infty[$?
2. Pour deux fonctions paires que peut-on dire sur la parité de la somme ? du produit ? et de la composée ? Et pour deux fonctions impaires ? Et si l'une est paire et l'autre impaire ?
3. On note $\{x\} = x - E(x)$ la partie fractionnaire de x. Tracer le graphe de la fonction $x \mapsto \{x\}$ et montrer qu'elle est périodique.
4. Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction définie par $f(x) = \frac{x}{1+x^2}$. Montrer que $|f|$ est majorée par $\frac{1}{2}$, étudier les variations de f (sans utiliser de dérivée) et tracer son graphe.

5. On considère la fonction $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \sin\big(\pi f(x)\big)$, où f est définie à la question précédente. Déduire de l'étude de f les variations, la parité, la périodicité de g et tracer son graphe.

2. Limites

2.1. Définitions

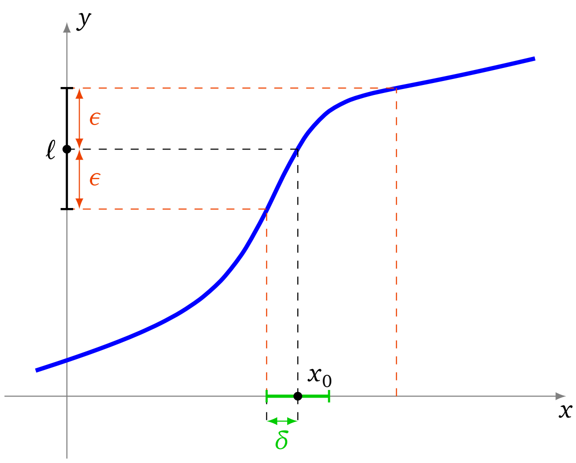
Limite en un point

Soit $f : I \to \mathbb{R}$ une fonction définie sur un intervalle I de $\mathbb{R}$. Soit $x_0 \in \mathbb{R}$ un point de I ou une extrémité de I.

Définition 7.
Soit $\ell \in \mathbb{R}$. On dit que ***f a pour limite ℓ en x_0*** si

$$\boxed{\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in I \quad |x - x_0| < \delta \implies |f(x) - \ell| < \epsilon}$$

On dit aussi que ***$f(x)$ tend vers ℓ lorsque x tend vers x_0***. On note alors $\lim_{x \to x_0} f(x) = \ell$ ou bien $\lim_{x_0} f = \ell$.

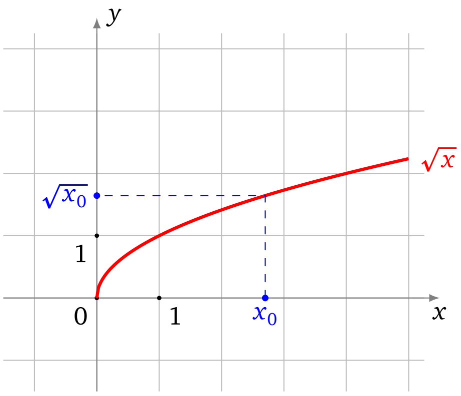


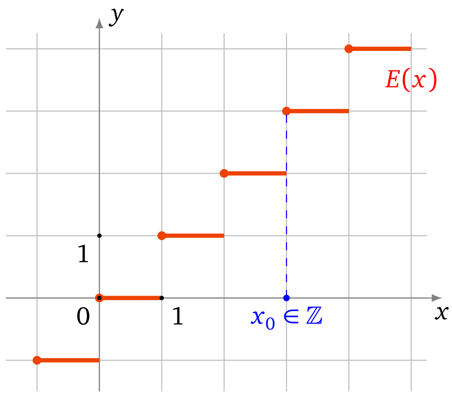
Remarque.

- L'inégalité $|x - x_0| < \delta$ équivaut à $x \in]x_0 - \delta, x_0 + \delta[$. L'inégalité $|f(x) - \ell| < \epsilon$ équivaut à $f(x) \in]\ell - \epsilon, \ell + \epsilon[$.
- On peut remplacer certaines inégalités strictes « $<$ » par des inégalités larges « $\leqslant$ » dans la définition : $\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in I \quad |x - x_0| \leqslant \delta \implies |f(x) - \ell| \leqslant \epsilon$
- Dans la définition de la limite
$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in I \quad |x - x_0| < \delta \implies |f(x) - \ell| < \epsilon$$
le quantificateur $\forall x \in I$ n'est là que pour être sûr que l'on puisse parler de $f(x)$. Il est souvent omis et l'existence de la limite s'écrit alors juste :
$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad |x - x_0| < \delta \implies |f(x) - \ell| < \epsilon.$$
- N'oubliez pas que l'ordre des quantificateurs est important, on ne peut pas échanger le $\forall \epsilon$ avec le $\exists \delta$: le δ dépend en général du ϵ. Pour marquer cette dépendance on peut écrire : $\forall \epsilon > 0 \quad \exists \delta(\epsilon) > 0 \ldots$

Exemple 5.

- $\lim_{x \to x_0} \sqrt{x} = \sqrt{x_0}$ pour tout $x_0 \geqslant 0$,
- la fonction partie entière E n'a pas de limite aux points $x_0 \in \mathbb{Z}$.





Soit f une fonction définie sur un ensemble de la forme $]a,x_0[\cup]x_0,b[$.

Définition 8.

- On dit que *f a pour limite $+\infty$ en x_0* si

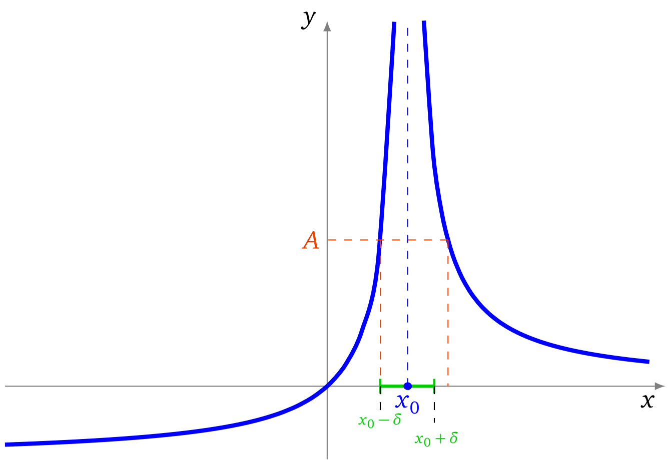
$$\forall A>0 \quad \exists \delta>0 \quad \forall x\in I \quad |x-x_0|<\delta \implies f(x)>A$$

On note alors $\lim_{x\to x_0} f(x)=+\infty$.

- On dit que *f a pour limite $-\infty$ en x_0* si

$$\forall A>0 \quad \exists \delta>0 \quad \forall x\in I \quad |x-x_0|<\delta \implies f(x)<-A$$

On note alors $\lim_{x\to x_0} f(x)=-\infty$.



Limite en l'infini

Soit $f : I\to\mathbb{R}$ une fonction définie sur un intervalle de la forme $I=]a,+\infty[$.

Définition 9.

- Soit $\ell\in\mathbb{R}$. On dit que *f a pour limite ℓ en $+\infty$* si

$$\forall \epsilon>0 \quad \exists B>0 \quad \forall x\in I \quad x>B \implies |f(x)-\ell|<\epsilon$$

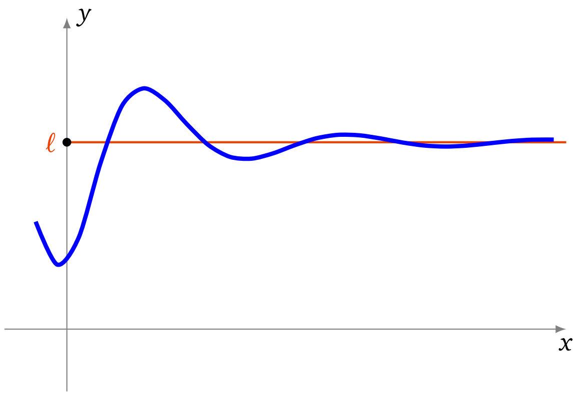
On note alors $\lim_{x\to+\infty} f(x)=\ell$ ou $\lim_{+\infty} f=\ell$.

- On dit que *f a pour limite $+\infty$ en $+\infty$* si

$$\forall A>0 \quad \exists B>0 \quad \forall x\in I \quad x>B \implies f(x)>A$$

On note alors $\lim_{x\to+\infty} f(x)=+\infty$.

On définirait de la même manière la limite en $-\infty$ pour des fonctions définies sur les intervalles du type $]-\infty,a[$.



Exemple 6.
On a les limites classiques suivantes pour tout $n \geqslant 1$:

- $\lim_{x \to +\infty} x^n = +\infty$ et $\lim_{x \to -\infty} x^n = \begin{cases} +\infty & \text{si } n \text{ est pair} \\ -\infty & \text{si } n \text{ est impair} \end{cases}$
- $\lim_{x \to +\infty} \left(\frac{1}{x^n}\right) = 0$ et $\lim_{x \to -\infty} \left(\frac{1}{x^n}\right) = 0$.

Exemple 7.
Soit $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ avec $a_n > 0$ et $Q(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$ avec $b_m > 0$.

$$\lim_{x \to +\infty} \frac{P(x)}{Q(x)} = \begin{cases} +\infty & \text{si } n > m \\ \frac{a_n}{b_m} & \text{si } n = m \\ 0 & \text{si } n < m \end{cases}$$

Limite à gauche et à droite

Soit f une fonction définie sur un ensemble de la forme $]a, x_0[\cup]x_0, b[$.

Définition 10.
- On appelle ***limite à droite*** en x_0 de f la limite de la fonction $f_{\big|]x_0,b[}$ en x_0 et on la note $\lim_{x_0^+} f$.
- On définit de même la ***limite à gauche*** en x_0 de f : la limite de la fonction $f_{\big|]a,x_0[}$ en x_0 et on la note $\lim_{x_0^-} f$.
- On note aussi $\lim_{\substack{x \to x_0 \\ x > x_0}} f(x)$ pour la limite à droite et $\lim_{\substack{x \to x_0 \\ x < x_0}} f(x)$ pour la limite à gauche.

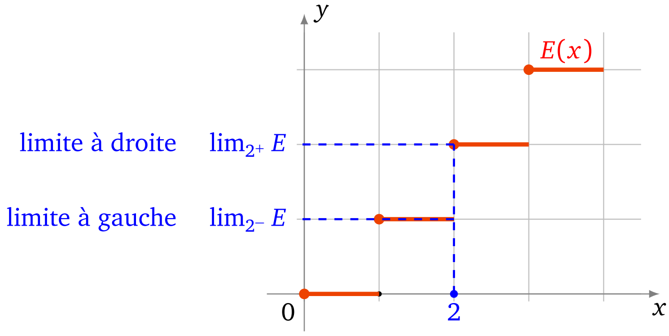
Dire que $f : I \to \mathbb{R}$ admet une limite $\ell \in \mathbb{R}$ à droite en x_0 signifie donc :

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad x_0 < x < x_0 + \delta \implies |f(x) - \ell| < \epsilon$$

Si la fonction f a une limite en x_0, alors ses limites à gauche et à droite en x_0 coïncident et valent $\lim_{x_0} f$.
Réciproquement, si f a une limite à gauche et une limite à droite en x_0 et si ces limites valent $f(x_0)$ (si f est bien définie en x_0) alors f admet une limite en x_0.

Exemple 8.
Considérons la fonction partie entière au point $x = 2$:
- comme pour tout $x \in]2,3[$ on a $E(x) = 2$, on a $\lim_{2^+} E = 2$,
- comme pour tout $x \in [1,2[$ on a $E(x) = 1$, on a $\lim_{2^-} E = 1$.

Ces deux limites étant différentes, on en déduit que E n'a pas de limite en 2.



2.2. Propriétés

Proposition 1.

Si une fonction admet une limite, alors cette limite est unique.

On ne donne pas la démonstration de cette proposition, qui est très similaire à celle de l'unicité de la limite pour les suites (un raisonnement par l'absurde).

Soient deux fonctions f et g. On suppose que x_0 est un réel, ou que $x_0 = \pm\infty$.

Proposition 2.

Si $\lim_{x_0} f = \ell \in \mathbb{R}$ *et* $\lim_{x_0} g = \ell' \in \mathbb{R}$, *alors :*

- $\lim_{x_0}(\lambda \cdot f) = \lambda \cdot \ell$ *pour tout* $\lambda \in \mathbb{R}$
- $\lim_{x_0}(f+g) = \ell + \ell'$
- $\lim_{x_0}(f \times g) = \ell \times \ell'$
- *si* $\ell \neq 0$, *alors* $\lim_{x_0} \frac{1}{f} = \frac{1}{\ell}$

De plus, si $\lim_{x_0} f = +\infty$ *(ou* $-\infty$*) alors* $\lim_{x_0} \frac{1}{f} = 0$.

Cette proposition se montre de manière similaire à la proposition analogue sur les limites de suites. Nous n'allons donc pas donner la démonstration de tous les résultats.

Démonstration. Montrons par exemple que si f tend en x_0 vers une limite ℓ non nulle, alors $\frac{1}{f}$ est bien définie dans un voisinage de x_0 et tend vers $\frac{1}{\ell}$.
Supposons $\ell > 0$, le cas $\ell < 0$ se montrerait de la même manière. Montrons tout d'abord que $\frac{1}{f}$ est bien définie et est bornée dans un voisinage de x_0 contenu dans l'intervalle I. Par hypothèse

$$\forall \epsilon' > 0 \quad \exists \delta > 0 \quad \forall x \in I \quad x_0 - \delta < x < x_0 + \delta \implies \ell - \epsilon' < f(x) < \ell + \epsilon'.$$

Si on choisit ϵ' tel que $0 < \epsilon' < \ell/2$, alors on voit qu'il existe un intervalle $J = I \cap]x_0 - \delta, x_0 + \delta[$ tel que pour tout x dans J, $f(x) > \ell/2 > 0$, c'est-à-dire, en posant $M = 2/\ell$:

$$\forall x \in J \quad 0 < \frac{1}{f(x)} < M.$$

Fixons à présent $\epsilon > 0$. Pour tout $x \in J$, on a

$$\left|\frac{1}{f(x)} - \frac{1}{\ell}\right| = \frac{|\ell - f(x)|}{f(x)\ell} < \frac{M}{\ell}|\ell - f(x)|.$$

Donc, si dans la définition précédente de la limite de f en x_0 on choisit $\epsilon' = \frac{\ell\epsilon}{M}$, alors on trouve qu'il existe un $\delta > 0$ tel que

$$\forall x \in J \quad x_0 - \delta < x < x_0 + \delta \implies \left|\frac{1}{f(x)} - \frac{1}{\ell}\right| < \frac{M}{\ell}|\ell - f(x)| < \frac{M}{\ell}\epsilon' = \epsilon.$$

□

Proposition 3.
Si $\lim_{x_0} f = \ell$ *et* $\lim_{\ell} g = \ell'$*, alors* $\lim_{x_0} g \circ f = \ell'$.

Ce sont des propriétés que l'on utilise sans s'en apercevoir !

Exemple 9.
Soit $x \mapsto u(x)$ une fonction et $x_0 \in \mathbb{R}$ tel que $u(x) \to 2$ lorsque $x \to x_0$. Posons $f(x) = \sqrt{1 + \frac{1}{u(x)^2} + \ln u(x)}$. Si elle existe, quelle est la limite de f en x_0 ?

- Tout d'abord comme $u(x) \to 2$ alors $u(x)^2 \to 4$ donc $\frac{1}{u(x)^2} \to \frac{1}{4}$ (lorsque $x \to x_0$).
- De même comme $u(x) \to 2$ alors, dans un voisinage de x_0, $u(x) > 0$ donc $\ln u(x)$ est bien définie dans ce voisinage et de plus $\ln u(x) \to \ln 2$ (lorsque $x \to x_0$).
- Cela entraîne que $1 + \frac{1}{u(x)^2} + \ln u(x) \to 1 + \frac{1}{4} + \ln 2$ lorsque $x \to x_0$. En particulier $1 + \frac{1}{u(x)^2} + \ln u(x) \geqslant 0$ dans un voisinage de x_0, donc $f(x)$ est bien définie dans un voisinage de x_0.
- Et par composition avec la racine carrée alors $f(x)$ a bien une limite en x_0 et $\lim_{x \to x_0} f(x) = \sqrt{1 + \frac{1}{4} + \ln 2}$.

Il y a des situations où l'on ne peut rien dire sur les limites. Par exemple si $\lim_{x_0} f = +\infty$ et $\lim_{x_0} g = -\infty$ alors on ne peut à priori rien dire sur la limite de $f + g$ (cela dépend vraiment de f et de g). On raccourci cela en $+\infty - \infty$ est une ***forme indéterminée***.

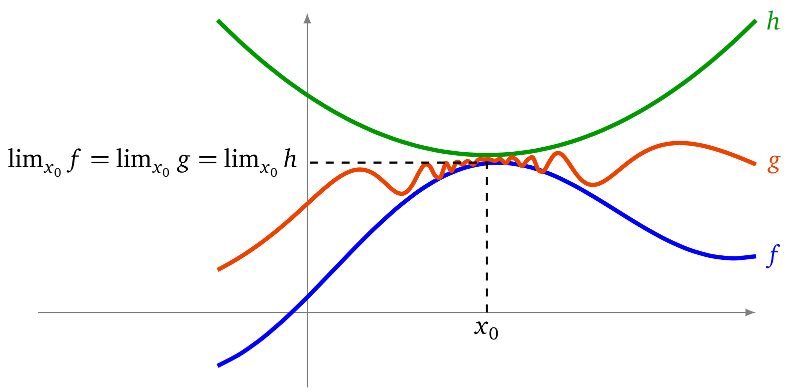
Voici une liste de formes indéterminées : $+\infty - \infty$; $0 \times \infty$; $\frac{\infty}{\infty}$; $\frac{0}{0}$; 1^∞ ; ∞^0.

Enfin voici une proposition très importante qui signifie qu'on peut passer à la limite dans une inégalité *large*.

Proposition 4.

- *Si* $f \leqslant g$ *et si* $\lim_{x_0} f = \ell \in \mathbb{R}$ *et* $\lim_{x_0} g = \ell' \in \mathbb{R}$*, alors* $\ell \leqslant \ell'$.
- *Si* $f \leqslant g$ *et si* $\lim_{x_0} f = +\infty$*, alors* $\lim_{x_0} g = +\infty$.
- *Théorème des gendarmes*

$$\boxed{\text{Si } f \leqslant g \leqslant h \text{ et si } \lim_{x_0} f = \lim_{x_0} h = \ell \in \mathbb{R}, \text{ alors } g \text{ a une limite en } x_0 \text{ et } \lim_{x_0} g = \ell.}$$



Mini-exercices.

1. Déterminer, si elle existe, la limite de $\frac{2x^2-x-2}{3x^2+2x+2}$ en 0. Et en $+\infty$?
2. Déterminer, si elle existe, la limite de $\sin\left(\frac{1}{x}\right)$ en $+\infty$. Et pour $\frac{\cos x}{\sqrt{x}}$?
3. En utilisant la définition de la limite (avec des ϵ), montrer que $\lim_{x \to 2}(3x+1) = 7$.
4. Montrer que si f admet une limite finie en x_0 alors il existe $\delta > 0$ tel que f soit bornée sur $]x_0 - \delta, x_0 + \delta[$.

5. Déterminer, si elle existe, $\lim_{x\to 0} \frac{\sqrt{1+x}-\sqrt{1+x^2}}{x}$. Et $\lim_{x\to 2} \frac{x^2-4}{x^2-3x+2}$?

3. Continuité en un point

3.1. Définition

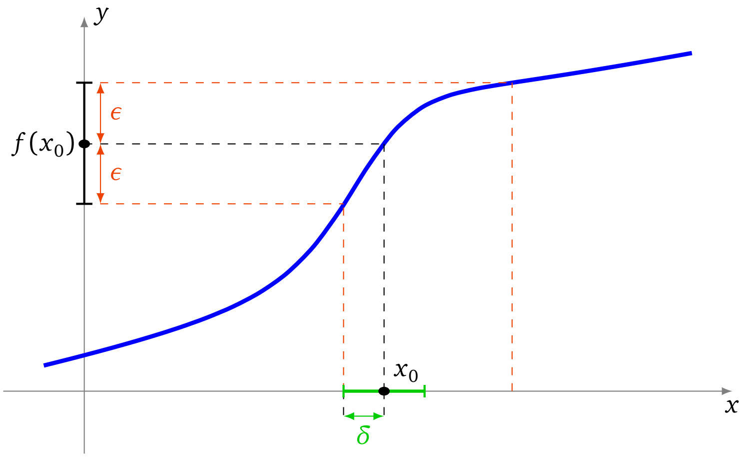
Soit I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction.

Définition 11.

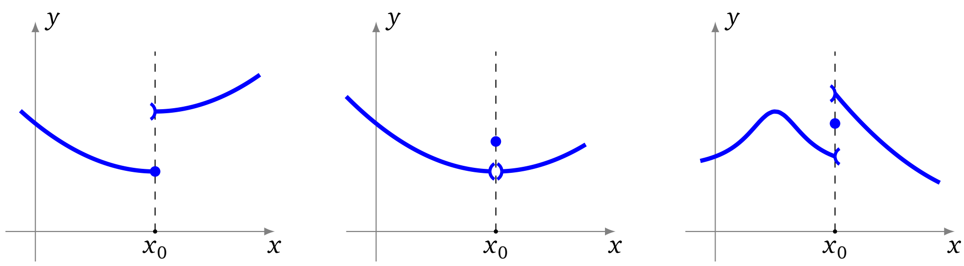
- On dit que f est ***continue en un point*** $x_0 \in I$ si

$$\boxed{\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in I \quad |x - x_0| < \delta \implies |f(x) - f(x_0)| < \epsilon}$$

  c'est-à-dire si f admet une limite en x_0 (cette limite vaut alors nécessairement $f(x_0)$).
- On dit que f est ***continue sur*** I si f est continue en tout point de I.



Intuitivement, une fonction est continue sur un intervalle, si on peut tracer son graphe « sans lever le crayon », c'est-à-dire si sa courbe représentative n'admet pas de saut.
Voici des fonctions qui ne sont pas continues en x_0 :



Exemple 10.
Les fonctions suivantes sont continues :

- une fonction constante sur un intervalle,
- la fonction racine carrée $x \mapsto \sqrt{x}$ sur $[0, +\infty[$,
- les fonctions sin et cos sur $\mathbb{R}$,
- la fonction valeur absolue $x \mapsto |x|$ sur $\mathbb{R}$,
- la fonction exp sur $\mathbb{R}$,
- la fonction ln sur $]0, +\infty[$.

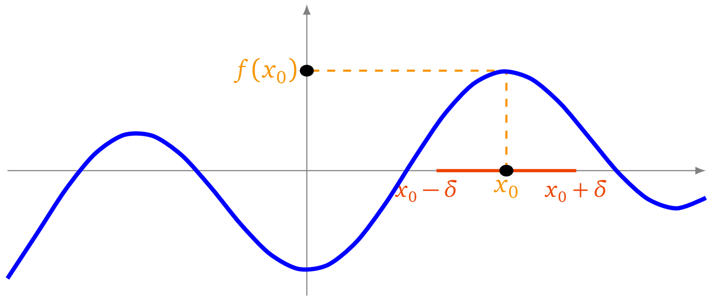
Par contre, la fonction partie entière E n'est pas continue aux points $x_0 \in \mathbb{Z}$, puisqu'elle n'admet pas de limite en ces points. Pour $x_0 \in \mathbb{R} \setminus \mathbb{Z}$, elle est continue en x_0.

3.2. Propriétés

La continuité assure par exemple que si la fonction n'est pas nulle en un point (qui est une propriété ponctuelle) alors elle n'est pas nulle autour de ce point (propriété locale). Voici l'énoncé :

Lemme 1.
Soit $f : I \to \mathbb{R}$ une fonction définie sur un intervalle I et x_0 un point de I. Si f est continue en x_0 et si $f(x_0) \neq 0$, alors il existe $\delta > 0$ tel que

$$\forall x \in]x_0 - \delta, x_0 + \delta[\quad f(x) \neq 0$$



Démonstration. Supposons par exemple que $f(x_0) > 0$, le cas $f(x_0) < 0$ se montrerait de la même manière. Écrivons ainsi la définition de la continuité de f en x_0 :

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in I \quad x \in]x_0 - \delta, x_0 + \delta[\implies f(x_0) - \epsilon < f(x) < f(x_0) + \epsilon.$$

Il suffit donc de choisir ϵ tel que $0 < \epsilon < f(x_0)$. Il existe alors bien un intervalle $J = I \cap]x_0 - \delta, x_0 + \delta[$ tel que pour tout x dans J, on a $f(x) > 0$. □

La continuité se comporte bien avec les opérations élémentaires. Les propositions suivantes sont des conséquences immédiates des propositions analogues sur les limites.

Proposition 5.
Soient $f, g : I \to \mathbb{R}$ deux fonctions continues en un point $x_0 \in I$. Alors
- *$\lambda \cdot f$ est continue en x_0 (pour tout $\lambda \in \mathbb{R}$),*
- *$f + g$ est continue en x_0,*
- *$f \times g$ est continue en x_0,*
- *si $f(x_0) \neq 0$, alors $\frac{1}{f}$ est continue en x_0.*

Exemple 11.
La proposition précédente permet de vérifier que d'autres fonctions usuelles sont continues :
- les fonctions puissance $x \mapsto x^n$ sur $\mathbb{R}$ (comme produit $x \times x \times \cdots$),
- les polynômes sur $\mathbb{R}$ (somme et produit de fonctions puissance et de fonctions constantes),
- les fractions rationnelles $x \mapsto \frac{P(x)}{Q(x)}$ sur tout intervalle où le polynôme $Q(x)$ ne s'annule pas.

La composition conserve la continuité (mais il faut faire attention en quels points les hypothèses s'appliquent).

Proposition 6.
Soient $f : I \to \mathbb{R}$ et $g : J \to \mathbb{R}$ deux fonctions telles que $f(I) \subset J$. Si f est continue en un point $x_0 \in I$ et si g est continue en $f(x_0)$, alors $g \circ f$ est continue en x_0.

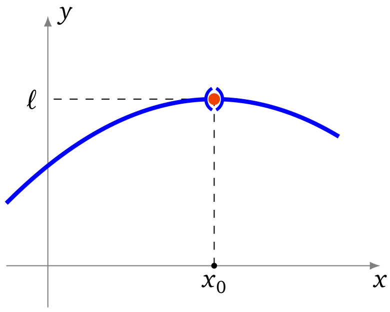
3.3. Prolongement par continuité

Définition 12.
Soit I un intervalle, x_0 un point de I et $f : I \setminus \{x_0\} \to \mathbb{R}$ une fonction.

- On dit que f est ***prolongeable par continuité*** en x_0 si f admet une limite finie en x_0. Notons alors $\ell = \lim_{x_0} f$.
- On définit alors la fonction $\tilde{f} : I \to \mathbb{R}$ en posant pour tout $x \in I$

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \neq x_0 \\ \ell & \text{si } x = x_0. \end{cases}$$

Alors $\tilde{f}$ est continue en x_0 et on l'appelle le ***prolongement par continuité*** de f en x_0.



Dans la pratique, on continuera souvent à noter f à la place de $\tilde{f}$.

Exemple 12.
Considérons la fonction f définie sur $\mathbb{R}^*$ par $f(x) = x \sin\left(\frac{1}{x}\right)$. Voyons si f admet un prolongement par continuité en 0 ?
Comme pour tout $x \in \mathbb{R}^*$ on a $|f(x)| \leqslant |x|$, on en déduit que f tend vers 0 en 0. Elle est donc prolongeable par continuité en 0 et son prolongement est la fonction $\tilde{f}$ définie sur $\mathbb{R}$ tout entier par :

$$\tilde{f}(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases}$$

3.4. Suites et continuité

Proposition 7.
Soit $f : I \to \mathbb{R}$ une fonction et x_0 un point de I. Alors :

$$\boxed{f \text{ est continue en } x_0 \iff \begin{array}{c} \text{pour toute suite } (u_n) \text{ qui converge vers } x_0 \\ \text{la suite } (f(u_n)) \text{ converge vers } f(x_0) \end{array}}$$

Démonstration.

$\Longrightarrow$ *On suppose que f est continue en x_0 et que (u_n) est une suite qui converge vers x_0 et on veut montrer que $(f(u_n))$ converge vers $f(x_0)$.*
Soit $\epsilon > 0$. Comme f est continue en x_0, il existe un $\delta > 0$ tel que

$$\forall x \in I \quad |x - x_0| < \delta \implies |f(x) - f(x_0)| < \epsilon.$$

Pour ce δ, comme (u_n) converge vers x_0, il existe $N \in \mathbb{N}$ tel que

$$\forall n \in \mathbb{N} \quad n \geqslant N \implies |u_n - x_0| < \delta.$$

On en déduit que, pour tout $n \geqslant N$, comme $|u_n - x_0| < \delta$, on a $|f(u_n) - f(x_0)| < \epsilon$. Comme c'est vrai pour tout $\epsilon > 0$, on peut maintenant conclure que $(f(u_n))$ converge vers $f(x_0)$.

$\Longleftarrow$ *On va montrer la contraposée : supposons que f n'est pas continue en x_0 et montrons qu'alors il existe une suite (u_n) qui converge vers x_0 et telle que $(f(u_n))$ ne converge pas vers $f(x_0)$.*

Par hypothèse, comme f n'est pas continue en x_0 :

$$\exists \epsilon_0 > 0 \quad \forall \delta > 0 \quad \exists x_\delta \in I \quad \text{tel que} \quad |x_\delta - x_0| < \delta \text{ et } |f(x_\delta) - f(x_0)| > \epsilon_0.$$

On construit la suite (u_n) de la façon suivante : pour tout $n \in \mathbb{N}^*$, on choisit dans l'assertion précédente $\delta = 1/n$ et on obtient qu'il existe u_n (qui est $x_{1/n}$) tel que

$$|u_n - x_0| < \frac{1}{n} \quad \text{et} \quad |f(u_n) - f(x_0)| > \epsilon_0.$$

La suite (u_n) converge vers x_0 alors que la suite $(f(u_n))$ ne peut pas converger vers $f(x_0)$.

□

Remarque.

On retiendra surtout l'implication : si f est continue sur I et si (u_n) est une suite convergente de limite ℓ, alors $(f(u_n))$ converge vers $f(\ell)$. On l'utilisera intensivement pour l'étude des suites récurrentes $u_{n+1} = f(u_n)$: si f est continue et $u_n \to \ell$, alors $f(\ell) = \ell$.

Mini-exercices.

1. Déterminer le domaine de définition et de continuité des fonctions suivantes : $f(x) = 1/\sin x$, $g(x) = 1/\sqrt{x + \frac{1}{2}}$, $h(x) = \ln(x^2 + x - 1)$.
2. Trouver les couples $(a, b) \in \mathbb{R}^2$ tels que la fonction f définie sur $\mathbb{R}$ par $f(x) = ax + b$ si $x < 0$ et $f(x) = \exp(x)$ si $x \geqslant 0$ soit continue sur $\mathbb{R}$. Et si on avait $f(x) = \frac{a}{x-1} + b$ pour $x < 0$?
3. Soit f une fonction continue telle que $f(x_0) = 1$. Montrer qu'il existe $\delta > 0$ tel que : pour tout $x \in]x_0 - \delta, x_0 + \delta[\quad f(x) > \frac{1}{2}$.
4. Étudier la continuité de $f : \mathbb{R} \to \mathbb{R}$ définie par : $f(x) = \sin(x)\cos\left(\frac{1}{x}\right)$ si $x \neq 0$ et $f(0) = 0$. Et pour $g(x) = xE(x)$?
5. La fonction définie par $f(x) = \frac{x^3+8}{|x+2|}$ admet-elle un prolongement par continuité en -2 ?
6. Soit la suite définie par $u_0 > 0$ et $u_{n+1} = \sqrt{u_n}$. Montrer que (u_n) admet une limite $\ell \in \mathbb{R}$ lorsque $n \to +\infty$. À l'aide de la fonction $f(x) = \sqrt{x}$ calculer cette limite.

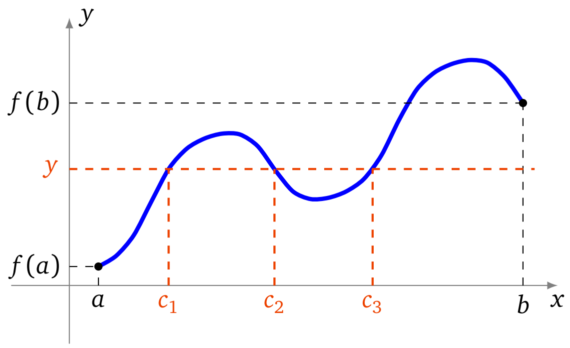
4. Continuité sur un intervalle

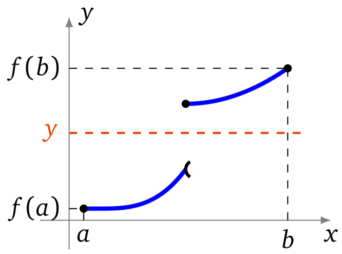
4.1. Le théorème des valeurs intermédiaires

Théorème 1 (Théorème des valeurs intermédiaires)**.**
Soit $f : [a, b] \to \mathbb{R}$ une fonction continue sur un segment.

Pour tout réel y compris entre $f(a)$ et $f(b)$,
il existe $c \in [a, b]$ tel que $f(c) = y$.

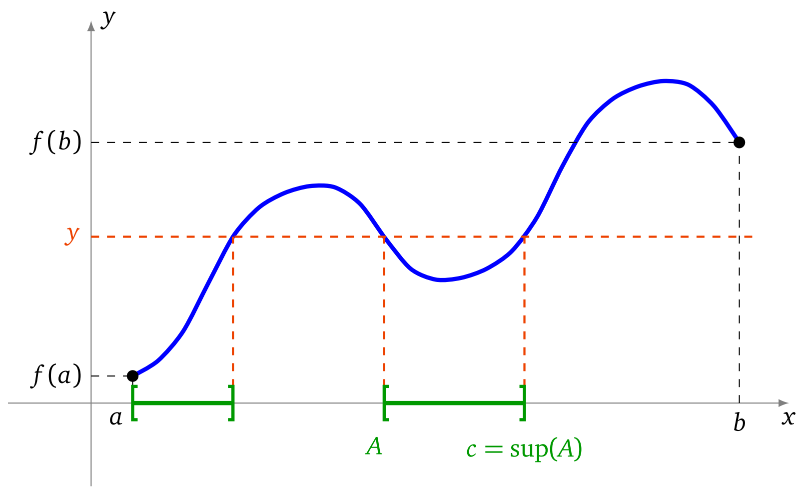
Une illustration du théorème des valeurs intermédiaires (figure de gauche), le réel c n'est pas nécessairement unique. De plus si la fonction n'est pas continue, le théorème n'est plus vrai (figure de droite).





Démonstration. Montrons le théorème dans le cas où $f(a) < f(b)$. On considère alors un réel y tel que $f(a) \leqslant y \leqslant f(b)$ et on veut montrer qu'il a un antécédent par f.

1. On introduit l'ensemble suivant
$$A = \Big\{x \in [a,b] \mid f(x) \leqslant y\Big\}.$$
Tout d'abord l'ensemble A est non vide (car $a \in A$) et il est majoré (car il est contenu dans $[a,b]$) : il admet donc une borne supérieure, que l'on note $c = \sup A$. Montrons que $f(c) = y$.



2. Montrons tout d'abord que $f(c) \leqslant y$. Comme $c = \sup A$, il existe une suite $(u_n)_{n\in\mathbb{N}}$ contenue dans A telle que (u_n) converge vers c. D'une part, pour tout $n \in \mathbb{N}$, comme $u_n \in A$, on a $f(u_n) \leqslant y$. D'autre part, comme f est continue en c, la suite $(f(u_n))$ converge vers $f(c)$. On en déduit donc, par passage à la limite, que $f(c) \leqslant y$.
3. Montrons à présent que $f(c) \geqslant y$. Remarquons tout d'abord que si $c = b$, alors on a fini, puisque $f(b) \geqslant y$. Sinon, pour tout $x \in]c,b]$, comme $x \notin A$, on a $f(x) > y$. Or, étant donné que f est continue en c, f admet une limite à droite en c, qui vaut $f(c)$ et on obtient $f(c) \geqslant y$.

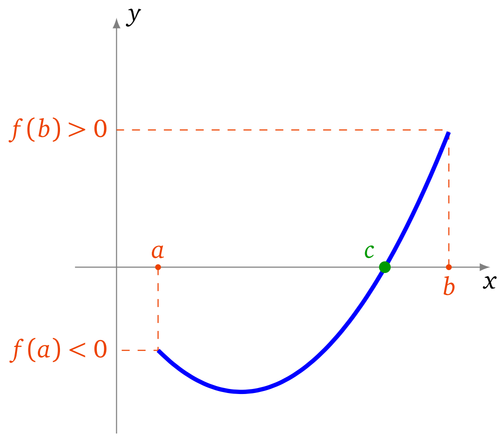
□

4.2. Applications du théorème des valeurs intermédiaires

Voici la version la plus utilisée du théorème des valeurs intermédiaires.

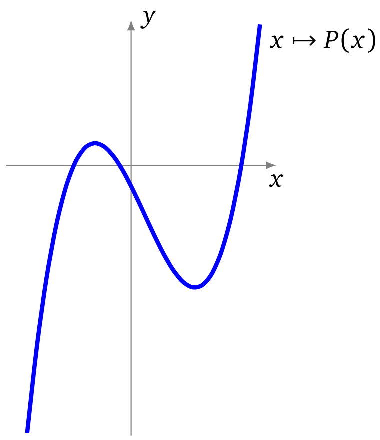
Corollaire 1.
Soit $f : [a,b] \to \mathbb{R}$ une fonction continue sur un segment.

$$\boxed{\text{Si } f(a)\cdot f(b) < 0, \text{ alors il existe } c \in]a,b[\text{ tel que } f(c) = 0.}$$



Démonstration. Il s'agit d'une application directe du théorème des valeurs intermédiaires avec $y = 0$. L'hypothèse $f(a) \cdot f(b) < 0$ signifiant que $f(a)$ et $f(b)$ sont de signes contraires. □

Exemple 13.
Tout polynôme de degré impair possède au moins une racine réelle.



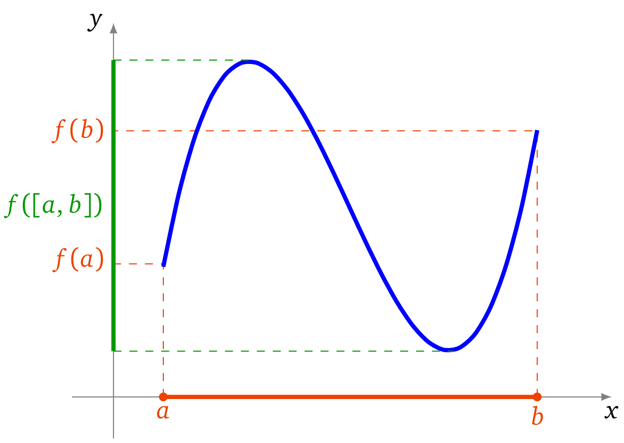
En effet, un tel polynôme s'écrit $P(x) = a_n x^n + \cdots + a_1 x + a_0$ avec n un entier impair. On peut supposer que le coefficient a_n est strictement positif. Alors on a $\lim_{-\infty} P = -\infty$ et $\lim_{+\infty} P = +\infty$. En particulier, il existe deux réels a et b tels que $f(a) < 0$ et $f(b) > 0$ et on conclut grâce au corollaire précédent.

Voici une formulation théorique du théorème des valeurs intermédiaires.

Corollaire 2.

Soit $f : I \to \mathbb{R}$ une fonction continue sur un intervalle I. Alors $f(I)$ est un intervalle.

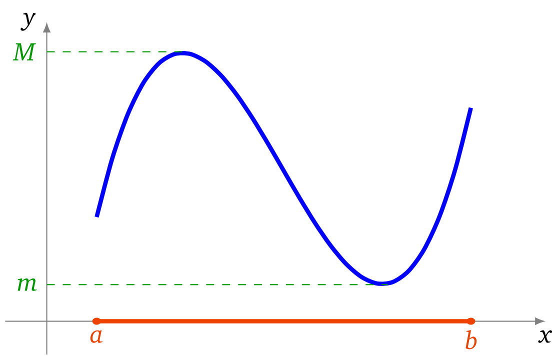
Attention ! Il serait faux de croire que l'image par une fonction f de l'intervalle $[a, b]$ soit l'intervalle $[f(a), f(b)]$ (voir la figure ci-dessous).



Démonstration. Soient $y_1, y_2 \in f(I)$, $y_1 \leqslant y_2$. Montrons que si $y \in [y_1, y_2]$, alors $y \in f(I)$. Par hypothèse, il existe $x_1, x_2 \in I$ tels que $y_1 = f(x_1)$, $y_2 = f(x_2)$ et donc y est compris entre $f(x_1)$ et $f(x_2)$. D'après le théorème des valeurs intermédiaires, comme f est continue, il existe donc $x \in I$ tel que $y = f(x)$, et ainsi $y \in f(I)$. □

4.3. Fonctions continues sur un segment

Théorème 2.
Soit $f : [a,b] \to \mathbb{R}$ une fonction continue sur un segment. Alors il existe deux réels m et M tels que $f([a,b]) = [m, M]$. Autrement dit, l'image d'un segment par une fonction continue est un segment.



Comme on sait déjà par le théorème des valeurs intermédiaires que $f([a,b])$ est un intervalle, le théorème précédent signifie exactement que

Si f est continue sur $[a,b]$ alors f est bornée sur $[a,b]$, et elle atteint ses bornes.

Donc m est le minimum de la fonction sur l'intervalle $[a,b]$ alors que M est le maximum.

Démonstration.

1. Montrons d'abord que f est bornée.
   - Pour $r \in \mathbb{R}$, on note $A_r = \{x \in [a,b] \mid f(x) \geqslant r\}$. Fixons r tel que $A_r \neq \varnothing$, comme $A_r \subset [a,b]$, le nombre $s = \sup A_r$ existe. Soit $x_n \to s$ avec $x_n \in A_r$. Par définition $f(x_n) \geqslant r$ donc, f étant continue, à la limite $f(s) \geqslant r$ et ainsi $s \in A_r$.

- Supposons par l'absurde que f ne soit pas bornée. Alors pour tout $n \geqslant 0$, A_n est non vide. Notons $s_n = \sup A_n$. Comme $f(x) \geqslant n+1$ implique $f(x) \geqslant n$ alors $A_{n+1} \subset A_n$, ce qui entraîne $s_{n+1} \leqslant s_n$. Bilan : (s_n) est une suite décroissante, minorée par a donc converge vers $\ell \in [a,b]$. Encore une fois f est continue donc $s_n \to \ell$, implique $f(s_n) \to f(\ell)$. Mais $f(s_n) \geqslant n$ donc $\lim f(s_n) = +\infty$. Cela contredit $\lim f(s_n) = f(\ell) < +\infty$. Conclusion : f est majorée.
- Un raisonnement tout à fait similaire prouve que f est aussi minorée, donc bornée. Par ailleurs on sait déjà que $f(I)$ est un intervalle (c'est le théorème des valeurs intermédiaires), donc maintenant $f(I)$ est un intervalle borné. Il reste à montrer qu'il du type $[m,M]$ (et pas $]m,M[$ par exemple).

2. Montrons maintenant que $f(I)$ est un intervalle fermé. Sachant déjà que $f(I)$ est un intervalle borné, notons m et M ses extrémités : $m = \inf f(I)$ et $M = \sup f(I)$. Supposons par l'absurde que $M \notin f(I)$. Alors pour $t \in [a,b]$, $M > f(t)$. La fonction $g : t \mapsto \frac{1}{M-f(t)}$ est donc bien définie. La fonction g est continue sur I donc d'après le premier point de cette preuve (appliqué à g) elle est bornée, disons par un réel K. Mais il existe $y_n \to M$, $y_n \in f(I)$. Donc il existe $x_n \in [a,b]$ tel que $y_n = f(x_n) \to M$ et alors $g(x_n) = \frac{1}{M-f(x_n)} \to +\infty$. Cela contredit que g soit une fonction bornée par K. Bilan : $M \in f(I)$. De même on a $m \in f(I)$. Conclusion finale : $f(I) = [m,M]$.

□

Mini-exercices.

1. Soient $P(x) = x^5 - 3x - 2$ et $f(x) = x2^x - 1$ deux fonctions définies sur $\mathbb{R}$. Montrer que l'équation $P(x) = 0$ a au moins une racine dans $[1,2]$; l'équation $f(x) = 0$ a au moins une racine dans $[0,1]$; l'équation $P(x) = f(x)$ a au moins une racine dans $]0,2[$.
2. Montrer qu'il existe $x > 0$ tel que $2^x + 3^x = 7^x$.
3. Dessiner le graphe d'une fonction continue $f : \mathbb{R} \to \mathbb{R}$ tel que $f(\mathbb{R}) = [0,1]$. Puis $f(\mathbb{R}) =]0,1[$; $f(\mathbb{R}) = [0,1[$; $f(\mathbb{R}) =]-\infty,1]$, $f(\mathbb{R}) =]-\infty,1[$.
4. Soient $f,g : [0,1] \to \mathbb{R}$ deux fonctions continues. Quelles sont, parmi les fonctions suivantes, celles dont on peut affirmer qu'elles sont bornées : $f+g$, $f \times g$, f/g ?
5. Soient f et g deux fonctions continues sur $[0,1]$ telles que $\forall x \in [0,1]\ f(x) < g(x)$. Montrer qu'il existe $m > 0$ tel que $\forall x \in [0,1]\ f(x) + m < g(x)$. Ce résultat est-il vrai si on remplace $[0,1]$ par $\mathbb{R}$?

5. Fonctions monotones et bijections

5.1. Rappels : injection, surjection, bijection

Dans cette section nous rappelons le matériel nécessaire concernant les applications bijectives.

Définition 13.
Soit $f : E \to F$ une fonction, où E et F sont des parties de $\mathbb{R}$.

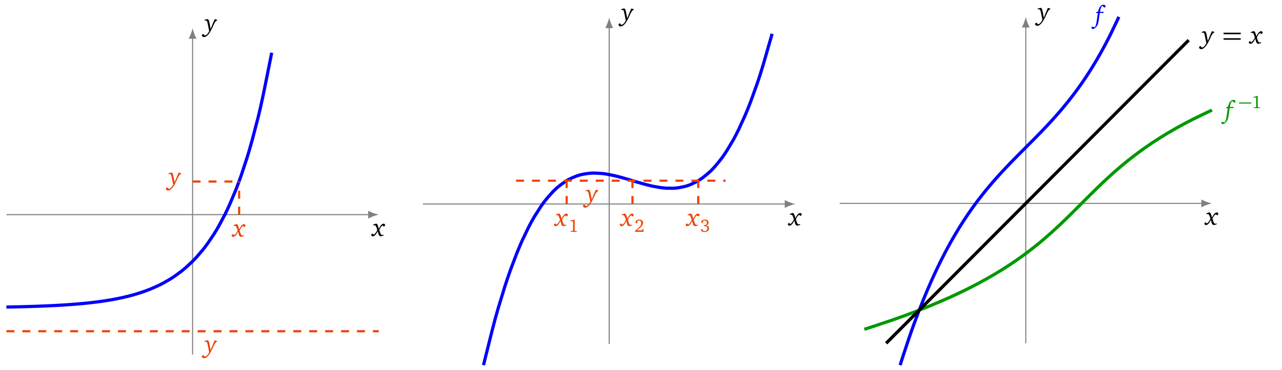
- f est *injective* si $\forall x, x' \in E \quad f(x) = f(x') \implies x = x'$;
- f est *surjective* si $\forall y \in F \quad \exists x \in E \quad y = f(x)$;
- f est *bijective* si f est à la fois injective et surjective, c'est-à-dire si $\forall y \in F \quad \exists! x \in E \quad y = f(x)$.

Proposition 8.
Si $f : E \to F$ est une fonction bijective alors il existe une unique application $g : F \to E$ telle que $g \circ f = \mathrm{id}_E$ et $f \circ g = \mathrm{id}_F$. La fonction g est la **bijection réciproque** *de f et se note f^{-1}.*

Remarque.

- On rappelle que l'*identité*, $\mathrm{id}_E : E \to E$ est simplement définie par $x \mapsto x$.
- $g \circ f = \mathrm{id}_E$ se reformule ainsi : $\forall x \in E \quad g\big(f(x)\big) = x$.
- Alors que $f \circ g = \mathrm{id}_F$ s'écrit : $\forall y \in F \quad f\big(g(y)\big) = y$.
- Dans un repère orthonormé les graphes des fonctions f et f^{-1} sont symétriques par rapport à la première bissectrice.

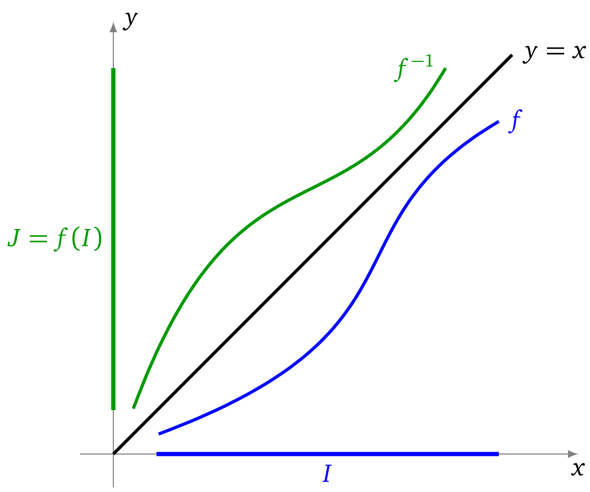
Voici le graphe d'une fonction injective (à gauche), d'une fonction surjective (à droite) et enfin le graphe d'une fonction bijective ainsi que le graphe de sa bijection réciproque.



5.2. Fonctions monotones et bijections

Voici un théorème très utilisé dans la pratique pour montrer qu'une fonction est bijective.

> **Théorème 3** (Théorème de la bijection).
> *Soit $f : I \to \mathbb{R}$ une fonction définie sur un intervalle I de $\mathbb{R}$. Si f est continue et strictement monotone sur I, alors*
>
> 1. *f établit une bijection de l'intervalle I dans l'intervalle image $J = f(I)$,*
> 2. *la fonction réciproque $f^{-1} : J \to I$ est continue et strictement monotone sur J et elle a le même sens de variation que f.*



En pratique, si on veut appliquer ce théorème à une fonction continue $f : I \to \mathbb{R}$, on découpe l'intervalle I en sous-intervalles sur lesquels la fonction f est strictement monotone.

Exemple 14.
Considérons la fonction carrée définie sur $\mathbb{R}$ par $f(x) = x^2$. La fonction f n'est pas strictement monotone sur $\mathbb{R}$: elle n'est pas même pas injective car un nombre et son opposé ont même carré. Cependant, en

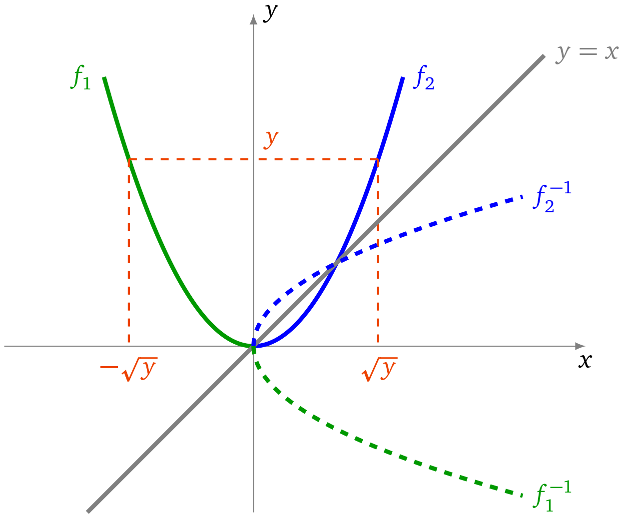
restreignant son ensemble de définition à $]-\infty,0]$ d'une part et à $[0,+\infty[$ d'autre part, on définit deux fonctions strictement monotones :

$$f_1 : \left\{ \begin{array}{r}]-\infty,0] \longrightarrow [0,+\infty[\\ x \longmapsto x^2 \end{array} \right. \qquad \text{et} \qquad f_2 : \left\{ \begin{array}{r} [0,+\infty[\longrightarrow [0,+\infty[\\ x \longmapsto x^2 \end{array} \right.$$

On remarque que $f(]-\infty,0]) = f([0,+\infty[) = [0,+\infty[$. D'après le théorème précédent, les fonctions f_1 et f_2 sont des bijections. Déterminons leurs fonctions réciproques $f_1^{-1} : [0,+\infty[\to]-\infty,0]$ et $f_2^{-1} : [0,+\infty[\to [0,+\infty[$. Soient deux réels x et y tels que $y \geqslant 0$. Alors

$$\begin{aligned} y = f(x) &\iff y = x^2 \\ &\iff x = \sqrt{y} \quad \text{ou} \quad x = -\sqrt{y}, \end{aligned}$$

c'est-à-dire y admet (au plus) deux antécédents, l'un dans $[0,+\infty[$ et l'autre dans $]-\infty,0]$. Et donc $f_1^{-1}(y) = -\sqrt{y}$ et $f_2^{-1}(y) = \sqrt{y}$. On vérifie bien que chacune des deux fonctions f_1 et f_2 a le même sens de variation que sa réciproque.



On remarque que la courbe totale en pointillé (à la fois la partie bleue et la verte), qui est l'image du graphe de f par la symétrie par rapport à la première bissectrice, ne peut pas être le graphe d'une fonction : c'est une autre manière de voir que f n'est pas bijective.

Généralisons en partie l'exemple précédent.

Exemple 15.

Soit $n \geqslant 1$. Soit $f : [0,+\infty[\to [0,+\infty[$ définie par $f(x) = x^n$. Alors f est continue et strictement croissante. Comme $\lim_{+\infty} f = +\infty$ alors f est une bijection. Sa bijection réciproque f^{-1} est notée : $x \mapsto x^{\frac{1}{n}}$ (ou aussi $x \mapsto \sqrt[n]{x}$) : c'est la fonction racine n-ième. Elle est continue et strictement croissante.

5.3. Démonstration

On établit d'abord un lemme utile à la démonstration du « théorème de la bijection ».

Lemme 2.

Soit $f : I \to \mathbb{R}$ une fonction définie sur un intervalle I de $\mathbb{R}$. Si f est strictement monotone sur I, alors f est injective sur I.

Démonstration. Soient $x, x' \in I$ tels que $f(x) = f(x')$. Montrons que $x = x'$. Si on avait $x < x'$, alors on aurait nécessairement $f(x) < f(x')$ ou $f(x) > f(x')$, suivant que f est strictement croissante, ou strictement décroissante. Comme c'est impossible, on en déduit que $x \geqslant x'$. En échangeant les rôles de x et de x', on montre de même que $x \leqslant x'$. On en conclut que $x = x'$ et donc que f est injective. $\square$

Démonstration du théorème.

1. D'après le lemme précédent, f est injective sur I. En restreignant son ensemble d'arrivée à son image $J = f(I)$, on obtient que f établit une bijection de I dans J. Comme f est continue, par le théorème des valeurs intermédiaires, l'ensemble J est un intervalle.

2. Supposons pour fixer les idées que f est strictement croissante.

   (a) Montrons que f^{-1} est strictement croissante sur J. Soient $y, y' \in J$ tels que $y < y'$. Notons $x = f^{-1}(y) \in I$ et $x' = f^{-1}(y') \in I$. Alors $y = f(x)$, $y' = f(x')$ et donc

   $$\begin{aligned} y < y' &\implies f(x) < f(x') \\ &\implies x < x' \qquad \text{(car } f \text{ est strictement croissante)} \\ &\implies f^{-1}(y) < f^{-1}(y'), \end{aligned}$$

   c'est-à-dire f^{-1} est strictement croissante sur J.

   (b) Montrons que f^{-1} est continue sur J. On se limite au cas où I est de la forme $]a, b[$, les autres cas se montrent de la même manière. Soit $y_0 \in J$. On note $x_0 = f^{-1}(y_0) \in I$. Soit $\epsilon > 0$. On peut toujours supposer que $[x_0 - \epsilon, x_0 + \epsilon] \subset I$. On cherche un réel $\delta > 0$ tel que pour tout $y \in J$ on ait

   $$y_0 - \delta < y < y_0 + \delta \implies f^{-1}(y_0) - \epsilon < f^{-1}(y) < f^{-1}(y_0) + \epsilon$$

   c'est-à-dire tel que pour tout $x \in I$ on ait

   $$y_0 - \delta < f(x) < y_0 + \delta \implies f^{-1}(y_0) - \epsilon < x < f^{-1}(y_0) + \epsilon.$$

   Or, comme f est strictement croissante, on a pour tout $x \in I$

   $$\begin{aligned} f(x_0 - \epsilon) < f(x) < f(x_0 + \epsilon) &\implies x_0 - \epsilon < x < x_0 + \epsilon \\ &\implies f^{-1}(y_0) - \epsilon < x < f^{-1}(y_0) + \epsilon. \end{aligned}$$

   Comme $f(x_0 - \epsilon) < y_0 < f(x_0 + \epsilon)$, on peut choisir le réel $\delta > 0$ tel que

   $$f(x_0 - \epsilon) < y_0 - \delta \quad \text{et} \quad f(x_0 + \epsilon) > y_0 + \delta$$

   et on a bien alors pour tout $x \in I$

   $$\begin{aligned} y_0 - \delta < f(x) < y_0 + \delta &\implies f(x_0 - \epsilon) < f(x) < f(x_0 + \epsilon) \\ &\implies f^{-1}(y_0) - \epsilon < x < f^{-1}(y_0) + \epsilon. \end{aligned}$$

   La fonction f^{-1} est donc continue sur J.

□

Mini-exercices.

1. Montrer que chacune des hypothèses « continue » et « strictement monotone » est nécessaire dans l'énoncé du théorème de la bijection.
2. Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^3 + x$. Montrer que f est bijective, tracer le graphe de f et de f^{-1}.
3. Soit $n \geqslant 1$. Montrer que $f(x) = 1 + x + x^2 + \cdots + x^n$ définit une bijection de l'intervalle $[0, 1]$ vers un intervalle à préciser.
4. Existe-t-il une fonction continue : $f : [0, 1[\to]0, 1[$ qui soit bijective ? $f : [0, 1[\to]0, 1[$ qui soit injective ? $f :]0, 1[\to [0, 1]$ qui soit surjective ?
5. Pour $y \in \mathbb{R}$ on considère l'équation $x + \exp x = y$. Montrer qu'il existe une unique solution y. Comment varie y en fonction de x ? Comme varie x en fonction de y ?

Auteurs du chapitre

Auteurs : Arnaud Bodin, Niels Borne, Laura Desideri
Dessins : Benjamin Boutin

Chapitre

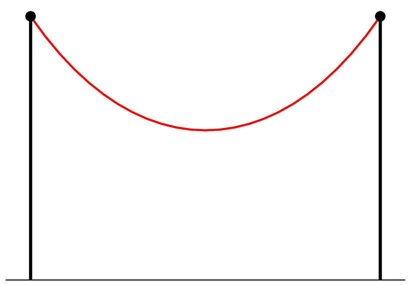
2

Fonctions usuelles

Vous connaissez déjà des fonctions classiques : exp, ln, cos, sin, tan. Dans ce chapitre il s'agit d'ajouter à notre catalogue de nouvelles fonctions : ch, sh, th, arccos, arcsin, arctan, Argch, Argsh, Argth.

Ces fonctions apparaissent naturellement dans la résolution de problèmes simples, en particulier issus de la physique. Par exemple lorsqu'un fil est suspendu entre deux poteaux (ou un collier tenu entre deux mains) alors la courbe dessinée est une *chaînette* dont l'équation fait intervenir le cosinus hyperbolique et un paramètre a (qui dépend de la longueur du fil et de l'écartement des poteaux) :

$$y = a \operatorname{ch}\left(\frac{x}{a}\right)$$



1. Logarithme et exponentielle

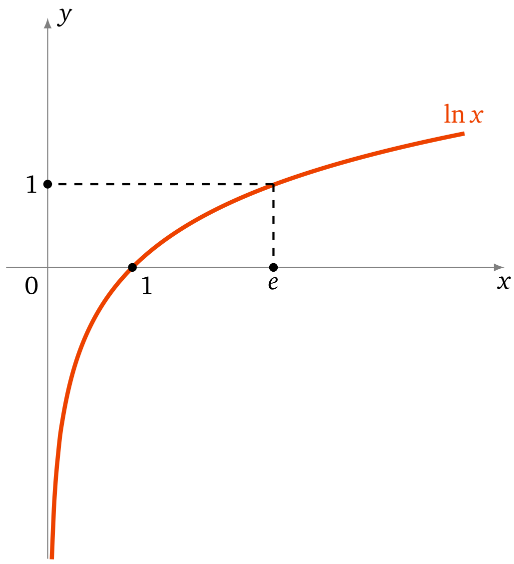
1.1. Logarithme

Proposition 1.
Il existe une unique fonction, notée $\ln :]0,+\infty[\to \mathbb{R}$ *telle que :*

$$\ln'(x) = \frac{1}{x} \quad (\textit{pour tout } x > 0) \qquad \textit{et} \qquad \ln(1) = 0.$$

De plus cette fonction vérifie (pour tout $a, b > 0$*) :*

1. $\ln(a \times b) = \ln a + \ln b$,
2. $\ln(\frac{1}{a}) = -\ln a$,
3. $\ln(a^n) = n \ln a$, *(pour tout* $n \in \mathbb{N}$*)*
4. $\ln$ *est une fonction continue, strictement croissante et définit une bijection de* $]0,+\infty[$ *sur* $\mathbb{R}$,
5. $\lim_{x\to 0} \frac{\ln(1+x)}{x} = 1$,
6. *la fonction* $\ln$ *est concave et* $\ln x \leqslant x - 1$ *(pour tout* $x > 0$*).*



Remarque.
$\ln x$ s'appelle le ***logarithme naturel*** ou aussi ***logarithme néperien***. Il est caractérisé par $\ln(e) = 1$. On définit le ***logarithme en base a*** par

$$\log_a(x) = \frac{\ln(x)}{\ln(a)}$$

De sorte que $\log_a(a) = 1$.
Pour $a = 10$ on obtient le ***logarithme décimal*** $\log_{10}$ qui vérifie $\log_{10}(10) = 1$ (et donc $\log_{10}(10^n) = n$). Dans la pratique on utilise l'équivalence :

$$\boxed{x = 10^y \iff y = \log_{10}(x)}$$

En informatique intervient aussi le logarithme en base 2 : $\log_2(2^n) = n$.

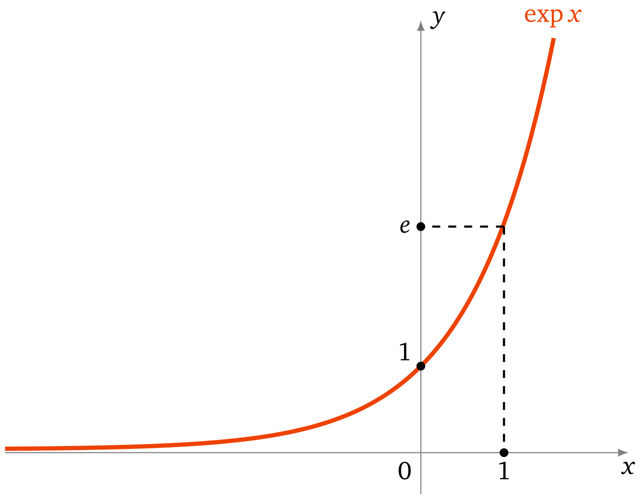
Démonstration. L'existence et l'unicité viennent de la théorie de l'intégrale : $\ln(x) = \int_1^x \frac{1}{t}\,dt$. Passons aux propriétés.

1. Posons $f(x) = \ln(xy) - \ln(x)$ où $y > 0$ est fixé. Alors $f'(x) = y\ln'(xy) - \ln'(x) = \frac{y}{xy} - \frac{1}{x} = 0$. Donc $x \mapsto f(x)$ a une dérivée nulle, donc est constante et vaut $f(1) = \ln(y) - \ln(1) = \ln(y)$. Donc $\ln(xy) - \ln(x) = \ln(y)$.
2. D'une part $\ln(a \times \frac{1}{a}) = \ln a + \ln\frac{1}{a}$, mais d'autre part $\ln(a \times \frac{1}{a}) = \ln(1) = 0$. Donc $\ln a + \ln\frac{1}{a} = 0$.
3. Similaire ou récurrence.
4. ln est dérivable donc continue, $\ln'(x) = \frac{1}{x} > 0$ donc la fonction est strictement croissante. Comme $\ln(2) > \ln(1) = 0$ alors $\ln(2^n) = n\ln(2) \to +\infty$ (lorsque $n \to +\infty$). Donc $\lim_{x\to+\infty} \ln x = +\infty$. De $\ln x = -\ln\frac{1}{x}$ on déduit $\lim_{x\to 0} \ln x = -\infty$. Par le théorème sur les fonctions continues et strictement croissantes, $\ln :]0, +\infty[\to \mathbb{R}$ est une bijection.
5. $\lim_{x\to 0} \frac{\ln(1+x)}{x}$ est la dérivée de ln au point $x_0 = 1$, donc cette limite existe et vaut $\ln'(1) = 1$.
6. $\ln'(x) = \frac{1}{x}$ est décroissante, donc la fonction ln est concave. Posons $f(x) = x - 1 - \ln x$; $f'(x) = 1 - \frac{1}{x}$. Par une étude de fonction f atteint son minimum en $x_0 = 1$. Donc $f(x) \geqslant f(1) = 0$. Donc $\ln x \leqslant x - 1$.

□

1.2. Exponentielle

Définition 1.
La bijection réciproque de $\ln :]0,+\infty[\to \mathbb{R}$ s'appelle la fonction *exponentielle*, notée $\exp : \mathbb{R} \to]0,+\infty[$.



Pour $x \in \mathbb{R}$ on note aussi e^x pour $\exp x$.

Proposition 2.
La fonction exponentielle vérifie les propriétés suivantes :

- $\boxed{\exp(\ln x) = x \text{ pour tout } x > 0}$ *et* $\boxed{\ln(\exp x) = x \text{ pour tout } x \in \mathbb{R}}$
- $\exp(a+b) = \exp(a) \times \exp(b)$
- $\exp(nx) = (\exp x)^n$
- $\exp : \mathbb{R} \to]0,+\infty[$ *est une fonction continue, strictement croissante vérifiant* $\lim_{x\to-\infty} \exp x = 0$ *et* $\lim_{x\to+\infty} \exp = +\infty$.
- *La fonction exponentielle est dérivable et* $\exp' x = \exp x$*, pour tout* $x \in \mathbb{R}$*. Elle est convexe et* $\exp x \geqslant 1+x$.

Remarque.
La fonction exponentielle est l'unique fonction qui vérifie $\exp'(x) = \exp(x)$ (pour tout $x \in \mathbb{R}$) et $\exp(1) = e$. Où $e \simeq 2,718\ldots$ est le nombre qui vérifie $\ln e = 1$.

Démonstration. Ce sont les propriétés du logarithme retranscrites pour sa bijection réciproque.
Par exemple pour la dérivée : on part de l'égalité $\ln(\exp x) = x$ que l'on dérive. Cela donne $\exp'(x) \times \ln'(\exp x) = 1$ donc $\exp'(x) \times \frac{1}{\exp x} = 1$ et ainsi $\exp'(x) = \exp x$. □

1.3. Puissance et comparaison

Par définition, pour $a > 0$ et $b \in \mathbb{R}$,

$$\boxed{a^b = \exp\big(b \ln a\big)}$$

Remarque.

- $\sqrt{a} = a^{\frac{1}{2}} = \exp\big(\frac{1}{2}\ln a\big)$
- $\sqrt[n]{a} = a^{\frac{1}{n}} = \exp\big(\frac{1}{n}\ln a\big)$ (la *racine n-ème* de a)
- On note aussi $\exp x$ par e^x ce qui se justifie par le calcul : $e^x = \exp\big(x \ln e\big) = \exp(x)$.
- Les fonctions $x \mapsto a^x$ s'appellent aussi des fonctions exponentielles et se ramènent systématiquement à la fonction exponentielle classique par l'égalité $a^x = \exp(x \ln a)$. Il ne faut surtout pas les confondre avec les fonctions puissances $x \mapsto x^a$.

Proposition 3.

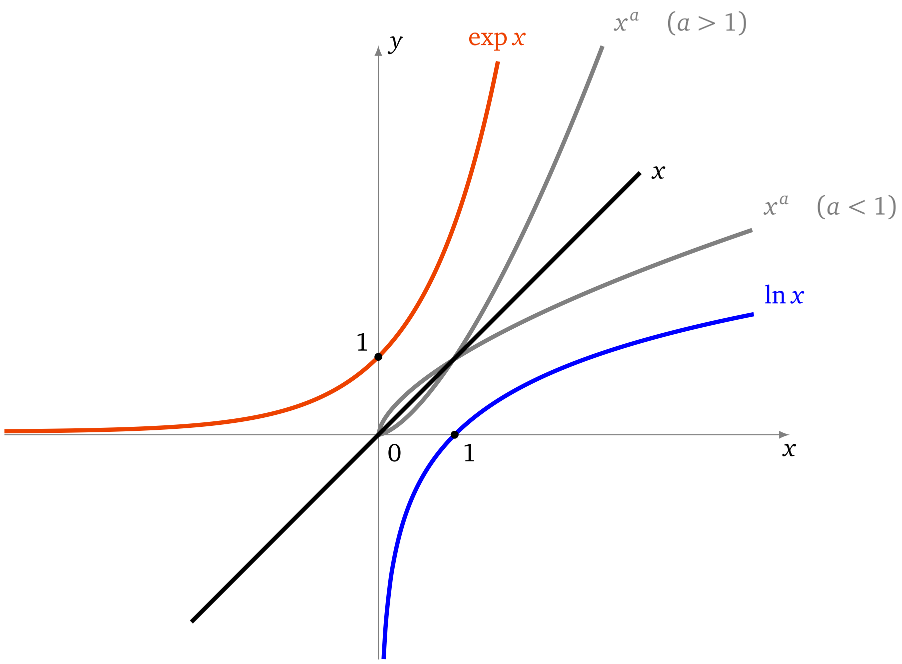
Soit $x, y > 0$ et $a, b \in \mathbb{R}$.

- $\boxed{x^{a+b} = x^a x^b}$
- $\boxed{x^{-a} = \frac{1}{x^a}}$
- $\boxed{(xy)^a = x^a y^a}$
- $\boxed{(x^a)^b = x^{ab}}$
- $\boxed{\ln(x^a) = a \ln x}$

Comparons les fonctions $\ln x$, $\exp x$ avec x :

Proposition 4.

$$\lim_{x\to+\infty} \frac{\ln x}{x} = 0 \qquad \textit{et} \qquad \lim_{x\to+\infty} \frac{\exp x}{x} = +\infty.$$



Démonstration.

1. On a vu $\ln x \leqslant x - 1$ (pour tout $x > 0$). Donc $\ln x \leqslant x$ donc $\frac{\ln\sqrt{x}}{\sqrt{x}} \leqslant 1$. Cela donne

$$0 \leqslant \frac{\ln x}{x} = \frac{\ln\left(\sqrt{x}^2\right)}{x} = 2\frac{\ln\sqrt{x}}{x} = 2\frac{\ln\sqrt{x}}{\sqrt{x}}\frac{1}{\sqrt{x}} \leqslant \frac{2}{\sqrt{x}}$$

   Cette double inégalité entraîne $\lim_{x\to+\infty} \frac{\ln x}{x} = 0$.

2. On a vu $\exp x \geqslant 1 + x$ (pour tout $x \in \mathbb{R}$). Donc $\exp x \to +\infty$ (lorsque $x \to +\infty$).

$$\frac{x}{\exp x} = \frac{\ln(\exp x)}{\exp x} = \frac{\ln u}{u}$$

   Lorsque $x \to +\infty$ alors $u = \exp x \to +\infty$ et donc par le premier point $\frac{\ln u}{u} \to 0$. Donc $\frac{x}{\exp x} \to 0$ et reste positive, ainsi $\lim_{x\to+\infty} \frac{\exp x}{x} = +\infty$.

$\square$

Mini-exercices.

1. Montrer que $\ln(1+e^x) = x + \ln(1+e^{-x})$, pour tout $x \in \mathbb{R}$.
2. Étudier la fonction $f(x) = \ln(x^2+1) - \ln(x) - 1$. Tracer son graphe. Résoudre l'équation $(f(x) = 0)$. Idem avec $g(x) = \frac{1+\ln x}{x}$. Idem avec $h(x) = x^x$.
3. Expliquer comment $\log_{10}$ permet de calculer le nombre de chiffres d'un entier n.
4. Montrer $\ln(1+x) \geqslant x - \frac{x^2}{2}$ pour $x \geqslant 0$ (faire une étude de fonction). Idem avec $e^x \geqslant 1 + x + \frac{x^2}{2}$ pour tout $x \geqslant 0$.
5. Calculer la limite de la suite définie par $u_n = \left(1+\frac{1}{n}\right)^n$ lorsque $n \to +\infty$. Idem avec $v_n = \left(\frac{1}{n}\right)^n$ et $w_n = n^{\frac{1}{n}}$.

2. Fonctions circulaires inverses

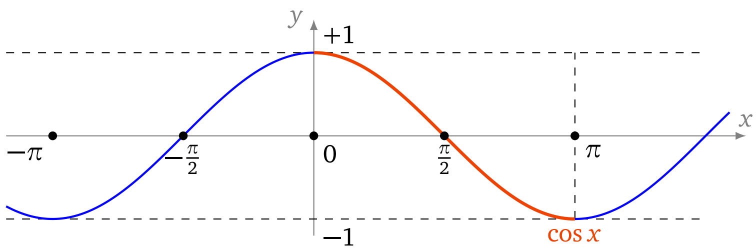
2.1. Arccosinus

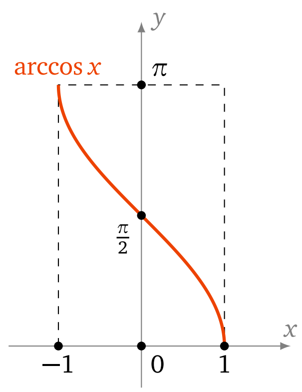
Considérons la fonction cosinus $\cos : \mathbb{R} \to [-1,1]$, $x \mapsto \cos x$. Pour obtenir une bijection à partir de cette fonction, il faut considérer la restriction de cosinus à l'intervalle $[0,\pi]$. Sur cet intervalle la fonction cosinus est continue et strictement décroissante, donc la restriction

$$\cos_| : [0,\pi] \to [-1,1]$$

est une bijection. Sa bijection réciproque est la fonction ***arccosinus*** :

$$\arccos : [-1,1] \to [0,\pi]$$





On a donc, par définition de la bijection réciproque :

$$\boxed{\begin{array}{ll} \cos\big(\arccos(x)\big) = x & \forall x \in [-1,1] \\ \arccos\big(\cos(x)\big) = x & \forall x \in [0,\pi] \end{array}}$$

Autrement dit :

$$\boxed{\text{Si} \quad x \in [0, \pi] \qquad \cos(x) = y \iff x = \arccos y}$$

Terminons avec la dérivée de arccos :

$$\boxed{\arccos'(x) = \frac{-1}{\sqrt{1-x^2}} \qquad \forall x \in]-1, 1[}$$

Démonstration. On démarre de l'égalité $\cos(\arccos x) = x$ que l'on dérive :

$$\begin{aligned}
& \cos(\arccos x) = x \\
\implies & -\arccos'(x) \times \sin(\arccos x) = 1 \\
\implies & \arccos'(x) = \frac{-1}{\sin(\arccos x)} \\
\implies & \arccos'(x) = \frac{-1}{\sqrt{1-\cos^2(\arccos x)}} \qquad (*) \\
\implies & \arccos'(x) = \frac{-1}{\sqrt{1-x^2}}
\end{aligned}$$

Le point crucial $(*)$ se justifie ainsi : on démarre de l'égalité $\cos^2 y + \sin^2 y = 1$, en substituant $y = \arccos x$ on obtient $\cos^2(\arccos x) + \sin^2(\arccos x) = 1$ donc $x^2 + \sin^2(\arccos x) = 1$. On en déduit : $\sin(\arccos x) = +\sqrt{1-x^2}$ (avec le signe $+$ car $\arccos x \in [0, \pi]$, et donc on a $\sin(\arccos x) \geqslant 0$). □

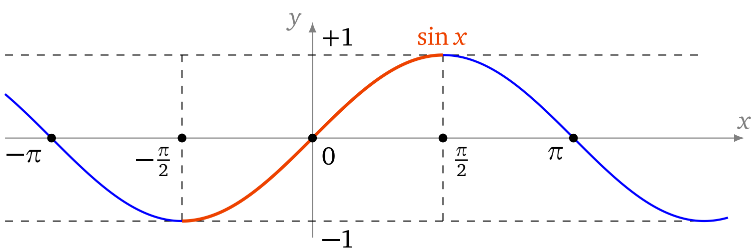
2.2. Arcsinus

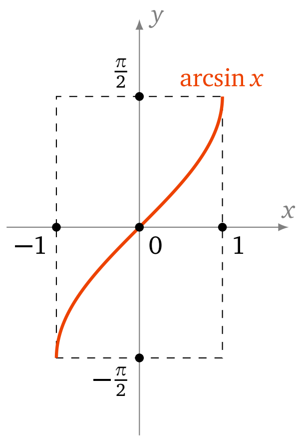
La restriction

$$\sin_| : [-\tfrac{\pi}{2}, +\tfrac{\pi}{2}] \to [-1, 1]$$

est une bijection. Sa bijection réciproque est la fonction ***arcsinus*** :

$$\arcsin : [-1, 1] \to [-\tfrac{\pi}{2}, +\tfrac{\pi}{2}]$$





$$\boxed{\begin{array}{ll} \sin\big(\arcsin(x)\big) = x & \forall x \in [-1, 1] \\ \arcsin\big(\sin(x)\big) = x & \forall x \in [-\frac{\pi}{2}, +\frac{\pi}{2}] \end{array}}$$

$$\boxed{\text{Si} \quad x \in [-\tfrac{\pi}{2}, +\tfrac{\pi}{2}] \qquad \sin(x) = y \iff x = \arcsin y}$$

$$\boxed{\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} \qquad \forall x \in]-1, 1[}$$

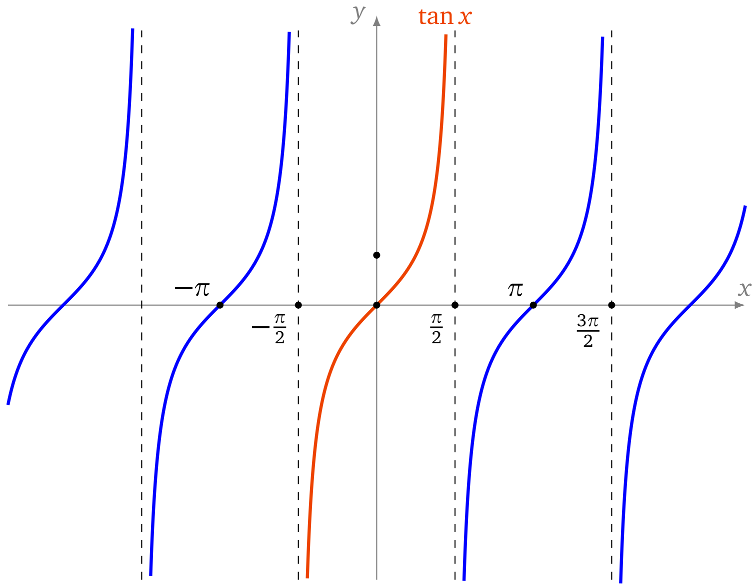
2.3. Arctangente

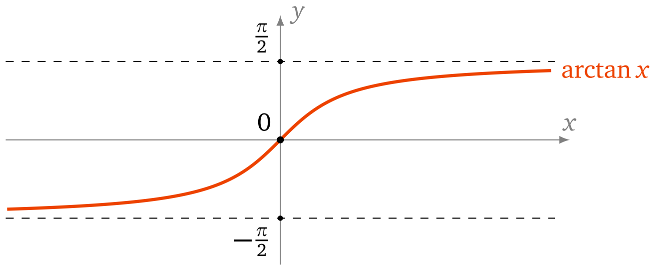
La restriction

$$\tan_| : \left]-\frac{\pi}{2}, +\frac{\pi}{2}\right[\to \mathbb{R}$$

est une bijection. Sa bijection réciproque est la fonction ***arctangente*** :

$$\arctan : \mathbb{R} \to \left]-\frac{\pi}{2}, +\frac{\pi}{2}\right[$$





$$\boxed{\begin{aligned} \tan\big(\arctan(x)\big) &= x \quad \forall x \in \mathbb{R} \\ \arctan\big(\tan(x)\big) &= x \quad \forall x \in \left]-\frac{\pi}{2}, +\frac{\pi}{2}\right[\end{aligned}}$$

$$\boxed{\text{Si} \quad x \in \left]-\frac{\pi}{2}, +\frac{\pi}{2}\right[\qquad \tan(x) = y \iff x = \arctan y}$$

$$\boxed{\arctan'(x) = \frac{1}{1+x^2} \qquad \forall x \in \mathbb{R}}$$

Mini-exercices.

1. Calculer les valeurs de arccos et arcsin en 0, 1, $\frac{1}{2}$, $\frac{\sqrt{2}}{2}$, $\frac{\sqrt{3}}{2}$. Idem pour arctan en 0, 1, $\sqrt{3}$ et $\frac{1}{\sqrt{3}}$.
2. Calculer $\arccos(\cos\frac{7\pi}{3})$. Idem avec $\arcsin(\sin\frac{7\pi}{3})$ et $\arctan(\tan\frac{7\pi}{3})$ (attention aux intervalles !)
3. Calculer $\cos(\arctan x)$, $\cos(\arcsin x)$, $\tan(\arcsin x)$.
4. Calculer la dérivée de $f(x) = \arctan\left(\frac{x}{\sqrt{1-x^2}}\right)$. En déduire que $f(x) = \arcsin x$, pour tout $x \in]-1,1[$.
5. Montrer que $\arccos x + \arcsin x = \frac{\pi}{2}$, pour tout $x \in [-1,1]$.

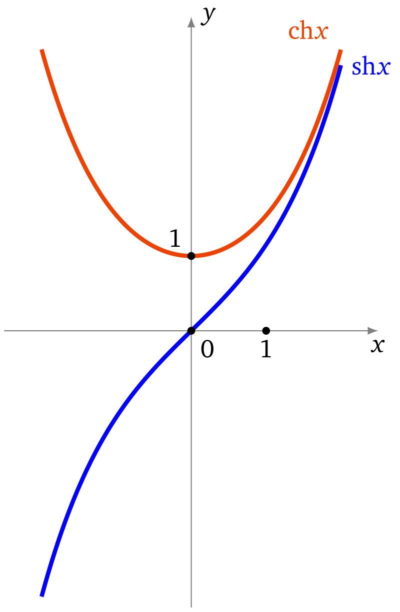
3. Fonctions hyperboliques et hyperboliques inverses

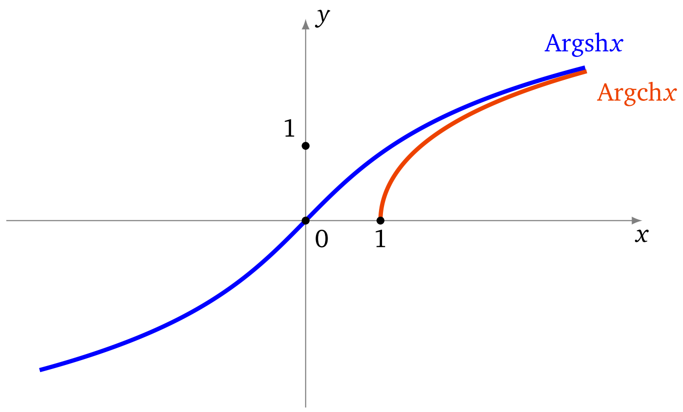
3.1. Cosinus hyperbolique et son inverse

Pour $x \in \mathbb{R}$, le *cosinus hyperbolique* est :

$$\boxed{\operatorname{ch} x = \frac{e^x + e^{-x}}{2}}$$

La restriction $\operatorname{ch}_| : [0, +\infty[\to [1, +\infty[$ est une bijection. Sa bijection réciproque est $\operatorname{Argch} : [1, +\infty[\to [0, +\infty[$.





3.2. Sinus hyperbolique et son inverse

Pour $x \in \mathbb{R}$, le *sinus hyperbolique* est :

$$\boxed{\operatorname{sh} x = \frac{e^x - e^{-x}}{2}}$$

$\operatorname{sh} : \mathbb{R} \to \mathbb{R}$ est une fonction continue, dérivable, strictement croissante vérifiant $\lim_{x \to -\infty} \operatorname{sh} x = -\infty$ et $\lim_{x \to +\infty} \operatorname{sh} x = +\infty$, c'est donc une bijection. Sa bijection réciproque est $\operatorname{Argsh} : \mathbb{R} \to \mathbb{R}$.

Proposition 5.

- $\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1$
- $\operatorname{ch}' x = \operatorname{sh} x, \ \operatorname{sh}' x = \operatorname{ch} x$
- $\operatorname{Argsh} : \mathbb{R} \to \mathbb{R}$ *est strictement croissante et continue.*
- Argsh *est dérivable et* $\operatorname{Argsh}' x = \frac{1}{\sqrt{x^2+1}}$.
- $\operatorname{Argsh} x = \ln\left(x + \sqrt{x^2+1}\right)$

Démonstration.

- $\operatorname{ch}^2 x - \operatorname{sh}^2 x = \frac{1}{4}\left[(e^x + e^{-x})^2 - (e^x - e^{-x})^2\right] = \frac{1}{4}\left[(e^{2x} + 2 + e^{-2x}) - (e^{2x} - 2 + e^{-2x})\right] = 1$.
- $\frac{d}{dx}(\operatorname{ch} x) = \frac{d}{dx}\frac{e^x + e^{-x}}{2} = \frac{e^x - e^{-x}}{2} = \operatorname{sh} x$. Idem pour la dérivée de $\operatorname{sh} x$.
- Car c'est la réciproque de sh.

- Comme la fonction $x \mapsto \operatorname{sh}' x$ ne s'annule pas sur $\mathbb{R}$ alors la fonction Argsh est dérivable sur $\mathbb{R}$. On calcule la dérivée par dérivation de l'égalité $\operatorname{sh}(\operatorname{Argsh} x) = x$:
$$\operatorname{Argsh}' x = \frac{1}{\operatorname{ch}(\operatorname{Argsh} x)} = \frac{1}{\sqrt{\operatorname{sh}^2(\operatorname{Argsh} x) + 1}} = \frac{1}{\sqrt{x^2+1}}$$
- Notons $f(x) = \ln\left(x + \sqrt{x^2+1}\right)$ alors
$$f'(x) = \frac{1 + \frac{x}{\sqrt{x^2+1}}}{x + \sqrt{x^2+1}} = \frac{1}{\sqrt{x^2+1}} = \operatorname{Argsh}' x$$
Comme de plus $f(0) = \ln(1) = 0$ et $\operatorname{Argsh} 0 = 0$ (car $\operatorname{sh} 0 = 0$), on en déduit que pour tout $x \in \mathbb{R}$, $f(x) = \operatorname{Argsh} x$.

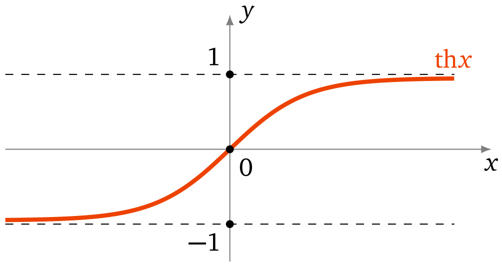
□

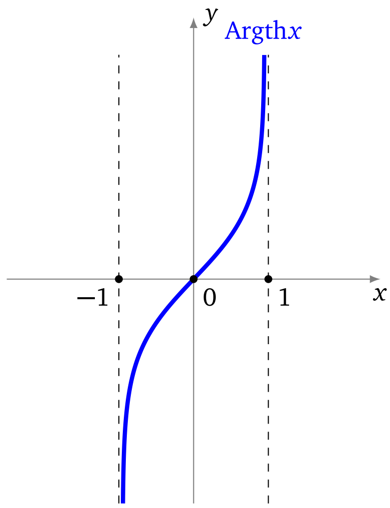
3.3. Tangente hyperbolique et son inverse

Par définition la ***tangente hyperbolique*** est :

$$\boxed{\operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x}}$$

La fonction $\operatorname{th} : \mathbb{R} \to]-1,1[$ est une bijection, on note $\operatorname{Argth} :]-1,1[\to \mathbb{R}$ sa bijection réciproque.





3.4. Trigonométrie hyperbolique

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1$$

$$\operatorname{ch}(a+b) = \operatorname{ch} a \cdot \operatorname{ch} b + \operatorname{sh} a \cdot \operatorname{sh} b$$
$$\operatorname{ch}(2a) = \operatorname{ch}^2 a + \operatorname{sh}^2 a = 2\operatorname{ch}^2 a - 1 = 1 + 2\operatorname{sh}^2 a$$

$$\operatorname{sh}(a+b) = \operatorname{sh} a \cdot \operatorname{ch} b + \operatorname{sh} b \cdot \operatorname{ch} a$$
$$\operatorname{sh}(2a) = 2\operatorname{sh} a \cdot \operatorname{ch} a$$

$$\operatorname{th}(a+b) = \frac{\operatorname{th} a + \operatorname{th} b}{1 + \operatorname{th} a \cdot \operatorname{th} b}$$

$$\begin{aligned}
\operatorname{ch}' x &= \operatorname{sh} x \\
\operatorname{sh}' x &= \operatorname{ch} x \\
\operatorname{th}' x &= 1 - \operatorname{th}^2 x = \frac{1}{\operatorname{ch}^2 x}
\end{aligned}$$

$$\begin{aligned}
\operatorname{Argch}' x &= \frac{1}{\sqrt{x^2-1}} \quad (x > 1) \\
\operatorname{Argsh}' x &= \frac{1}{\sqrt{x^2+1}} \\
\operatorname{Argth}' x &= \frac{1}{1-x^2} \quad (|x| < 1)
\end{aligned}$$

$$\begin{aligned}
\operatorname{Argch} x &= \ln\left(x + \sqrt{x^2-1}\right) \quad (x \geqslant 1) \\
\operatorname{Argsh} x &= \ln\left(x + \sqrt{x^2+1}\right) \quad (x \in \mathbb{R}) \\
\operatorname{Argth} x &= \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) \quad (-1 < x < 1)
\end{aligned}$$

Mini-exercices.

1. Dessiner les courbes paramétrées $t \mapsto (\cos t, \sin t)$ et $t \mapsto (\operatorname{ch} t, \operatorname{sh} t)$. Pourquoi cos et sin s'appellent des fonctions trigonométriques *circulaires* alors que ch et sh sont des fonctions trigonométriques *hyperboliques* ?
2. Prouver par le calcul la formule $\operatorname{ch}(a+b) = \ldots$ En utilisant que $\cos x = \frac{e^{ix}+e^{-ix}}{2}$ retrouver la formule pour $\cos(a+b)$.
3. Résoudre l'équation $\operatorname{sh} x = 3$.
4. Montrer que $\frac{\operatorname{sh}(2x)}{1+\operatorname{ch}(2x)} = \operatorname{th} x$.
5. Calculer les dérivées des fonctions définies par : $\operatorname{th}(1+x^2)$, $\ln(\operatorname{ch} x)$, $\operatorname{Argch}(\exp x)$, $\operatorname{Argth}(\cos x)$.

Auteurs du chapitre Arnaud Bodin, Niels Borne, Laura Desideri

Chapitre 3

Les suites

Objectif

L'étude des suites numériques a pour objet la compréhension de l'évolution de séquences de nombres (réels, complexes ...). Ceci permet de modéliser de nombreux phénomènes de la vie quotidienne. Supposons par exemple que l'on place une somme S à un taux annuel de 10%. Si S_n représente la somme que l'on obtiendra après n années, on a

$$S_0 = S \quad S_1 = S \times 1,1 \quad \dots \quad S_n = S \times (1,1)^n \ .$$

Au bout de $n = 10$ ans, on possédera donc $S_{10} = S \times (1,1)^{10} \approx S \times 2,59$: la somme de départ avec les intérêts cumulés.

1. Définitions

1.1. Définition d'une suite

Définition 1.

- Une *suite* est une application $u : \mathbb{N} \to \mathbb{R}$.
- Pour $n \in \mathbb{N}$, on note $u(n)$ par u_n et on l'appelle n-ème *terme* ou *terme général* de la suite.

La suite est notée u, ou plus souvent $(u_n)_{n\in\mathbb{N}}$ ou simplement (u_n). Il arrive fréquemment que l'on considère des suites définies à partir d'un certain entier naturel n_0 plus grand que 0, on note alors $(u_n)_{n\geqslant n_0}$.

Exemple 1.

- $(\sqrt{n})_{n\geqslant 0}$ est la suite de termes : 0, 1, $\sqrt{2}$, $\sqrt{3}$,...
- $((-1)^n)_{n\geqslant 0}$ est la suite qui alterne $+1, -1, +1, -1, \dots$
- La suite $(S_n)_{n\geqslant 0}$ de l'introduction définie par $S_n = S \times (1,1)^n$,
- $(F_n)_{n\geqslant 0}$ définie par $F_0 = 1$, $F_1 = 1$ et la relation $F_{n+2} = F_{n+1} + F_n$ pour $n \in \mathbb{N}$ (suite de Fibonacci). Les premiers termes sont 1, 1, 2, 3, 5, 8, 13, ...Chaque terme est la somme des deux précédents.
- $\left(\frac{1}{n^2}\right)_{n\geqslant 1}$. Les premiers termes sont 1, $\frac{1}{4}$, $\frac{1}{9}$, $\frac{1}{16}$, ...

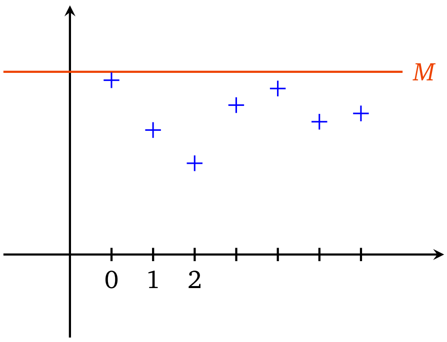
1.2. Suite majorée, minorée, bornée

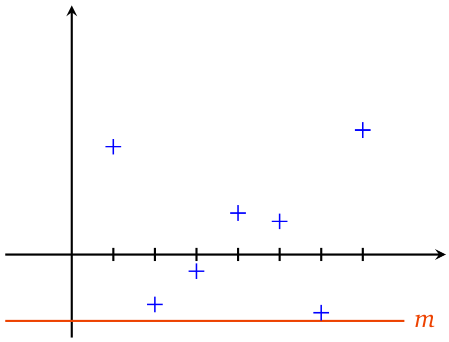
Définition 2.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite.

- $(u_n)_{n\in\mathbb{N}}$ est *majorée* si $\quad \exists M \in \mathbb{R} \quad \forall n \in \mathbb{N} \quad u_n \leqslant M$.
- $(u_n)_{n\in\mathbb{N}}$ est *minorée* si $\quad \exists m \in \mathbb{R} \quad \forall n \in \mathbb{N} \quad u_n \geqslant m$.

- $(u_n)_{n\in\mathbb{N}}$ est ***bornée*** si elle est majorée et minorée, ce qui revient à dire :
$$\exists M \in \mathbb{R} \quad \forall n \in \mathbb{N} \quad |u_n| \leqslant M.$$

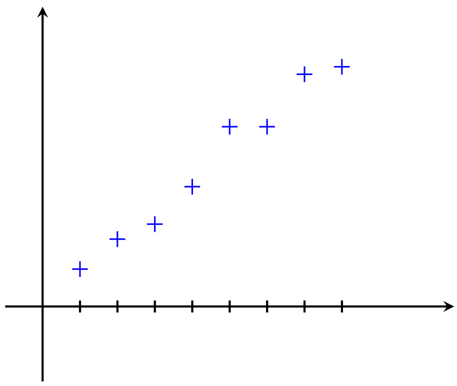




1.3. Suite croissante, décroissante

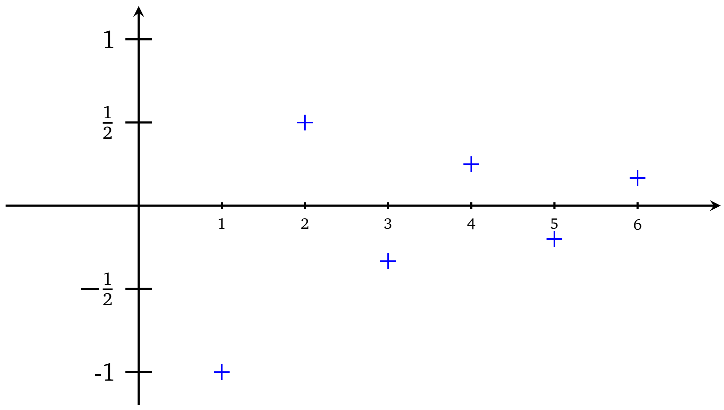
Définition 3.
Soit $(u_n)_{n\in\mathbb{N}}$ une suite.
- $(u_n)_{n\in\mathbb{N}}$ est ***croissante*** si $\quad \forall n \in \mathbb{N} \quad u_{n+1} \geqslant u_n$.
- $(u_n)_{n\in\mathbb{N}}$ est ***strictement croissante*** si $\quad \forall n \in \mathbb{N} \quad u_{n+1} > u_n$.
- $(u_n)_{n\in\mathbb{N}}$ est ***décroissante*** si $\quad \forall n \in \mathbb{N} \quad u_{n+1} \leqslant u_n$.
- $(u_n)_{n\in\mathbb{N}}$ est ***strictement décroissante*** si $\quad \forall n \in \mathbb{N} \quad u_{n+1} < u_n$.
- $(u_n)_{n\in\mathbb{N}}$ est ***monotone*** si elle est croissante ou décroissante.
- $(u_n)_{n\in\mathbb{N}}$ est ***strictement monotone*** si elle est strictement croissante ou strictement décroissante.

Voici un exemple d'une suite croissante (mais pas strictement croissante) :



Remarque.
- $(u_n)_{n\in\mathbb{N}}$ est croissante si et seulement si $\forall n \in \mathbb{N} \quad u_{n+1} - u_n \geqslant 0$.
- Si $(u_n)_{n\in\mathbb{N}}$ est une suite à termes strictement positifs, elle est croissante si et seulement si $\forall n \in \mathbb{N} \quad \frac{u_{n+1}}{u_n} \geqslant 1$.

Exemple 2.
- La suite $(S_n)_{n\geqslant 0}$ de l'introduction est strictement croissante car $S_{n+1}/S_n = 1,1 > 1$.
- La suite $(u_n)_{n\geqslant 1}$ définie par $u_n = (-1)^n/n$ pour $n \geqslant 1$, n'est ni croissante ni décroissante. Elle est majorée par $1/2$ (borne atteinte en $n = 2$), minorée par -1 (borne atteinte en $n = 1$).



- La suite $\left(\frac{1}{n}\right)_{n\geqslant 1}$ est une suite strictement décroissante. Elle est majorée par 1 (borne atteinte pour $n = 1$), elle est minorée par 0 mais cette valeur n'est jamais atteinte.

Mini-exercices.

1. La suite $\left(\frac{n}{n+1}\right)_{n\in\mathbb{N}}$ est-elle monotone ? Est-elle bornée ?
2. La suite $\left(\frac{n\sin(n!)}{1+n^2}\right)_{n\in\mathbb{N}}$ est-elle bornée ?
3. Réécrire les phrases suivantes en une phrase mathématique. Écrire ensuite la négation mathématique de chacune des phrases. (a) La suite $(u_n)_{n\in\mathbb{N}}$ est majorée par 7. (b) La suite $(u_n)_{n\in\mathbb{N}}$ est constante. (c) La suite $(u_n)_{n\in\mathbb{N}}$ est strictement positive à partir d'un certain rang. (d) $(u_n)_{n\in\mathbb{N}}$ n'est pas strictement croissante.
4. Est-il vrai qu'une suite croissante est minorée ? Majorée ?
5. Soit $x > 0$ un réel. Montrer que la suite $\left(\frac{x^n}{n!}\right)_{n\in\mathbb{N}}$ est décroissante à partir d'un certain rang.

2. Limites

2.1. Introduction

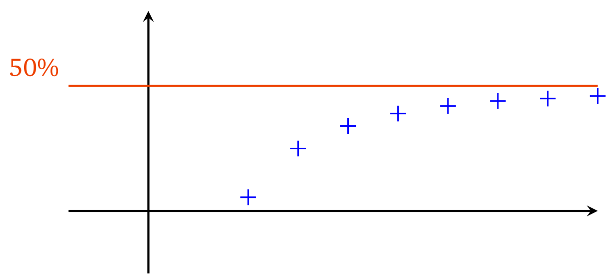
Pour un trajet au prix normal de 20 euros on achète une carte d'abonnement de train à 50 euros et on obtient chaque billet à 10 euros. La publicité affirme « 50% de réduction ». Qu'en pensez-vous ?
Pour modéliser la situation en termes de suites, on pose pour un entier $n \geqslant 1$:

$$u_n = 20n$$
$$v_n = 10n + 50$$

u_n est le prix payé au bout de n achats au tarif plein, et v_n celui au tarif réduit, y compris le prix de l'abonnement. La réduction est donc, en pourcentage :

$$1 - \frac{v_n}{u_n} = \frac{u_n - v_n}{u_n} = \frac{10n - 50}{20n} = 0,5 - \frac{5}{2n} \xrightarrow[n\to+\infty]{} 0,5$$

Il faut donc une infinité de trajets pour arriver à 50% de réduction !



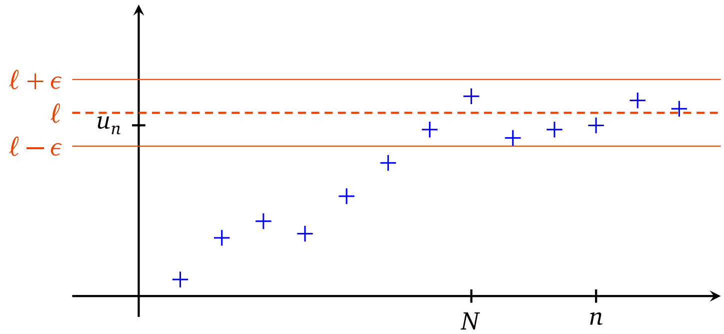
2.2. Limite finie, limite infinie

Soit $(u_n)_{n\in\mathbb{N}}$ une suite.

Définition 4.
La suite $(u_n)_{n\in\mathbb{N}}$ a pour *limite* $\ell \in \mathbb{R}$ si : pour tout $\epsilon > 0$, il existe un entier naturel N tel que si $n \geqslant N$ alors $|u_n - \ell| \leqslant \epsilon$:

$$\boxed{\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad (n \geqslant N \implies |u_n - \ell| \leqslant \epsilon)}$$

On dit aussi que la suite $(u_n)_{n\in\mathbb{N}}$ *tend vers* ℓ. Autrement dit : u_n est proche d'aussi près que l'on veut de ℓ, à partir d'un certain rang.



Définition 5.

1. La suite $(u_n)_{n\in\mathbb{N}}$ *tend vers* $+\infty$ si :
$$\forall A > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad (n \geqslant N \implies u_n \geqslant A)$$
2. La suite $(u_n)_{n\in\mathbb{N}}$ *tend vers* $-\infty$ si :
$$\forall A > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad (n \geqslant N \implies u_n \leqslant -A)$$

Remarque.

1. On note $\lim_{n\to+\infty} u_n = \ell$ ou parfois $u_n \xrightarrow[n\to+\infty]{} \ell$, et de même pour une limite $\pm\infty$.
2. $\lim_{n\to+\infty} u_n = -\infty \iff \lim_{n\to+\infty} -u_n = +\infty$.
3. On raccourcit souvent la phrase logique en :
$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \qquad (n \geqslant N \implies |u_n - \ell| \leqslant \epsilon).$$
Noter que N dépend de ϵ et qu'on ne peut pas échanger l'ordre du « pour tout » et du « il existe ».
4. L'inégalité $|u_n - \ell| \leqslant \epsilon$ signifie $\ell - \epsilon \leqslant u_n \leqslant \ell + \epsilon$. On aurait aussi pu définir la limite par la phrase : $\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \quad (n \geqslant N \implies |u_n - \ell| < \epsilon)$, où l'on a remplacé la dernière inégalité large par une inégalité stricte.

Définition 6.
Une suite $(u_n)_{n\in\mathbb{N}}$ est *convergente* si elle admet une limite ***finie***. Elle est *divergente* sinon (c'est-à-dire soit la suite tend vers $\pm\infty$, soit elle n'admet pas de limite).

On va pouvoir parler de ***la*** limite, si elle existe, car il y a unicité de la limite :

Proposition 1.
Si une suite est convergente, sa limite est unique.

Démonstration. On procède par l'absurde. Soit $(u_n)_{n\in\mathbb{N}}$ une suite convergente ayant deux limites $\ell \neq \ell'$. Choisissons $\epsilon > 0$ tel que $\epsilon < \frac{|\ell-\ell'|}{2}$.
Comme $\lim_{n\to+\infty} u_n = \ell$, il existe N_1 tel que $n \geqslant N_1$ implique $|u_n - \ell| < \epsilon$.
De même $\lim_{n\to+\infty} u_n = \ell'$, il existe N_2 tel que $n \geqslant N_2$ implique $|u_n - \ell'| < \epsilon$.
Notons $N = \max(N_1, N_2)$, on a alors pour ce N :

$$|u_N - \ell| < \epsilon \quad \text{et} \quad |u_N - \ell'| < \epsilon$$

Donc $|\ell - \ell'| = |\ell - u_N + u_N - \ell'| \leqslant |\ell - u_N| + |u_N - \ell'|$ d'après l'inégalité triangulaire. On en tire $|\ell - \ell'| \leqslant \epsilon + \epsilon = 2\epsilon < |\ell - \ell'|$. On vient d'aboutir à l'inégalité $|\ell - \ell'| < |\ell - \ell'|$ qui est impossible. Bilan : notre hypothèse de départ est fausse et donc $\ell = \ell'$. □

2.3. Propriétés des limites

Proposition 2.

1. $\lim_{n\to+\infty} u_n = \ell \iff \lim_{n\to+\infty}(u_n - \ell) = 0 \iff \lim_{n\to+\infty}|u_n - \ell| = 0$,
2. $\lim_{n\to+\infty} u_n = \ell \implies \lim_{n\to+\infty}|u_n| = |\ell|$.

Démonstration. Cela résulte directement de la définition. □

Proposition 3 (Opérations sur les limites).
Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites convergentes.

1. *Si $\lim_{n\to+\infty} u_n = \ell$, où $\ell \in \mathbb{R}$, alors pour $\lambda \in \mathbb{R}$ on a $\lim_{n\to+\infty} \lambda u_n = \lambda\ell$.*
2. *Si $\lim_{n\to+\infty} u_n = \ell$ et $\lim_{n\to+\infty} v_n = \ell'$, où $\ell, \ell' \in \mathbb{R}$, alors*
$$\lim_{n\to+\infty}(u_n + v_n) = \ell + \ell'$$
$$\lim_{n\to+\infty}(u_n \times v_n) = \ell \times \ell'$$
3. *Si $\lim_{n\to+\infty} u_n = \ell$ où $\ell \in \mathbb{R}^* = \mathbb{R}\setminus\{0\}$ alors $u_n \neq 0$ pour n assez grand et $\lim_{n\to+\infty} \frac{1}{u_n} = \frac{1}{\ell}$.*

Nous ferons la preuve dans la section suivante.
Nous utilisons continuellement ces propriétés, le plus souvent sans nous en rendre compte.

Exemple 3.
Si $u_n \to \ell$ avec $\ell \neq \pm 1$, alors

$$u_n(1 - 3u_n) - \frac{1}{u_n^2 - 1} \xrightarrow[n\to+\infty]{} \ell(1 - 3\ell) - \frac{1}{\ell^2 - 1}.$$

Proposition 4 (Opérations sur les limites infinies).
Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites telles que $\lim_{n\to+\infty} v_n = +\infty$.

1. $\lim_{n\to+\infty} \frac{1}{v_n} = 0$
2. *Si $(u_n)_{n\in\mathbb{N}}$ est minorée alors $\lim_{n\to+\infty}(u_n + v_n) = +\infty$.*
3. *Si $(u_n)_{n\in\mathbb{N}}$ est minorée par un nombre $\lambda > 0$ alors $\lim_{n\to+\infty}(u_n \times v_n) = +\infty$.*
4. *Si $\lim_{n\to+\infty} u_n = 0$ et $u_n > 0$ pour n assez grand alors $\lim_{n\to+\infty} \frac{1}{u_n} = +\infty$.*

Exemple 4.
La suite $(\sqrt{n})$ tend vers $+\infty$, donc la suite $(\frac{1}{\sqrt{n}})$ tend vers 0.

2.4. Des preuves !

Nous n'allons pas tout prouver mais seulement quelques résultats importants. Les autres se démontrent de manière tout à fait semblable.
Commençons par prouver un résultat assez facile (le premier point de la proposition 4) :

$$\text{« } \textit{Si} \quad \lim u_n = +\infty \quad \textit{alors} \quad \lim \frac{1}{u_n} = 0. \text{ »}$$

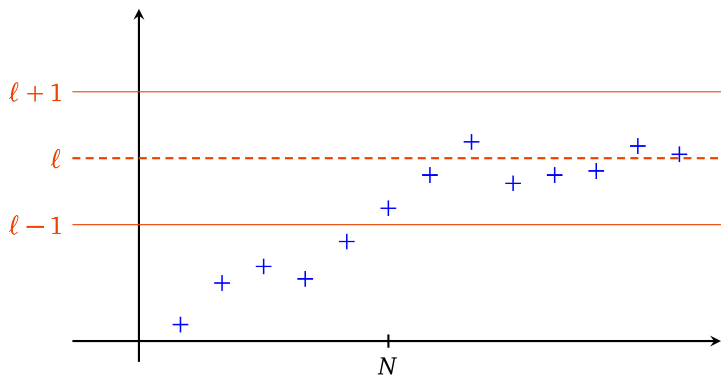
Démonstration. Fixons $\epsilon > 0$. Comme $\lim_{n\to+\infty} u_n = +\infty$, il existe un entier naturel N tel que $n \geqslant N$ implique $u_n \geqslant \frac{1}{\epsilon}$. On obtient alors $0 \leqslant \frac{1}{u_n} \leqslant \epsilon$ pour $n \geqslant N$. On a donc montré que $\lim_{n\to+\infty} \frac{1}{u_n} = 0$. □

Afin de prouver que la limite d'un produit est le produit des limites nous aurons besoin d'un peu de travail.

Proposition 5.
Toute suite convergente est bornée.

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$ une suite convergeant vers le réel ℓ. En appliquant la définition de limite (définition 4) avec $\epsilon = 1$, on obtient qu'il existe un entier naturel N tel que pour $n \geqslant N$ on ait $|u_n - \ell| \leqslant 1$, et donc pour $n \geqslant N$ on a

$$|u_n| = |\ell + (u_n - \ell)| \leqslant |\ell| + |u_n - \ell| \leqslant |\ell| + 1.$$



Donc si on pose

$$M = \max(|u_0|, |u_1|, \cdots, |u_{N-1}|, |\ell| + 1)$$

on a alors $\forall n \in \mathbb{N} \;\; |u_n| \leqslant M$. □

Proposition 6.
Si la suite $(u_n)_{n\in\mathbb{N}}$ *est bornée et* $\lim_{n\to+\infty} v_n = 0$ *alors* $\lim_{n\to+\infty}(u_n \times v_n) = 0$.

Exemple 5.
Si $(u_n)_{n\geqslant 1}$ est la suite donnée par $u_n = \cos(n)$ et $(v_n)_{n\geqslant 1}$ est celle donnée par $v_n = \frac{1}{\sqrt{n}}$, alors $\lim_{n\to+\infty}(u_n v_n) = 0$.

Démonstration. La suite $(u_n)_{n\in\mathbb{N}}$ est bornée, on peut donc trouver un réel $M > 0$ tel que pour tout entier naturel n on ait $|u_n| \leqslant M$. Fixons $\epsilon > 0$. On applique la définition de limite (définition 4) à la suite $(v_n)_{n\in\mathbb{N}}$ pour $\epsilon' = \frac{\epsilon}{M}$. Il existe donc un entier naturel N tel que $n \geqslant N$ implique $|v_n| \leqslant \epsilon'$. Mais alors pour $n \geqslant N$ on a :

$$|u_n v_n| = |u_n||v_n| \leqslant M \times \epsilon' = \epsilon.$$

On a bien montré que $\lim_{n\to+\infty}(u_n \times v_n) = 0$. □

Prouvons maintenant la formule concernant le produit de deux limites (voir proposition 3).

$$\text{« } \textit{Si} \quad \lim u_n = \ell \quad \textit{et} \quad \lim v_n = \ell' \quad \textit{alors} \quad \lim u_n v_n = \ell\ell'. \text{ »}$$

Démonstration de la formule concernant le produit de deux limites. Le principe est d'écrire :

$$u_n v_n - \ell\ell' = (u_n - \ell)v_n + \ell(v_n - \ell')$$

D'après la proposition 6, la suite de terme général $\ell(v_n - \ell')$ tend vers 0. Par la même proposition il en est de même de la suite de terme général $(u_n - \ell)v_n$, car la suite convergente $(v_n)_{n\in\mathbb{N}}$ est bornée. On conclut que $\lim_{n\to+\infty}(u_n v_n - \ell\ell') = 0$, ce qui équivaut à $\lim_{n\to+\infty} u_n v_n = \ell\ell'$. □

2.5. Formes indéterminées

Dans certaines situations, on ne peut rien dire à priori sur la limite, il faut faire une étude au cas par cas.

Exemple 6.

1. « $+\infty - \infty$ » Cela signifie que si $u_n \to +\infty$ et $v_n \to -\infty$ il faut faire faire l'étude en fonction de chaque suite pour déterminer $\lim(u_n + v_n)$ comme le prouve les exemples suivants.

$$\lim_{n\to+\infty} (e^n - \ln(n)) = +\infty$$
$$\lim_{n\to+\infty} (n - n^2) = -\infty$$
$$\lim_{n\to+\infty} \left(\left(n + \frac{1}{n}\right) - n\right) = 0$$

2. « $0 \times \infty$ »

$$\lim_{n\to+\infty} \frac{1}{\ln n} \times e^n = +\infty$$
$$\lim_{n\to+\infty} \frac{1}{n} \times \ln n = 0$$
$$\lim_{n\to+\infty} \frac{1}{n} \times (n+1) = 1$$

3. « $\frac{\infty}{\infty}$ », « $\frac{0}{0}$ », « 1^∞ », ...

2.6. Limite et inégalités

Proposition 7.

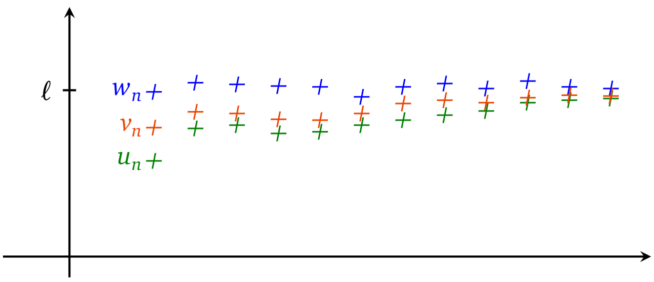
1. *Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites convergentes telles que : $\forall n \in \mathbb{N}$, $u_n \leqslant v_n$. Alors*

$$\lim_{n\to+\infty} u_n \leqslant \lim_{n\to+\infty} v_n$$

2. *Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites telles que $\lim_{n\to+\infty} u_n = +\infty$ et $\forall n \in \mathbb{N}$, $v_n \geqslant u_n$. Alors $\lim_{n\to+\infty} v_n = +\infty$.*
3. *Théorème des « gendarmes » : si $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ sont trois suites telles que*

$$\forall n \in \mathbb{N} \quad u_n \leqslant v_n \leqslant w_n$$

et $\lim_{n\to+\infty} u_n = \ell = \lim_{n\to+\infty} w_n$, alors la suite $(v_n)_{n\in\mathbb{N}}$ est convergente et $\lim_{n\to+\infty} v_n = \ell$.

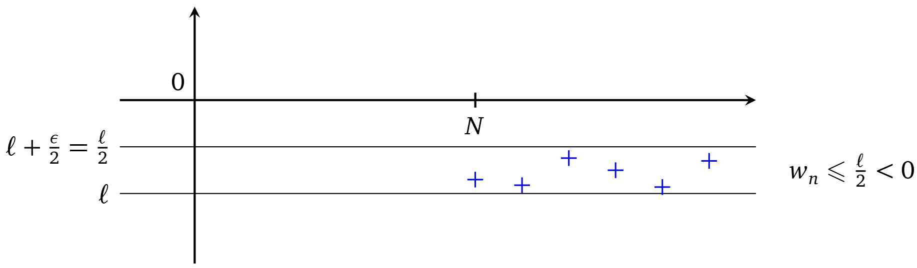


Remarque.

1. Soit $(u_n)_{n\in\mathbb{N}}$ une suite convergente telle que : $\forall n \in \mathbb{N}$, $u_n \geqslant 0$. Alors $\lim_{n\to+\infty} u_n \geqslant 0$.
2. Attention, si $(u_n)_{n\in\mathbb{N}}$ est une suite convergente telle que : $\forall n \in \mathbb{N}$, $u_n > 0$, on ne peut affirmer que la limite est strictement positive mais seulement que $\lim_{n\to+\infty} u_n \geqslant 0$. Par exemple la suite $(u_n)_{n\in\mathbb{N}}$ donnée par $u_n = \frac{1}{n+1}$ est à termes strictement positifs, mais converge vers zéro.

Démonstration de la proposition 7.

1. En posant $w_n = v_n - u_n$, on se ramène à montrer que si une suite $(w_n)_{n\in\mathbb{N}}$ vérifie $\forall n \in \mathbb{N}$, $w_n \geqslant 0$ et converge, alors $\lim_{n\to+\infty} w_n \geqslant 0$. On procède par l'absurde en supposant que $\ell = \lim_{n\to+\infty} w_n < 0$. En prenant $\epsilon = |\frac{\ell}{2}|$ dans la définition de limite (définition 4), on obtient qu'il existe un entier naturel N tel que $n \geqslant N$ implique $|w_n - \ell| < \epsilon = -\frac{\ell}{2}$. En particulier on a pour $n \geqslant N$ que $w_n < \ell - \frac{\ell}{2} = \frac{\ell}{2} < 0$, une contradiction.



2. Laissé en exercice.
3. En soustrayant la suite $(u_n)_{n\in\mathbb{N}}$, on se ramène à montrer l'énoncé suivant : si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont deux suites telles que : $\forall n \in \mathbb{N}$, $0 \leqslant u_n \leqslant v_n$ et $\lim_{n\to+\infty} v_n = 0$, alors (u_n) converge et $\lim_{n\to+\infty} u_n = 0$. Soit $\epsilon > 0$ et N un entier naturel tel que $n \geqslant N$ implique $|v_n| < \epsilon$. Comme $|u_n| = u_n \leqslant v_n = |v_n|$, on a donc : $n \geqslant N$ implique $|u_n| < \epsilon$. On a bien montré que $\lim_{n\to+\infty} u_n = 0$.

□

Exemple 7 (Exemple d'application du théorème des « gendarmes »).
Trouver la limite de la suite $(u_n)_{n\in\mathbb{N}}$ de terme général :

$$u_n = 2 + \frac{(-1)^n}{1+n+n^2}$$

Mini-exercices.

1. Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par $u_n = \frac{2n+1}{n+2}$. En utilisant la définition de la limite montrer que $\lim_{n\to+\infty} u_n = 2$. Trouver explicitement un rang à partir duquel $1,999 \leqslant u_n \leqslant 2,001$.
2. Déterminer la limite ℓ de la suite $(u_n)_{n\in\mathbb{N}^*}$ de terme général : $\frac{n+\cos n}{n-\sin n}$ et trouver un entier N tel que si $n \geqslant N$, on ait $|u_n - \ell| \leqslant 10^{-2}$.
3. La suite $(u_n)_{n\in\mathbb{N}}$ de terme général $(-1)^n e^n$ admet-elle une limite ? Et la suite de terme général $\frac{1}{u_n}$?
4. Déterminer la limite de la suite $(u_n)_{n\geqslant 1}$ de terme général $\sqrt{n+1} - \sqrt{n}$. Idem avec $v_n = \frac{\cos n}{\sin n + \ln n}$. Idem avec $w_n = \frac{n!}{n^n}$.

3. Exemples remarquables

3.1. Suite géométrique

Proposition 8 (Suite géométrique).
On fixe un réel a. Soit $(u_n)_{n\in\mathbb{N}}$ la suite de terme général : $u_n = a^n$.

1. *Si $a = 1$, on a pour tout $n \in \mathbb{N}$: $u_n = 1$.*
2. *Si $a > 1$, alors $\lim_{n\to+\infty} u_n = +\infty$.*
3. *Si $-1 < a < 1$, alors $\lim_{n\to+\infty} u_n = 0$.*
4. *Si $a \leqslant -1$, la suite $(u_n)_{n\in\mathbb{N}}$ diverge.*

Démonstration.

1. est évident.
2. Écrivons $a = 1 + b$ avec $b > 0$. Alors le binôme de Newton s'écrit $a^n = (1+b)^n = 1 + nb + \binom{n}{2}b^2 + \cdots + \binom{n}{k}b^k + \cdots + b^n$. Tous les termes sont positifs, donc pour tout entier naturel n on a : $a^n \geqslant 1 + nb$. Or $\lim_{n\to+\infty}(1+nb) = +\infty$ car $b > 0$. On en déduit que $\lim_{n\to+\infty} a^n = +\infty$.
3. Si $a = 0$, le résultat est clair. Sinon, on pose $b = |\frac{1}{a}|$. Alors $b > 1$ et d'après le point précédent $\lim_{n\to+\infty} b^n = +\infty$. Comme pour tout entier naturel n on a : $|a|^n = \frac{1}{b^n}$, on en déduit que $\lim_{n\to+\infty} |a|^n = 0$, et donc aussi $\lim_{n\to+\infty} a^n = 0$.
4. Supposons par l'absurde que la suite $(u_n)_{n\in\mathbb{N}}$ converge vers le réel ℓ. De $a^2 \geqslant 1$, on déduit que pour tout entier naturel n, on a $a^{2n} \geqslant 1$. En passant à la limite, il vient $\ell \geqslant 1$. Comme de plus pour tout entier naturel n on a $a^{2n+1} \leqslant a \leqslant -1$, il vient en passant de nouveau à la limite $\ell \leqslant -1$. Mais comme on a déjà $\ell \geqslant 1$, on obtient une contradiction, et donc (u_n) ne converge pas.

□

3.2. Série géométrique

Proposition 9 (Série géométrique).
Soit a un réel, $a \neq 1$. En notant $\sum_{k=0}^{n} a^k = 1 + a + a^2 + \cdots + a^n$, on a :

$$\boxed{\sum_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}}$$

Démonstration. En multipliant par $1-a$ on fait apparaître une somme télescopique (presque tous les termes s'annulent) :

$$(1-a)\big(1+a+a^2+\cdots+a^n\big) = \big(1+a+a^2+\cdots+a^n\big) - \big(a+a^2+\cdots+a^{n+1}\big) = 1-a^{n+1}.$$

□

Remarque.
Si $a \in]-1,1[$ et $(u_n)_{n\in\mathbb{N}}$ est la suite de terme général : $u_n = \sum_{k=0}^{n} a^k$, alors $\lim_{n\to+\infty} u_n = \frac{1}{1-a}$. De manière plus frappante, on peut écrire :

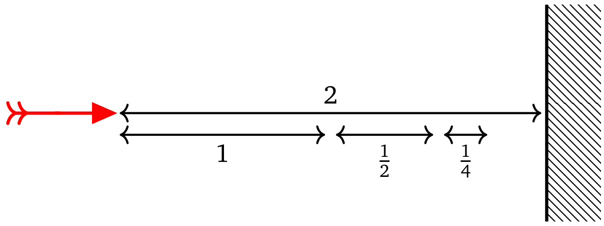
$$1 + a + a^2 + a^3 + \cdots = \frac{1}{1-a}$$

Enfin, ces formules sont aussi valables si $a \in \mathbb{C} \setminus \{1\}$. Si $a = 1$, alors $1 + a + a^2 + \cdots + a^n = n+1$.

Exemple 8.
L'exemple précédent avec $a = \frac{1}{2}$ donne

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2.$$

Cette formule était difficilement concevable avant l'avènement du calcul infinitésimal et a été popularisée sous le nom du ***paradoxe de Zénon***. On tire une flèche à 2 mètres d'une cible. Elle met un certain laps de temps pour parcourir la moitié de la distance, à savoir un mètre. Puis il lui faut encore du temps pour parcourir la moitié de la distance restante, et de nouveau un certain temps pour la moitié de la distance encore restante. On ajoute ainsi une infinité de durées non nulles, et Zénon en conclut que la flèche n'atteint jamais sa cible !
L'explication est bien donnée par l'égalité ci-dessus : la somme d'une infinité de termes peut bien être une valeur finie ! ! Par exemple si la flèche va à une vitesse de 1 m/s, alors elle parcoure la première moitié en 1 s, le moitié de la distance restante en $\frac{1}{2}$ s, etc. Elle parcoure bien toute la distance en $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2$ secondes !



3.3. Suites telles que $\left|\frac{u_{n+1}}{u_n}\right| < \ell < 1$

Théorème 1.
Soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels non nuls. On suppose qu'il existe un réel ℓ tel que pour tout entier naturel n (ou seulement à partir d'un certain rang) on ait :

$$\left|\frac{u_{n+1}}{u_n}\right| < \ell < 1.$$

Alors $\lim_{n\to+\infty} u_n = 0$.

Démonstration. On suppose que la propriété $\left|\frac{u_{n+1}}{u_n}\right| < \ell < 1$ est vraie pour tout entier naturel n (la preuve dans le cas où cette propriété n'est vraie qu'à partir d'un certain rang n'est pas très différente). On écrit

$$\frac{u_n}{u_0} = \frac{u_1}{u_0} \times \frac{u_2}{u_1} \times \frac{u_3}{u_2} \times \cdots \times \frac{u_n}{u_{n-1}}$$

ce dont on déduit

$$\left|\frac{u_n}{u_0}\right| < \ell \times \ell \times \ell \times \cdots \times \times \ell = \ell^n$$

et donc $|u_n| < |u_0|\ell^n$. Comme $\ell < 1$, on a $\lim_{n\to+\infty} \ell^n = 0$. On conclut que $\lim_{n\to+\infty} u_n = 0$. □

Corollaire 1.
Soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels non nuls.

$$\boxed{\textit{Si } \lim_{n\to+\infty} \frac{u_{n+1}}{u_n} = 0, \textit{ alors } \lim_{n\to+\infty} u_n = 0.}$$

Exemple 9.
Soit $a \in \mathbb{R}$. Alors $\lim_{n\to+\infty} \frac{a^n}{n!} = 0$.

Démonstration. Si $a = 0$, le résultat est évident. Supposons $a \neq 0$, et posons $u_n = \frac{a^n}{n!}$. Alors

$$\frac{u_{n+1}}{u_n} = \frac{a^{n+1}}{(n+1)!} \cdot \frac{n!}{a^n} = \frac{a}{n+1}.$$

Pour conclure, on peut ou bien directement utiliser le corollaire : comme $\lim \frac{u_{n+1}}{u_n} = 0$ (car a est fixe), on a $\lim u_n = 0$. Ou bien, comme $\frac{u_{n+1}}{u_n} = \frac{a}{n+1}$, on déduit par le théorème que pour $n \geqslant N > 2|a|$ on a :

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{|a|}{n+1} \leqslant \frac{|a|}{N+1} < \frac{|a|}{N} < \frac{1}{2} = \ell$$

et donc $\lim_{n\to+\infty} u_n = 0$. □

Remarque.

1. Avec les notations du théorème, si on a pour tout entier naturel n à partir d'un certain rang : $\left|\frac{u_{n+1}}{u_n}\right| > \ell > 1$, alors la suite $(u_n)_{n\in\mathbb{N}}$ diverge. En effet, il suffit d'appliquer le théorème à la suite de terme général $\frac{1}{|u_n|}$ pour voir que $\lim_{n\to+\infty} |u_n| = +\infty$.
2. Toujours avec les notations du théorème, si $\ell = 1$ on ne peut rien dire.

Exemple 10.
Pour un nombre réel a, $a > 0$, calculer $\lim_{n\to+\infty} \sqrt[n]{a}$.
On va montrer que $\lim_{n\to+\infty} \sqrt[n]{a} = 1$. Si $a = 1$, c'est clair. Supposons $a > 1$. Écrivons $a = 1 + h$, avec $h > 0$. Comme

$$\left(1+\frac{h}{n}\right)^n \geqslant 1 + n\frac{h}{n} = 1 + h = a$$

(voir la preuve de la proposition 8) on a en appliquant la fonction racine n-ème, $\sqrt[n]{\cdot}$:

$$1 + \frac{h}{n} \geqslant \sqrt[n]{a} \geqslant 1.$$

On peut conclure grâce au théorème « des gendarmes » que $\lim_{n\to+\infty} \sqrt[n]{a} = 1$. Enfin, si $a < 1$, on applique le cas précédent à $b = \frac{1}{a} > 1$.

3.4. Approximation des réels par des décimaux

Proposition 10.
Soit $a \in \mathbb{R}$. Posons

$$u_n = \frac{E(10^n a)}{10^n}.$$

Alors u_n est une approximation décimale de a à 10^{-n} près, en particulier $\lim_{n\to+\infty} u_n = a$.

Exemple 11.
$\pi = 3,14159265\ldots$

$$\begin{aligned}
u_0 &= \tfrac{E(10^0\pi)}{10^0} = E(\pi) = 3\\
u_1 &= \tfrac{E(10^1\pi)}{10^1} = \tfrac{E(31,415...)}{10} = 3,1\\
u_2 &= \tfrac{E(10^2\pi)}{10^2} = \tfrac{E(314,15...)}{100} = 3,14\\
u_3 &= 3,141
\end{aligned}$$

Démonstration. D'après la définition de la partie entière, on a

$$E(10^n a) \leqslant 10^n a < E(10^n a) + 1$$

donc

$$u_n \leqslant a < u_n + \frac{1}{10^n}$$

ou encore
$$0 \leqslant a - u_n < \frac{1}{10^n}.$$
Or la suite de terme général $\frac{1}{10^n}$ est une suite géométrique de raison $\frac{1}{10}$, donc elle tend vers 0. On en déduit que $\lim_{n\to+\infty} u_n = a$.

□

Exercice 1.
Montrer que la suite $(u_n)_{n\in\mathbb{N}}$ de la proposition 10 est croissante.

Remarque.

1. Les u_n sont des nombres décimaux, en particulier ce sont des nombres rationnels.
2. Ceci fournit une démonstration de la densité de $\mathbb{Q}$ dans $\mathbb{R}$. Pour $\epsilon > 0$, et $I =]a-\epsilon, a+\epsilon[$, alors pour n assez grand, $u_n \in I \cap \mathbb{Q}$.

Mini-exercices.

1. Déterminer la limite de la suite $(u_n)_{n\in\mathbb{N}}$ de terme général $5^n - 4^n$.
2. Soit $v_n = 1 + a + a^2 + \cdots + a^n$. Pour quelle valeur de $a \in \mathbb{R}$ la suite $(v_n)_{n\geqslant 1}$ a pour limite 3 (lorsque $n \to +\infty$) ?
3. Calculer la limite de $\frac{1+2+2^2+\cdots+2^n}{2^n}$.
4. Montrer que la somme des racines n-èmes de l'unité est nulle.
5. Montrer que si $\sin(\frac{\theta}{2}) \neq 0$ alors $\frac{1}{2} + \cos(\theta) + \cos(2\theta) + \cdots + \cos(n\theta) = \frac{\sin\left((n+\frac{1}{2})\theta\right)}{2\sin(\frac{\theta}{2})}$ (penser à $e^{\mathrm{i}\theta}$).
6. Soit $(u_n)_{n\geqslant 2}$ la suite de terme général $u_n = \ln(1+\frac{1}{2}) \times \ln(1+\frac{1}{3}) \times \cdots \times \ln(1+\frac{1}{n})$. Déterminer la limite de $\frac{u_{n+1}}{u_n}$. Que peut-on en déduire ?
7. Déterminer la limite de $\frac{\pi^n}{1\times3\times5\times\cdots\times(2n+1)}$ (où $\pi = 3,14\ldots$).
8. Soit a un réel. Montrer que pour tout $\epsilon > 0$ il existe un couple $(m,n) \in \mathbb{Z} \times \mathbb{N}$ (et même une infinité) tel que $\left|a - \frac{m}{2^n}\right| \leqslant \epsilon$.

4. Théorème de convergence

4.1. Toute suite convergente est bornée

Revenons sur une propriété importante que nous avons déjà démontrée dans la section sur les limites.

Proposition 11.
Toute suite convergente est bornée.

La réciproque est fausse mais nous allons ajouter une hypothèse supplémentaire pour obtenir des résultats.

4.2. Suite monotone

Théorème 2.

Toute suite croissante et majorée est convergente.

Remarque.
Et aussi :
- Toute suite décroissante et minorée est convergente.
- Une suite croissante et qui n'est pas majorée tend vers $+\infty$.
- Une suite décroissante et qui n'est pas minorée tend vers $-\infty$.

Démonstration du théorème 2. Notons $A = \{u_n | n \in \mathbb{N}\} \subset \mathbb{R}$. Comme la suite $(u_n)_{n\in\mathbb{N}}$ est majorée, disons par le réel M, l'ensemble A est majoré par M, et de plus il est non vide. Donc d'après le théorème $\mathbb{R}4$ du chapitre sur les réels, l'ensemble A admet une borne supérieure : notons $\ell = \sup A$. Montrons que $\lim_{n\to+\infty} u_n = \ell$. Soit $\epsilon > 0$. Par la caractérisation de la borne supérieure, il existe un élément u_N de A tel que $\ell - \epsilon < u_N \leqslant \ell$. Mais alors pour $n \geqslant N$ on a $\ell - \epsilon < u_N \leqslant u_n \leqslant \ell$, et donc $|u_n - \ell| \leqslant \epsilon$. □

4.3. Deux exemples

La limite $\zeta(2)$

Soit $(u_n)_{n\geqslant 1}$ la suite de terme général :

$$u_n = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2}.$$

- La suite $(u_n)_{n\geqslant 1}$ est croissante : en effet $u_{n+1} - u_n = \frac{1}{(n+1)^2} > 0$.
- Montrons par récurrence que pour tout entier naturel $n \geqslant 1$ on a $u_n \leqslant 2 - \frac{1}{n}$.
  - Pour $n = 1$, on a $u_1 = 1 \leqslant 1 = 2 - \frac{1}{1}$.
  - Fixons $n \geqslant 1$ pour lequel on suppose $u_n \leqslant 2 - \frac{1}{n}$. Alors $u_{n+1} = u_n + \frac{1}{(n+1)^2} \leqslant 2 - \frac{1}{n} + \frac{1}{(n+1)^2}$. Or $\frac{1}{(n+1)^2} \leqslant \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, donc $u_{n+1} \leqslant 2 - \frac{1}{n+1}$, ce qui achève la récurrence.
- Donc la suite $(u_n)_{n\geqslant 1}$ est croissante et majorée par 2 : elle converge.

On note $\zeta(2)$ cette limite, vous montrerez plus tard qu'en fait $\zeta(2) = \frac{\pi^2}{6}$.

Suite harmonique

C'est la suite $(u_n)_{n\geqslant 1}$ de terme général :

$$u_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$

Calculons $\lim_{n\to+\infty} u_n$.
- La suite $(u_n)_{n\geqslant 1}$ est croissante : en effet $u_{n+1} - u_n = \frac{1}{n+1} > 0$.
- Minoration de $u_{2^p} - u_{2^{p-1}}$. On a $u_2 - u_1 = 1 + \frac{1}{2} - 1 = \frac{1}{2}$; $u_4 - u_2 = \frac{1}{3} + \frac{1}{4} > \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, et en général :
$$u_{2^p} - u_{2^{p-1}} = \underbrace{\frac{1}{2^{p-1}+1} + \frac{1}{2^{p-1}+2} + \cdots + \frac{1}{2^p}}_{2^{p-1}=2^p-2^{p-1} \text{ termes } \geqslant \frac{1}{2^p}} > 2^{p-1} \times \frac{1}{2^p} = \frac{1}{2}$$
- $\lim_{n\to+\infty} u_n = +\infty$. En effet
$$u_{2^p} - 1 = u_{2^p} - u_1 = (u_2 - u_1) + (u_4 - u_2) + \cdots + (u_{2^p} - u_{2^{p-1}}) \geqslant \frac{p}{2}$$
donc la suite $(u_n)_{n\geqslant 1}$ est croissante mais n'est pas bornée, donc elle tend vers $+\infty$.

4.4. Suites adjacentes

Définition 7.
Les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont dites *adjacentes* si

1. $(u_n)_{n\in\mathbb{N}}$ est croissante et $(v_n)_{n\in\mathbb{N}}$ est décroissante,
2. pour tout $n \geqslant 0$, on a $u_n \leqslant v_n$,

3. $\lim_{n\to+\infty}(v_n - u_n) = 0$.

Théorème 3.

Si les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont adjacentes, elles convergent vers la même limite.

Il y a donc deux résultats dans ce théorème, la convergence de (u_n) et (v_n) et en plus l'égalité des limites. Les termes de la suites sont ordonnées ainsi :

$$u_0 \leqslant u_1 \leqslant u_2 \leqslant \cdots \leqslant u_n \leqslant \cdots\cdots \leqslant v_n \leqslant \cdots \leqslant v_2 \leqslant v_1 \leqslant v_0$$

Démonstration.

- La suite $(u_n)_{n\in\mathbb{N}}$ est croissante et majorée par v_0, donc elle converge vers une limite ℓ.
- La suite $(v_n)_{n\in\mathbb{N}}$ est décroissante et minorée par u_0, donc elle converge vers une limite ℓ'.
- Donc $\ell' - \ell = \lim_{n\to+\infty}(v_n - u_n) = 0$, d'où $\ell' = \ell$.

□

Exemple 12.

Reprenons l'exemple de $\zeta(2)$. Soient (u_n) et (v_n) les deux suites définies pour $n \geqslant 1$ par

$$u_n = \sum_{k=1}^{n} \frac{1}{k^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} \quad \text{et} \quad v_n = u_n + \frac{2}{n+1}.$$

Montrons que (u_n) et (v_n) sont deux suites adjacentes :

1. (a) (u_n) est croissante car $u_{n+1} - u_n = \frac{1}{(n+1)^2} > 0$.

   (b) (v_n) est décroissante :
   $v_{n+1} - v_n = \frac{1}{(n+1)^2} + \frac{2}{n+2} - \frac{2}{n+1} = \frac{n+2+2(n+1)^2-2(n+1)(n+2)}{(n+2)(n+1)^2} = \frac{-n}{(n+2)(n+1)^2} < 0$
2. Pour tout $n \geqslant 1$: $v_n - u_n = \frac{2}{n+1} > 0$, donc $u_n \leqslant v_n$.
3. Enfin comme $v_n - u_n = \frac{2}{n+1}$ alors $\lim(v_n - u_n) = 0$.

Les suites (u_n) et (v_n) sont deux suites adjacentes, elles convergent donc vers une même limite finie ℓ. Nous avons en plus l'encadrement $u_n \leqslant \ell \leqslant v_n$ pour tout $n \geqslant 1$. Ceci fournit des approximations de la limite : par exemple pour $n = 3$, $1 + \frac{1}{4} + \frac{1}{9} \leqslant \ell \leqslant 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{2}$ donc $1,3611\ldots \leqslant \ell \leqslant 1,8611\ldots$

Exercice 2.

Soit $(u_n)_{n\geqslant 1}$ la suite de terme général :

$$u_n = 1 + \frac{1}{2^3} + \frac{1}{3^3} + \cdots + \frac{1}{n^3}.$$

Montrer que la suite $(u_n)_{n\geqslant 1}$ converge (on pourra considérer la suite $(v_n)_{n\geqslant 1}$ de terme général $v_n = u_n + \frac{1}{n^2}$).

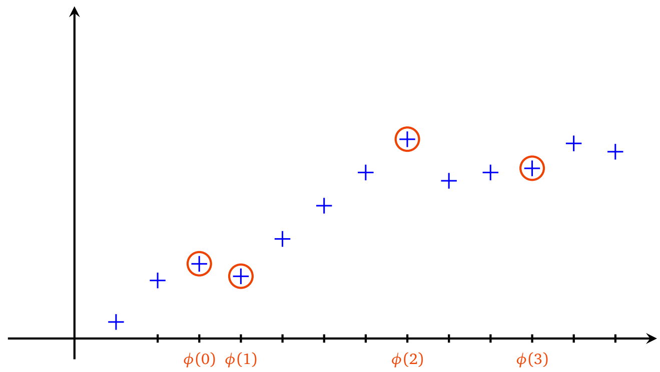
Remarque.

On note $\zeta(3)$ cette limite. On l'appelle aussi constante d'Apéry qui a prouvé en 1978 que $\zeta(3) \notin \mathbb{Q}$.

4.5. Théorème de Bolzano-Weierstrass

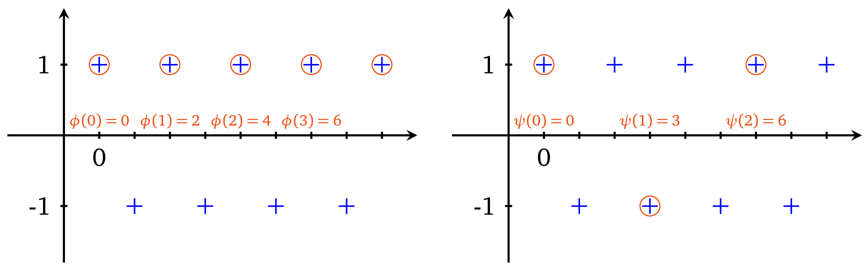
Définition 8.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite. Une ***suite extraite*** ou ***sous-suite*** de $(u_n)_{n\in\mathbb{N}}$ est une suite de la forme $(u_{\phi(n)})_{n\in\mathbb{N}}$, où $\phi : \mathbb{N} \to \mathbb{N}$ est une application strictement croissante.



Exemple 13.
Soit la suite $(u_n)_{n\in\mathbb{N}}$ de terme général $u_n = (-1)^n$.

- Si on considère $\phi : \mathbb{N} \to \mathbb{N}$ donnée par $\phi(n) = 2n$, alors la suite extraite correspondante a pour terme général $u_{\phi(n)} = (-1)^{2n} = 1$, donc la suite $(u_{\phi(n)})_{n\in\mathbb{N}}$ est constante égale à 1.
- Si on considère $\psi : \mathbb{N} \to \mathbb{N}$ donnée par $\psi(n) = 3n$, alors la suite extraite correspondante a pour terme général $u_{\psi(n)} = (-1)^{3n} = \big((-1)^3\big)^n = (-1)^n$. La suite $(u_{\psi(n)})_{n\in\mathbb{N}}$ est donc égale à $(u_n)_{n\in\mathbb{N}}$.



Proposition 12.
Soit $(u_n)_{n\in\mathbb{N}}$ une suite. Si $\lim_{n\to+\infty} u_n = \ell$, alors pour toute suite extraite $(u_{\phi(n)})_{n\in\mathbb{N}}$ on a $\lim_{n\to+\infty} u_{\phi(n)} = \ell$.

Démonstration. Soit $\epsilon > 0$. D'après la définition de limite (définition 4), il existe un entier naturel N tel que $n \geqslant N$ implique $|u_n - \ell| < \epsilon$. Comme l'application ϕ est strictement croissante, on montre facilement par récurrence que pour tout n, on a $\phi(n) \geqslant n$. Ceci implique en particulier que si $n \geqslant N$, alors aussi $\phi(n) \geqslant N$, et donc $|u_{\phi(n)} - \ell| < \epsilon$. Donc la définition de limite (définition 4) s'applique aussi à la suite extraite. □

Corollaire 2.
Soit $(u_n)_{n\in\mathbb{N}}$ une suite. Si elle admet une sous-suite divergente, ou bien si elle admet deux sous-suites convergeant vers des limites distinctes, alors elle diverge.

Exemple 14.
Soit la suite $(u_n)_{n\in\mathbb{N}}$ de terme général $u_n = (-1)^n$. Alors $(u_{2n})_{n\in\mathbb{N}}$ converge vers 1, et $(u_{2n+1})_{n\in\mathbb{N}}$ converge vers -1 (en fait ces deux sous-suites sont constantes). On en déduit que la suite $(u_n)_{n\in\mathbb{N}}$ diverge.

Exercice 3.
Soit $(u_n)_{n\in\mathbb{N}}$ une suite. On suppose que les deux sous-suites $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$ convergent vers la même limite ℓ. Montrer que $(u_n)_{n\in\mathbb{N}}$ converge également vers ℓ.

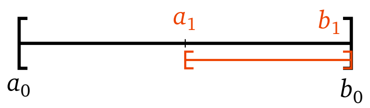
Terminons par un résultat théorique très important.

Théorème 4 (Théorème de Bolzano-Weierstrass)**.**
Toute suite bornée admet une sous-suite convergente.

Exemple 15.

1. On considère la suite $(u_n)_{n\in\mathbb{N}}$ de terme général $u_n = (-1)^n$. Alors on peut considérer les deux sous-suites $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$.
2. On considère la suite $(v_n)_{n\in\mathbb{N}}$ de terme général $v_n = \cos n$. Le théorème affirme qu'il existe une sous-suite convergente, mais il est moins facile de l'expliciter.

Démonstration du théorème 4. On procède par dichotomie. L'ensemble des valeurs de la suite est par hypothèse contenu dans un intervalle $[a,b]$. Posons $a_0 = a$, $b_0 = b$, $\phi(0) = 0$. Au moins l'un des deux intervalles $\left[a_0, \frac{a_0+b_0}{2}\right]$ ou $\left[\frac{a_0+b_0}{2}, b_0\right]$ contient u_n pour une infinité d'indices n. On note $[a_1, b_1]$ un tel intervalle, et on note $\phi(1)$ un entier $\phi(1) > \phi(0)$ tel que $u_{\phi(1)} \in [a_1, b_1]$.



En itérant cette construction, on construit pour tout entier naturel n un intervalle $[a_n, b_n]$, de longueur $\frac{b-a}{2^n}$, et un entier $\phi(n) > \phi(n-1)$ tel que $u_{\phi(n)} \in [a_n, b_n]$. Notons que par construction la suite $(a_n)_{n\in\mathbb{N}}$ est croissante et la suite $(b_n)_{n\in\mathbb{N}}$ est décroissante.
Comme de plus $\lim_{n\to+\infty}(b_n - a_n) = \lim_{n\to+\infty} \frac{b-a}{2^n} = 0$, les suites $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont adjacentes et donc convergent vers une même limite ℓ. On peut appliquer le théorème « des gendarmes » pour conclure que $\lim_{n\to+\infty} u_{\phi(n)} = \ell$. $\square$

Mini-exercices.

1. Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par $u_0 = 1$ et pour $n \geqslant 1$, $u_n = \sqrt{2 + u_{n-1}}$. Montrer que cette suite est croissante et majorée par 2. Que peut-on en conclure ?
2. Soit $(u_n)_{n\geqslant 2}$ la suite définie par $u_n = \frac{\ln 4}{\ln 5} \times \frac{\ln 6}{\ln 7} \times \frac{\ln 8}{\ln 9} \times \cdots \times \frac{\ln(2n)}{\ln(2n+1)}$. Étudier la croissance de la suite. Montrer que la suite (u_n) converge.
3. Soit $N \geqslant 1$ un entier et $(u_n)_{n\in\mathbb{N}}$ la suite de terme général $u_n = \cos(\frac{n\pi}{N})$. Montrer que la suite diverge.
4. Montrer que les suites de terme général $u_n = \sum_{k=1}^{n} \frac{1}{k!}$ et $v_n = u_n + \frac{1}{n\cdot(n!)}$ sont adjacentes. Que peut-on en déduire ?
5. Soit $(u_n)_{n\geqslant 1}$ la suite de terme général $\sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}$. On considère les deux suites extraites de terme général $v_n = u_{2n}$ et $w_n = u_{2n+1}$. Montrer que les deux suites $(v_n)_{n\geqslant 1}$ et $(w_n)_{n\geqslant 1}$ sont adjacentes. En déduire que la suite $(u_n)_{n\geqslant 1}$ converge.
6. Montrer qu'une suite bornée et divergente admet deux sous-suites convergeant vers des valeurs distinctes.

5. Suites récurrentes

Les suites récurrentes définies par une fonction forment une catégorie essentielle de suites.
Ce paragraphe est l'aboutissement de notre étude des suites, mais sa lecture nécessite aussi la maîtrise préalable de l'étude de fonctions (voir « Limites et fonctions continues »).

5.1. Suite récurrente définie par une fonction

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction. Une *suite récurrente* est définie par son premier terme et une relation permettant de calculer les termes de proche en proche :

$$u_0 \in \mathbb{R} \quad \text{et} \quad u_{n+1} = f(u_n) \quad \text{pour } n \geqslant 0.$$

Une suite récurrente est donc définie par deux données : un terme initial u_0, et une relation de récurrence $u_{n+1} = f(u_n)$. La suite s'écrit ainsi :

$$u_0, \quad u_1 = f(u_0), \quad u_2 = f(u_1) = f(f(u_0)), \quad u_3 = f(u_2) = f(f(f(u_0))), \ldots$$

Le comportement d'une telle suite peut très vite devenir complexe.

Exemple 16.
Soit $f(x) = 1 + \sqrt{x}$. Fixons $u_0 = 2$ et définissons pour $n \geqslant 0$: $u_{n+1} = f(u_n)$. C'est-à-dire $u_{n+1} = 1 + \sqrt{u_n}$. Alors les premiers termes de la suite sont :

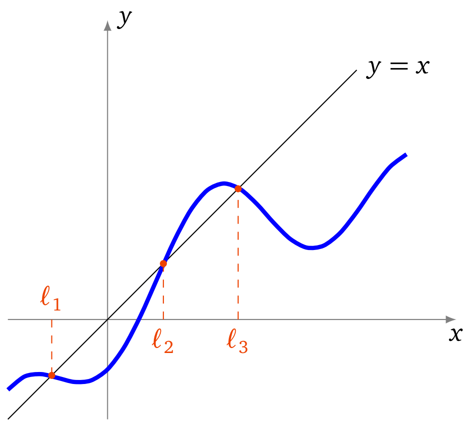
$$2, \quad 1+\sqrt{2}, \quad 1+\sqrt{1+\sqrt{2}}, \quad 1+\sqrt{1+\sqrt{1+\sqrt{2}}}, \quad 1+\sqrt{1+\sqrt{1+\sqrt{1+\sqrt{2}}}}, \ldots$$

Une suite récurrente donnée n'est pas forcément convergente. Lorsqu'elle admet une limite, l'ensemble des valeurs possibles est restreint par le résultat essentiel suivant.

Proposition 13.
Si f est une fonction continue et la suite récurrente (u_n) converge vers ℓ, alors ℓ est une solution de l'équation :

$$\boxed{f(\ell) = \ell}$$

Si on arrive à montrer que la limite existe, cette proposition affirme qu'elle est à chercher parmi les solutions de l'équation $f(\ell) = \ell$.



Une valeur ℓ, vérifiant $f(\ell) = \ell$ est un *point fixe* de f. La preuve est très simple et utilise essentiellement la continuité de la fonction f :

Démonstration. Lorsque $n \to +\infty$, $u_n \to \ell$ et donc aussi $u_{n+1} \to \ell$. Comme $u_n \to \ell$ et que f est continue alors la suite $(f(u_n)) \to f(\ell)$. La relation $u_{n+1} = f(u_n)$ devient à la limite (lorsque $n \to +\infty$) : $\ell = f(\ell)$. □

Nous allons étudier en détail deux cas particuliers, celui ou la fonction est croissante, puis celui ou la fonction est décroissante.

5.2. Cas d'une fonction croissante

Commençons par remarquer que pour une fonction croissante, le comportement de la suite (u_n) définie par récurrence est assez simple :

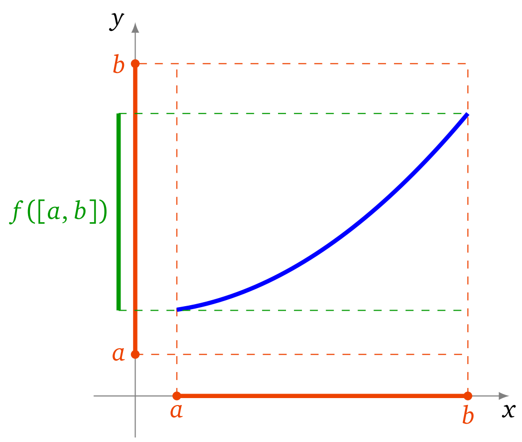
- Si $u_1 \geqslant u_0$ alors (u_n) est croissante.
- Si $u_1 \leqslant u_0$ alors (u_n) est décroissante.

La preuve est facile par récurrence : par exemple si $u_1 \geqslant u_0$, alors comme f est croissante on a $u_2 = f(u_1) \geqslant f(u_0) = u_1$. Partant de $u_2 \geqslant u_1$ on en déduit $u_3 \geqslant u_2$,...

Voici le résultat principal :

Proposition 14.
Si $f : [a,b] \to [a,b]$ une fonction continue et **croissante**, *alors quelque soit $u_0 \in [a,b]$, la suite récurrente (u_n) est monotone et converge vers $\ell \in [a,b]$ vérifiant* $\boxed{f(\ell) = \ell}$.

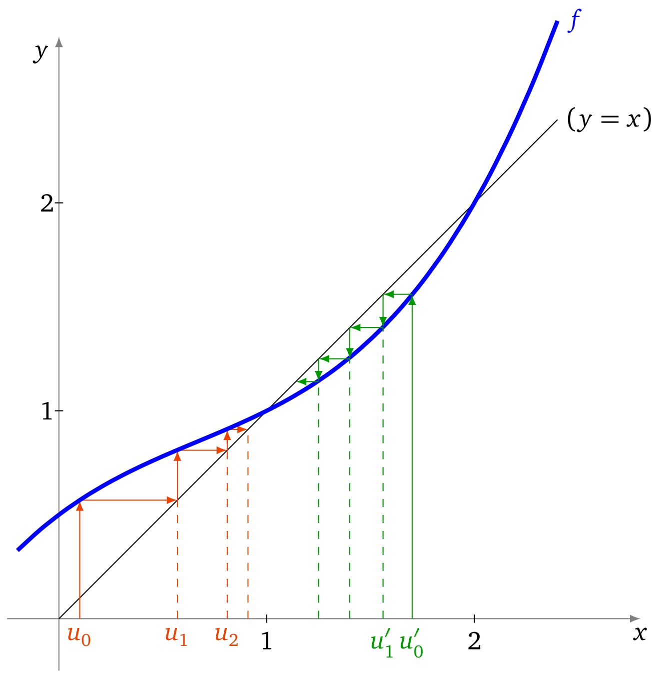
Il y a une hypothèse importante qui est un peu cachée : f va de l'intervalle $[a,b]$ dans lui-même. Dans la pratique, pour appliquer cette proposition, il faut commencer par choisir $[a,b]$ et vérifier que $f([a,b]) \subset [a,b]$.



Démonstration. La preuve est une conséquence des résultats précédents. Par exemple si $u_1 \geqslant u_0$ alors la suite (u_n) est croissante, comme par ailleurs elle est majorée par b, elle converge vers un réel ℓ. Par la proposition 13, on a $f(\ell) = \ell$. Si $u_1 \leqslant u_0$, (u_n) est une décroissante et minorée par a, et la conclusion est la même. $\square$

Exemple 17.
Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \frac{1}{4}(x^2 - 1)(x-2) + x$ et $u_0 \in [0,2]$. Étudions la suite (u_n) définie par récurrence : $u_{n+1} = f(u_n)$ (pour tout $n \geqslant 0$).

1. Étude de f
   (a) f est continue sur $\mathbb{R}$.
   (b) f est dérivable sur $\mathbb{R}$ et $f'(x) > 0$.
   (c) Sur l'intervalle $[0,2]$, f est strictement croissante.
   (d) Et comme $f(0) = \frac{1}{2}$ et $f(2) = 2$ alors $f([0,2]) \subset [0,2]$.
2. Graphe de f



Voici comment tracer la suite : on trace le graphe de f et la bissectrice $(y=x)$. On part d'une valeur u_0 (en rouge) sur l'axe des abscisses, la valeur $u_1 = f(u_0)$ se lit sur l'axe des ordonnées, mais on reporte la valeur de u_1 sur l'axe des abscisses par symétrie par rapport à la bissectrice. On recommence : $u_2 = f(u_1)$ se lit sur l'axe des ordonnées et on le reporte sur l'axe des abscisses, etc. On obtient ainsi une sorte d'escalier, et graphiquement on conjecture que la suite est croissante et tend vers 1. Si on part d'une autre valeur initiale u'_0 (en vert), c'est le même principe, mais cette fois on obtient un escalier qui descend.

3. Calcul des points fixes.
   Cherchons les valeurs x qui vérifient $(f(x)=x)$, autrement dit $(f(x)-x=0)$, mais
   $$f(x)-x = \frac{1}{4}(x^2-1)(x-2) \tag{1}$$
   Donc les points fixes sont les $\{-1,1,2\}$. La limite de (u_n) est donc à chercher parmi ces 3 valeurs.
4. Premier cas : $u_0 = 1$ ou $u_0 = 2$.
   Alors $u_1 = f(u_0) = u_0$ et par récurrence la suite (u_n) est constante (et converge donc vers u_0).
5. Deuxième cas : $0 \leqslant u_0 < 1$.
   - Comme $f([0,1]) \subset [0,1]$, la fonction f se restreint sur l'intervalle $[0,1]$ en une fonction $f : [0,1] \to [0,1]$.
   - De plus sur $[0,1]$, $f(x)-x \geqslant 0$. Cela se déduit de l'étude de f ou directement de l'expression (1).
   - Pour $u_0 \in [0,1[$, $u_1 = f(u_0) \geqslant u_0$ d'après le point précédent. Comme f est croissante, par récurrence, comme on l'a vu, la suite (u_n) est croissante.
   - La suite (u_n) est croissante et majorée par 1, donc elle converge. Notons ℓ sa limite.
   - D'une part ℓ doit être un point fixe de f : $f(\ell) = \ell$. Donc $\ell \in \{-1,1,2\}$.
   - D'autre part la suite (u_n) étant croissante avec $u_0 \geqslant 0$ et majorée par 1, donc $\ell \in [0,1]$.
   - Conclusion : si $0 \leqslant u_0 < 1$ alors (u_n) converge vers $\ell = 1$.
6. Troisième cas : $1 < u_0 < 2$.
   La fonction f se restreint en $f : [1,2] \to [1,2]$. Sur l'intervalle $[1,2]$, f est croissante mais cette fois $f(x) \leqslant x$. Donc $u_1 \leqslant u_0$, et la suite (u_n) est décroissante. La suite (u_n) étant minorée par 1, elle converge. Si on note ℓ sa limite alors d'une part $f(\ell) = \ell$, donc $\ell \in \{-1,1,2\}$, et d'autre part $\ell \in [1,2[$. Conclusion : (u_n) converge vers $\ell = 1$.

Le graphe de f joue un rôle très important, il faut le tracer même si on ne le demande pas explicitement. Il permet de se faire une idée très précise du comportement de la suite : Est-elle croissante ? Est-elle positive ?

Semble-t-elle converger ? Vers quelle limite ? Ces indications sont essentielles pour savoir ce qu'il faut montrer lors de l'étude de la suite.

5.3. Cas d'une fonction décroissante

Proposition 15.
Soit $f : [a,b] \to [a,b]$ une fonction continue et **décroissante**. *Soit $u_0 \in [a,b]$ et la suite récurrente (u_n) définie par $u_{n+1} = f(u_n)$. Alors :*

- *La sous-suite (u_{2n}) converge vers une limite ℓ vérifiant $f \circ f(\ell) = \ell$.*
- *La sous-suite (u_{2n+1}) converge vers une limite ℓ' vérifiant $f \circ f(\ell') = \ell'$.*

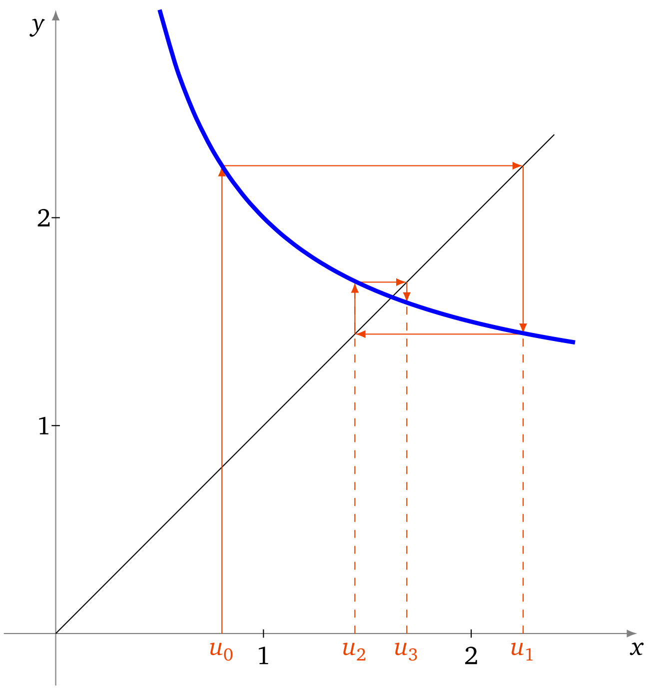
Il se peut (ou pas !) que $\ell = \ell'$.

Démonstration. La preuve se déduit du cas croissant. La fonction f étant décroissante, la fonction $f \circ f$ est croissante. Et on applique la proposition 14 à la fonction $f \circ f$ et à la sous-suite (u_{2n}) définie par récurrence $u_2 = f \circ f(u_0)$, $u_4 = f \circ f(u_2)$,...
De même en partant de u_1 et $u_3 = f \circ f(u_1)$,... □

Exemple 18.

$$f(x) = 1 + \frac{1}{x}, \qquad u_0 > 0, \qquad u_{n+1} = f(u_n) = 1 + \frac{1}{u_n}$$

1. Étude de f. La fonction $f :]0, +\infty[\to]0, +\infty[$ est une fonction continue et strictement décroissante.
2. Graphe de f.



Le principe pour tracer la suite est le même qu'auparavant : on place u_0, on trace $u_1 = f(u_0)$ sur l'axe des ordonnées et on le reporte par symétrie sur l'axe des abscisses,... On obtient ainsi une sorte d'escargot, et graphiquement on conjecture que la suite converge vers le point fixe de f. En plus on note que la suite des termes de rang pair semble une suite croissante, alors que la suite des termes de rang impair semble décroissante.

3. Points fixes de $f \circ f$.

$$f \circ f(x) = f\big(f(x)\big) = f\big(1+\frac{1}{x}\big) = 1 + \frac{1}{1+\frac{1}{x}} = 1 + \frac{x}{x+1} = \frac{2x+1}{x+1}$$

Donc

$$f \circ f(x) = x \iff \frac{2x+1}{x+1} = x \iff x^2 - x - 1 = 0 \iff x \in \left\{\frac{1-\sqrt{5}}{2}, \frac{1+\sqrt{5}}{2}\right\}$$

Comme la limite doit être positive, le seul point fixe à considérer est $\ell = \frac{1+\sqrt{5}}{2}$.

Attention ! Il y a un unique point fixe, mais on ne peut pas conclure à ce stade car f est définie sur $]0,+\infty[$ qui n'est pas un intervalle compact.

4. Premier cas $0 < u_0 \leqslant \ell = \frac{1+\sqrt{5}}{2}$.

Alors, $u_1 = f(u_0) \geqslant f(\ell) = \ell$; et par une étude de $f \circ f(x) - x$, on obtient que : $u_2 = f \circ f(u_0) \geqslant u_0$; $u_1 \geqslant f \circ f(u_1) = u_3$.

Comme $u_2 \geqslant u_0$ et $f \circ f$ est croissante, la suite (u_{2n}) est croissante. De même $u_3 \leqslant u_1$, donc la suite (u_{2n+1}) est décroissante. De plus comme $u_0 \leqslant u_1$, en appliquant f un nombre *pair* de fois, on obtient que $u_{2n} \leqslant u_{2n+1}$. La situation est donc la suivante :

$$u_0 \leqslant u_2 \leqslant \cdots \leqslant u_{2n} \leqslant \cdots \leqslant u_{2n+1} \leqslant \cdots \leqslant u_3 \leqslant u_1$$

La suite (u_{2n}) est croissante et majorée par u_1, donc elle converge. Sa limite ne peut être que l'unique point fixe de $f \circ f$: $\ell = \frac{1+\sqrt{5}}{2}$.

La suite (u_{2n+1}) est décroissante et minorée par u_0, donc elle converge aussi vers $\ell = \frac{1+\sqrt{5}}{2}$.

On en conclut que la suite (u_n) converge vers $\ell = \frac{1+\sqrt{5}}{2}$.

5. Deuxième cas $u_0 \geqslant \ell = \frac{1+\sqrt{5}}{2}$.

On montre de la même façon que (u_{2n}) est décroissante et converge vers $\frac{1+\sqrt{5}}{2}$, et que (u_{2n+1}) est croissante et converge aussi vers $\frac{1+\sqrt{5}}{2}$.

Mini-exercices.

1. Soit $f(x) = \frac{1}{9}x^3 + 1$, $u_0 = 0$ et pour $n \geqslant 0$: $u_{n+1} = f(u_n)$. Étudier en détail la suite (u_n) : (a) montrer que $u_n \geqslant 0$; (b) étudier et tracer le graphe de g ; (c) tracer les premiers termes de (u_n) ; (d) montrer que (u_n) est croissante ; (e) étudier la fonction $g(x) = f(x) - x$; (f) montrer que f admet deux points fixes sur $\mathbb{R}_+$, $0 < \ell < \ell'$; (g) montrer que $f([0,\ell]) \subset [0,\ell]$; (h) en déduire que (u_n) converge vers ℓ.
2. Soit $f(x) = 1 + \sqrt{x}$, $u_0 = 2$ et pour $n \geqslant 0$: $u_{n+1} = f(u_n)$. Étudier en détail la suite (u_n).
3. Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par : $u_0 \in [0,1]$ et $u_{n+1} = u_n - u_n^2$. Étudier en détail la suite (u_n).
4. Étudier la suite définie par $u_0 = 4$ et $u_{n+1} = \frac{4}{u_n+2}$.

Auteurs du chapitre

Auteurs : Arnaud Bodin, Niels Borne, Laura Desideri

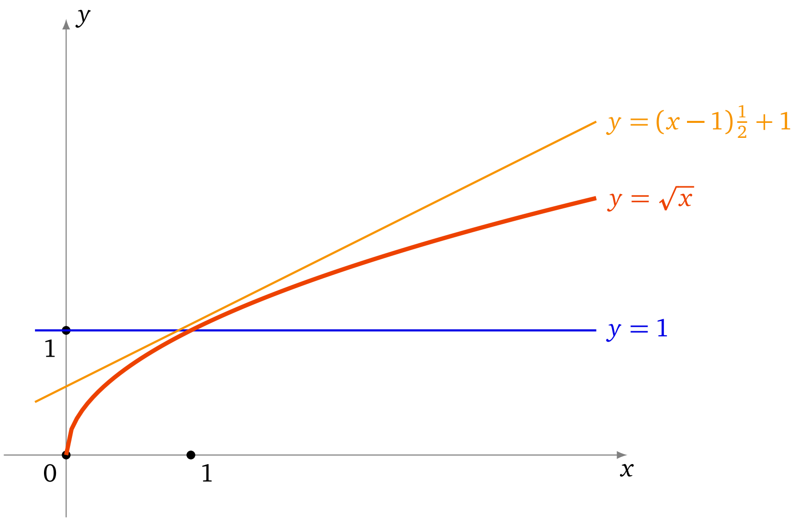
Dessins : Benjamin Boutin

Chapitre
4

Dérivée d'une fonction

Objectif

Nous souhaitons calculer $\sqrt{1,01}$ ou du moins en trouver une valeur approchée. Comme $1,01$ est proche de 1 et que $\sqrt{1} = 1$ on se doute bien que $\sqrt{1,01}$ sera proche de 1. Peut-on être plus précis ? Si l'on appelle f la fonction définie par $f(x) = \sqrt{x}$, alors la fonction f est une fonction continue en $x_0 = 1$. La continuité nous affirme que pour x suffisamment proche de x_0, $f(x)$ est proche de $f(x_0)$. Cela revient à dire que pour x au voisinage de x_0 on approche $f(x)$ par la constante $f(x_0)$.



Nous pouvons faire mieux qu'approcher notre fonction par une droite horizontale ! Essayons avec une droite quelconque. Quelle droite se rapproche le plus du graphe de f autour de x_0 ? Elle doit passer par le point $(x_0, f(x_0))$ et doit « coller » le plus possible au graphe : c'est la tangente au graphe en x_0. Une équation de la tangente est

$$y = (x - x_0)f'(x_0) + f(x_0)$$

où $f'(x_0)$ désigne le nombre dérivé de f en x_0.

On sait que pour $f(x) = \sqrt{x}$, on a $f'(x) = \frac{1}{2\sqrt{x}}$. Une équation de la tangente en $x_0 = 1$ est donc $y = (x-1)\frac{1}{2} + 1$. Et donc pour x proche de 1 on a $f(x) \approx (x-1)\frac{1}{2} + 1$. Qu'est-ce que cela donne pour notre calcul de $\sqrt{1,01}$? On pose $x = 1,01$ donc $f(x) \approx 1 + \frac{1}{2}(x-1) = 1 + \frac{0,01}{2} = 1,005$. Et c'est effectivement une très bonne de approximation de $\sqrt{0,01} = 1,00498\ldots$. En posant $h = x - 1$ on peut reformuler notre approximation en : $\sqrt{1+h} \approx 1 + \frac{1}{2}h$ qui est valable pour h proche de 0.

Dans ce chapitre nous allons donc définir ce qu'est la dérivée d'une fonction et établir les formules des dérivées des fonctions usuelles. Enfin, pour connaître l'erreur des approximations, il nous faudra travailler beaucoup plus afin d'obtenir le théorème des accroissements finis.

1. Dérivée

1.1. Dérivée en un point

Soit I un intervalle ouvert de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction. Soit $x_0 \in I$.

Définition 1.
f est ***dérivable en x_0*** si le ***taux d'accroissement*** $\frac{f(x)-f(x_0)}{x-x_0}$ a une limite finie lorsque x tend vers x_0. La limite s'appelle alors le ***nombre dérivé*** de f en x_0 et est noté $f'(x_0)$. Ainsi

$$\boxed{f'(x_0) = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}}$$

Définition 2.
f est ***dérivable sur I*** si f est dérivable en tout point $x_0 \in I$. La fonction $x \mapsto f'(x)$ est la ***fonction dérivée*** de f, elle se note f' ou $\frac{df}{dx}$.

Exemple 1.
La fonction définie par $f(x) = x^2$ est dérivable en tout point $x_0 \in \mathbb{R}$. En effet :

$$\frac{f(x)-f(x_0)}{x-x_0} = \frac{x^2-x_0^2}{x-x_0} = \frac{(x-x_0)(x+x_0)}{x-x_0} = x + x_0 \xrightarrow[x\to x_0]{} 2x_0.$$

On a même montré que le nombre dérivé de f en x_0 est $2x_0$, autrement dit : $f'(x) = 2x$.

Exemple 2.
Montrons que la dérivée de $f(x) = \sin x$ est $f'(x) = \cos x$. Nous allons utiliser les deux assertions suivantes :

$$\frac{\sin x}{x} \xrightarrow[x\to 0]{} 1 \qquad \text{et} \qquad \sin p - \sin q = 2 \sin \frac{p-q}{2} \cdot \cos \frac{p+q}{2}.$$

Remarquons déjà que la première assertion prouve $\frac{f(x)-f(0)}{x-0} = \frac{\sin x}{x} \to 1$ et donc f est dérivable en $x_0 = 0$ et $f'(0) = 1$.
Pour x_0 quelconque on écrit :

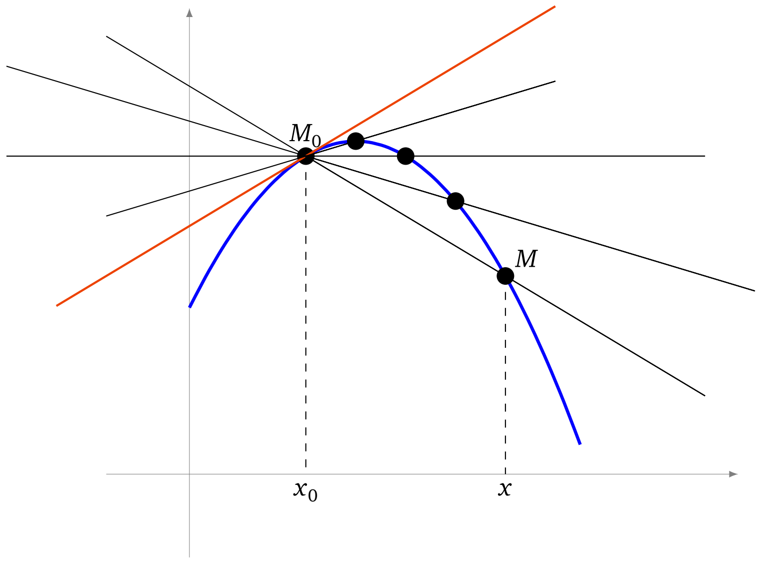
$$\frac{f(x)-f(x_0)}{x-x_0} = \frac{\sin x - \sin x_0}{x-x_0} = \frac{\sin \frac{x-x_0}{2}}{\frac{x-x_0}{2}} \cdot \cos \frac{x+x_0}{2}.$$

Lorsque $x \to x_0$ alors d'une part $\cos \frac{x+x_0}{2} \to \cos x_0$ et d'autre part en posant $u = \frac{x-x_0}{2}$ alors $u \to 0$ et on a $\frac{\sin u}{u} \to 1$. Ainsi $\frac{f(x)-f(x_0)}{x-x_0} \to \cos x_0$ et donc $f'(x) = \cos x$.

1.2. Tangente

La droite qui passe par les points distincts $(x_0, f(x_0))$ et $(x, f(x))$ a pour coefficient directeur $\frac{f(x)-f(x_0)}{x-x_0}$. À la limite on trouve que le coefficient directeur de la tangente est $f'(x_0)$. Une équation de la ***tangente*** au point $(x_0, f(x_0))$ est donc :

$$\boxed{y = (x-x_0)f'(x_0) + f(x_0)}$$



1.3. Autres écritures de la dérivée

Voici deux autres formulations de la dérivabilité de f en x_0.

Proposition 1.

- *f est dérivable en x_0 si et seulement si $\lim_{h\to 0} \dfrac{f(x_0+h)-f(x_0)}{h}$ existe et est finie.*
- *f est dérivable en x_0 si et seulement s'il existe $\ell \in \mathbb{R}$ (qui sera $f'(x_0)$) et une fonction $\epsilon : I \to \mathbb{R}$ telle que $\epsilon(x) \xrightarrow[x\to x_0]{} 0$ avec*

$$f(x) = f(x_0) + (x-x_0)\ell + (x-x_0)\epsilon(x).$$

Démonstration. Il s'agit juste de reformuler la définition de $f'(x_0)$. Par exemple, après division par $x - x_0$, la deuxième écriture devient

$$\frac{f(x)-f(x_0)}{x-x_0} = \ell + \epsilon(x).$$

□

Proposition 2.

Soit I un intervalle ouvert, $x_0 \in I$ et soit $f : I \to \mathbb{R}$ une fonction.

- *Si f est dérivable en x_0 alors f est continue en x_0.*
- *Si f est dérivable sur I alors f est continue sur I.*

Démonstration. Supposons f dérivable en x_0 et montrons qu'elle est aussi continue en ce point. Voici une démonstration concise : partant de l'écriture alternative donnée dans la proposition 1, nous écrivons

$$f(x) = f(x_0) + \underbrace{(x-x_0)\ell}_{\to 0} + \underbrace{(x-x_0)\epsilon(x)}_{\to 0}.$$

Donc $f(x) \to f(x_0)$ lorsque $x \to x_0$ et ainsi f est continue en x_0.

On reprend cette démonstration sans utiliser les limites mais uniquement la définition de continuité et dérivabilité : fixons $\epsilon' > 0$ et écrivons $f(x) = f(x_0) + (x-x_0)\ell + (x-x_0)\epsilon(x)$ grâce à la proposition 1, où $\epsilon(x) \xrightarrow[x\to x_0]{} 0$ et $\ell = f'(x_0)$. Choisissons $\delta > 0$ de sorte qu'il vérifie tous les points suivants :

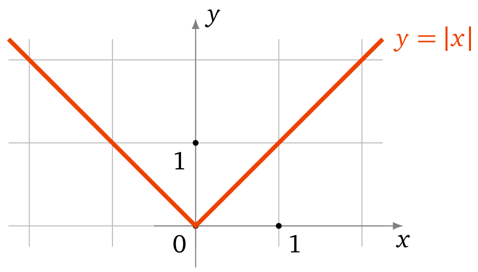
- $\delta \leqslant 1$,
- $\delta|\ell| < \epsilon'$,
- si $|x-x_0| < \delta$ alors $|\epsilon(x)| < \epsilon'$ (c'est possible car $\epsilon(x) \to 0$).

Alors l'égalité ci-dessus devient :

$$\begin{aligned}
\left|f(x)-f(x_0)\right| &= \left|(x-x_0)\ell+(x-x_0)\epsilon(x)\right| \\
&\leqslant |x-x_0|\cdot|\ell|+|x-x_0|\cdot|\epsilon(x)| \\
&\leqslant \delta|\ell| \quad + \quad \delta\epsilon' \qquad \text{pour } |x-x_0|<\delta \\
&\leqslant \epsilon'+\epsilon'=2\epsilon'
\end{aligned}$$

Nous venons de prouver que si $|x-x_0|<\delta$ alors $\left|f(x)-f(x_0)\right|<2\epsilon'$, ce qui exprime exactement que f est continue en x_0. □

Remarque.
La réciproque est **fausse** : par exemple, la fonction valeur absolue est continue en 0 mais n'est pas dérivable en 0.



En effet, le taux d'accroissement de $f(x)=|x|$ en $x_0=0$ vérifie :

$$\frac{f(x)-f(0)}{x-0}=\frac{|x|}{x}=\begin{cases}+1 & \text{si } x>0\\ -1 & \text{si } x<0\end{cases}.$$

Il y a bien une limite à droite (qui vaut $+1$), une limite à gauche (qui vaut -1) mais elles ne sont pas égales : il n'y a pas de limite en 0. Ainsi f n'est pas dérivable en $x=0$.
Cela se lit aussi sur le dessin, il y a une demi-tangente à droite, une demi-tangente à gauche, mais elles ont des directions différentes.

Mini-exercices.

1. Montrer que la fonction $f(x)=x^3$ est dérivable en tout point $x_0\in\mathbb{R}$ et que $f'(x_0)=3x_0^2$.
2. Montrer que la fonction $f(x)=\sqrt{x}$ est dérivable en tout point $x_0>0$ et que $f'(x_0)=\frac{1}{2\sqrt{x_0}}$.
3. Montrer que la fonction $f(x)=\sqrt{x}$ (qui est continue en $x_0=0$) n'est pas dérivable en $x_0=0$.
4. Calculer l'équation de la tangente (T_0) à la courbe d'équation $y=x^3-x^2-x$ au point d'abscisse $x_0=2$. Calculer x_1 afin que la tangente (T_1) au point d'abscisse x_1 soit parallèle à (T_0).
5. Montrer que si une fonction f est paire et dérivable, alors f' est une fonction impaire.

2. Calcul des dérivées

2.1. Somme, produit,...

Proposition 3.
Soient $f, g : I \to \mathbb{R}$ deux fonctions dérivables sur I. Alors pour tout $x \in I$:
- $(f+g)'(x) = f'(x) + g'(x)$
- $(\lambda f)'(x) = \lambda f'(x)$ *où λ est un réel fixé*
- $(f \times g)'(x) = f'(x)g(x) + f(x)g'(x)$
- $\left(\frac{1}{f}\right)'(x) = -\frac{f'(x)}{f(x)^2}$ *(si $f(x) \neq 0$)*
- $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$ *(si $g(x) \neq 0$)*

Remarque.
Il est plus facile de mémoriser les égalités de fonctions :

$$\boxed{(f+g)' = f' + g' \qquad (\lambda f)' = \lambda f' \qquad (f \times g)' = f'g + fg'}$$

$$\boxed{\left(\frac{1}{f}\right)' = -\frac{f'}{f^2} \qquad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}}$$

Démonstration. Prouvons par exemple $(f \times g)' = f'g + fg'$.
Fixons $x_0 \in I$. Nous allons réécrire le taux d'accroissement de $f(x) \times g(x)$:

$$\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} g(x) + \frac{g(x) - g(x_0)}{x - x_0} f(x_0)$$
$$\xrightarrow[x \to x_0]{} f'(x_0)g(x_0) + g'(x_0)f(x_0).$$

Ceci étant vrai pour tout $x_0 \in I$ la fonction $f \times g$ est dérivable sur I de dérivée $f'g + fg'$. □

2.2. Dérivée de fonctions usuelles

Le tableau de gauche est un résumé des principales formules à connaître, x est une variable. Le tableau de droite est celui des compositions (voir paragraphe suivant), u représente une fonction $x \mapsto u(x)$.

Fonction	**Dérivée**
x^n	$nx^{n-1} \quad (n \in \mathbb{Z})$
$\frac{1}{x}$	$-\frac{1}{x^2}$
$\sqrt{x}$	$\frac{1}{2}\frac{1}{\sqrt{x}}$
x^α	$\alpha x^{\alpha-1} \quad (\alpha \in \mathbb{R})$
e^x	e^x
$\ln x$	$\frac{1}{x}$
$\cos x$	$-\sin x$
$\sin x$	$\cos x$
$\tan x$	$1 + \tan^2 x = \frac{1}{\cos^2 x}$

Fonction	**Dérivée**
u^n	$nu'u^{n-1} \quad (n \in \mathbb{Z})$
$\frac{1}{u}$	$-\frac{u'}{u^2}$
$\sqrt{u}$	$\frac{1}{2}\frac{u'}{\sqrt{u}}$
u^α	$\alpha u'u^{\alpha-1} \quad (\alpha \in \mathbb{R})$
e^u	$u'e^u$
$\ln u$	$\frac{u'}{u}$
$\cos u$	$-u' \sin u$
$\sin u$	$u' \cos u$
$\tan u$	$u'(1 + \tan^2 u) = \frac{u'}{\cos^2 u}$

Remarque.

- Notez que les formules pour x^n, $\frac{1}{x}$, $\sqrt{x}$ et x^α sont aussi des conséquences de la dérivée de l'exponentielle. Par exemple $x^\alpha = e^{\alpha \ln x}$ et donc
$$\frac{d}{dx}(x^\alpha) = \frac{d}{dx}(e^{\alpha \ln x}) = \alpha\frac{1}{x}e^{\alpha \ln x} = \alpha\frac{1}{x}x^\alpha = \alpha x^{\alpha-1}.$$
- Si vous devez dériver une fonction avec un exposant dépendant de x il faut absolument repasser à la forme exponentielle. Par exemple si $f(x) = 2^x$ alors on réécrit d'abord $f(x) = e^{x \ln 2}$ pour pouvoir calculer $f'(x) = \ln 2 \cdot e^{x\ln 2} = \ln 2 \cdot 2^x$.

2.3. Composition

Proposition 4.
Si f est dérivable en x et g est dérivable en $f(x)$ alors $g\circ f$ est dérivable en x de dérivée :
$$\boxed{\big(g\circ f\big)'(x) = g'\big(f(x)\big)\cdot f'(x)}$$

Démonstration. La preuve est similaire à celle ci-dessus pour le produit en écrivant cette fois :
$$\begin{aligned}\frac{g\circ f(x) - g\circ f(x_0)}{x-x_0} &= \frac{g\big(f(x)\big) - g\big(f(x_0)\big)}{f(x)-f(x_0)} \times \frac{f(x)-f(x_0)}{x-x_0} \\ &\xrightarrow[x\to x_0]{} g'\big(f(x_0)\big)\times f'(x_0).\end{aligned}$$
∎

Exemple 3.
Calculons la dérivée de $\ln(1+x^2)$. Nous avons $g(x) = \ln(x)$ avec $g'(x) = \frac{1}{x}$; et $f(x) = 1+x^2$ avec $f'(x) = 2x$. Alors la dérivée de $\ln(1+x^2) = g\circ f(x)$ est
$$\big(g\circ f\big)'(x) = g'\big(f(x)\big)\cdot f'(x) = g'\big(1+x^2\big)\cdot 2x = \frac{2x}{1+x^2}.$$

Corollaire 1.
Soit I un intervalle ouvert. Soit $f : I \to J$ dérivable et bijective dont on note $f^{-1} : J \to I$ la bijection réciproque. Si f' ne s'annule pas sur I alors f^{-1} est dérivable et on a pour tout $x \in J$:
$$\boxed{\big(f^{-1}\big)'(x) = \frac{1}{f'\big(f^{-1}(x)\big)}}$$

Démonstration. Notons $g = f^{-1}$ la bijection réciproque de f. Soit $y_0 \in J$ et $x_0 \in I$ tel que $y_0 = f(x_0)$. Le taux d'accroissement de g en y_0 est :
$$\frac{g(y)-g(y_0)}{y-y_0} = \frac{g(y)-x_0}{f\big(g(y)\big)-f(x_0)}$$
Lorsque $y \to y_0$ alors $g(y) \to g(y_0) = x_0$ et donc ce taux d'accroissement tend vers $\frac{1}{f'(x_0)}$. Ainsi $g'(y_0) = \frac{1}{f'(x_0)}$. ∎

Remarque.
Il peut être plus simple de retrouver la formule à chaque fois en dérivant l'égalité
$$f\big(g(x)\big) = x$$
où $g = f^{-1}$ est la bijection réciproque de f.

En effet à droite la dérivée de x est 1 ; à gauche la dérivée de $f\big(g(x)\big) = f \circ g(x)$ est $f'\big(g(x)\big) \cdot g'(x)$. L'égalité $f\big(g(x)\big) = x$ conduit donc à l'égalité des dérivées :

$$f'\big(g(x)\big) \cdot g'(x) = 1.$$

Mais $g = f^{-1}$ donc

$$\big(f^{-1}\big)'(x) = \frac{1}{f'\big(f^{-1}(x)\big)}.$$

Exemple 4.
Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction définie par $f(x) = x + \exp(x)$. Étudions f en détail.
Tout d'abord :

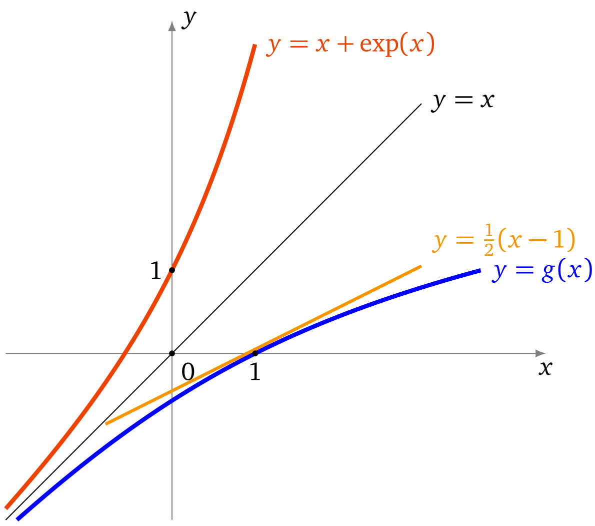
1. f est dérivable car f est la somme de deux fonctions dérivables. En particulier f est continue.
2. f est strictement croissante car f est la somme de deux fonctions strictement croissante.
3. f est une bijection car $\lim_{x\to-\infty} f(x) = -\infty$ et $\lim_{x\to+\infty} f(x) = +\infty$.
4. $f'(x) = 1 + \exp(x)$ ne s'annule jamais (pour tout $x \in \mathbb{R}$).

Notons $g = f^{-1}$ la bijection réciproque de f. Même si on ne sait pas a priori exprimer g, on peut malgré tout connaître des informations sur cette fonction : par le corollaire ci-dessus g est dérivable et l'on calcule g' en dérivant l'égalité $f\big(g(x)\big) = x$. Ce qui donne $f'\big(g(x)\big) \cdot g'(x) = 1$ et donc ici

$$g'(x) = \frac{1}{f'\big(g(x)\big)} = \frac{1}{1 + \exp\big(g(x)\big)}.$$

Pour cette fonction f particulière on peut préciser davantage : comme $f\big(g(x)\big) = x$ alors $g(x) + \exp\big(g(x)\big) = x$ donc $\exp\big(g(x)\big) = x - g(x)$. Cela conduit à :

$$g'(x) = \frac{1}{1 + x - g(x)}.$$



Par exemple $f(0) = 1$ donc $g(1) = 0$ et donc $g'(1) = \frac{1}{2}$. Autrement dit $\big(f^{-1}\big)'(1) = \frac{1}{2}$. L'équation de la tangente au graphe de f^{-1} au point d'abscisse $x_0 = 1$ est donc $y = \frac{1}{2}(x-1)$.

2.4. Dérivées successives

Soit $f : I \to \mathbb{R}$ une fonction dérivable et soit f' sa dérivée. Si la fonction $f' : I \to \mathbb{R}$ est aussi dérivable on note $f'' = (f')'$ la ***dérivée seconde*** de f. Plus généralement on note :

$$f^{(0)} = f, \quad f^{(1)} = f', \quad f^{(2)} = f'' \quad \text{et} \quad f^{(n+1)} = \left(f^{(n)}\right)'$$

Si la ***dérivée*** n***-ième*** $f^{(n)}$ existe on dit que f est n ***fois dérivable***.

Théorème 1 (Formule de Leibniz).

$$\boxed{\left(f \cdot g\right)^{(n)} = f^{(n)} \cdot g + \binom{n}{1} f^{(n-1)} \cdot g^{(1)} + \cdots + \binom{n}{k} f^{(n-k)} \cdot g^{(k)} + \cdots + f \cdot g^{(n)}}$$

Autrement dit :

$$\left(f \cdot g\right)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} \cdot g^{(k)}.$$

La démonstration est similaire à celle de la formule du binôme de Newton et les coefficients que l'on obtient sont les mêmes.

Exemple 5.

- Pour $n = 1$ on retrouve $(f \cdot g)' = f'g + fg'$.
- Pour $n = 2$, on a $(f \cdot g)'' = f''g + 2f'g' + fg''$.

Exemple 6.

Calculons les dérivées n-ème de $\exp(x) \cdot (x^2 + 1)$ pour tout $n \geqslant 0$. Notons $f(x) = \exp(x)$ alors $f'(x) = \exp(x)$, $f''(x) = \exp(x)$,..., $f^{(k)}(x) = \exp(x)$. Notons $g(x) = x^2 + 1$ alors $g'(x) = 2x$, $g''(x) = 2$ et pour $k \geqslant 3$, $g^{(k)}(x) = 0$.

Appliquons la formule de Leibniz :

$$\left(f \cdot g\right)^{(n)}(x) = f^{(n)}(x) \cdot g(x) + \binom{n}{1} f^{(n-1)}(x) \cdot g^{(1)}(x) + \\ + \binom{n}{2} f^{(n-2)}(x) \cdot g^{(2)}(x) + \binom{n}{3} f^{(n-3)}(x) \cdot g^{(3)}(x) + \cdots$$

On remplace $f^{(k)}(x) = \exp(x)$ et on sait que $g^{(3)}(x) = 0$, $g^{(4)}(x) = 0, \ldots$ Donc cette somme ne contient que les trois premiers termes :

$$\left(f \cdot g\right)^{(n)}(x) = \exp(x) \cdot (x^2 + 1) + \binom{n}{1} \exp(x) \cdot 2x + \binom{n}{2} \exp(x) \cdot 2.$$

Que l'on peut aussi écrire :

$$\left(f \cdot g\right)^{(n)}(x) = \exp(x) \cdot \left(x^2 + 2nx + n(n-1) + 1\right).$$

Mini-exercices.

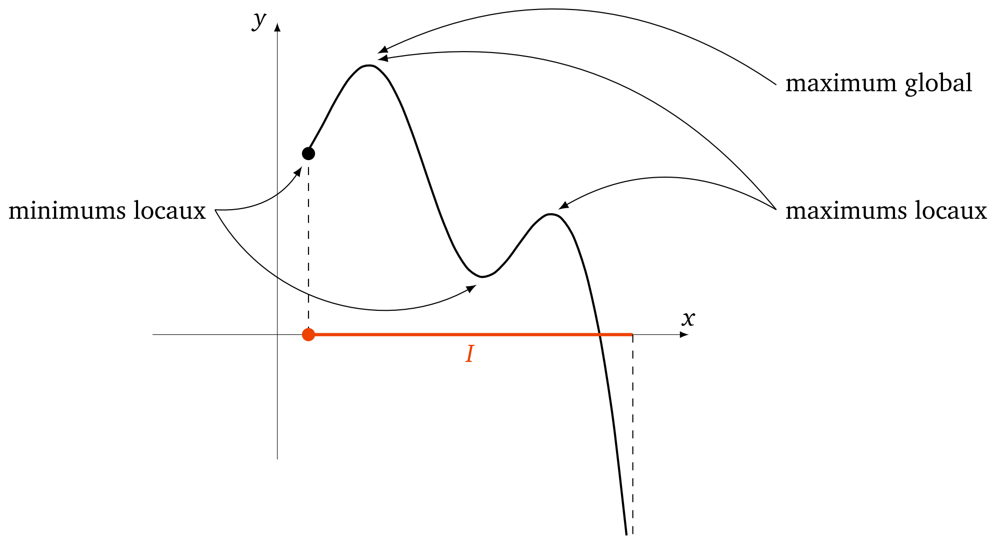
1. Calculer les dérivées des fonctions suivantes : $f_1(x) = x \ln x$, $f_2(x) = \sin \frac{1}{x}$, $f_3(x) = \sqrt{1 + \sqrt{1 + x^2}}$, $f_4(x) = \left(\ln(\frac{1+x}{1-x})\right)^{\frac{1}{3}}$, $f_5(x) = x^x$, $f_6(x) = \arctan x + \arctan \frac{1}{x}$.
2. On note $\Delta(f) = \frac{f'}{f}$. Calculer $\Delta(f \times g)$.
3. Soit $f :]1, +\infty[\to]-1, +\infty[$ définie par $f(x) = x\ln(x) - x$. Montrer que f est une bijection. Notons $g = f^{-1}$. Calculer $g(0)$ et $g'(0)$.
4. Calculer les dérivées successives de $f(x) = \ln(1 + x)$.
5. Calculer les dérivées successives de $f(x) = \ln(x) \cdot x^3$.

3. Extremum local, théorème de Rolle

3.1. Extremum local

Soit $f : I \to \mathbb{R}$ une fonction définie sur un intervalle I.

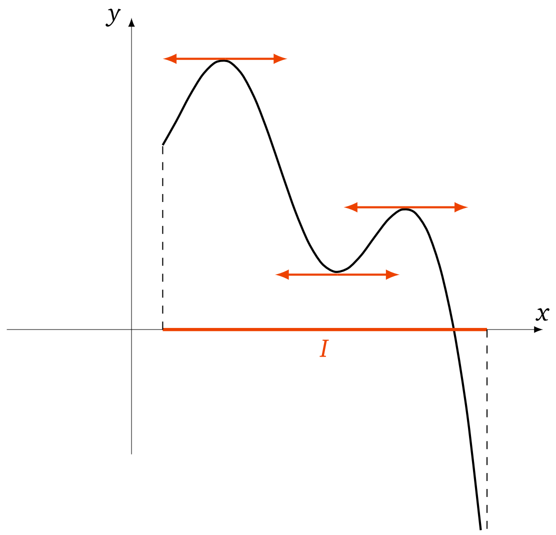
Définition 3.
- On dit que x_0 est un ***point critique*** de f si $f'(x_0) = 0$.
- On dit que f admet un ***maximum local en*** x_0 (resp. un ***minimum local en*** x_0) s'il existe un intervalle ouvert J contenant x_0 tel que
$$\text{pour tout } x \in I \cap J \quad f(x) \leqslant f(x_0)$$
(resp. $f(x) \geqslant f(x_0)$).
- On dit que f admet un ***extremum local en*** x_0 si f admet un maximum local ou un minimum local en ce point.



Dire que f a un maximum local en x_0 signifie que $f(x_0)$ est la plus grande des valeurs $f(x)$ pour les x proches de x_0. On dit que $f : I \to \mathbb{R}$ admet un ***maximum global*** en x_0 si pour toutes les autres valeurs $f(x)$, $x \in I$, on a $f(x) \leqslant f(x_0)$ (on ne regarde donc pas seulement les $f(x)$ pour x proche de x_0). Bien sûr un maximum global est aussi un maximum local, mais la réciproque est fausse.

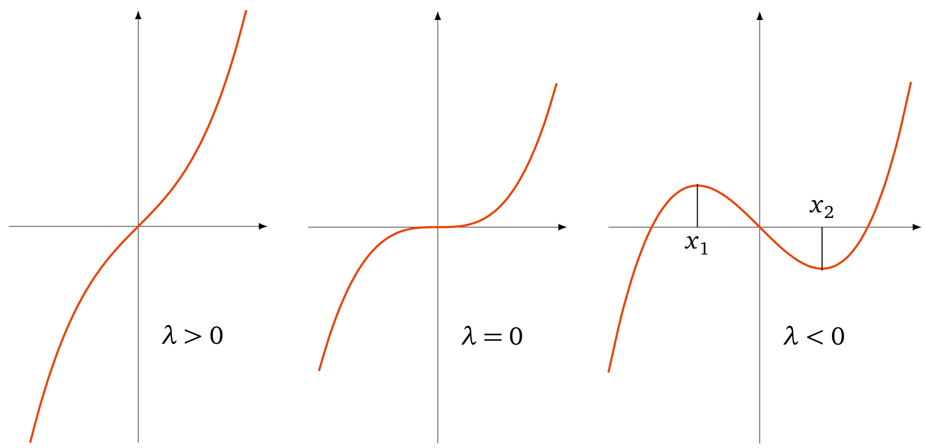
Théorème 2.
Soit I un intervalle ouvert et $f : I \to \mathbb{R}$ une fonction dérivable. Si f admet un maximum local (ou un minimum local) en x_0 alors $f'(x_0) = 0$.

En d'autres termes, un maximum local (ou un minimum local) x_0 est toujours un point critique. Géométriquement, au point $(x_0, f(x_0))$ la tangente au graphe est horizontale.



Exemple 7.
Étudions les extremums de la fonction f_λ définie par $f_\lambda(x) = x^3 + \lambda x$ en fonction du paramètre $\lambda \in \mathbb{R}$. La dérivée est $f'_\lambda(x) = 3x^2 + \lambda$. Si x_0 est un extremum local alors $f'_\lambda(x_0) = 0$.

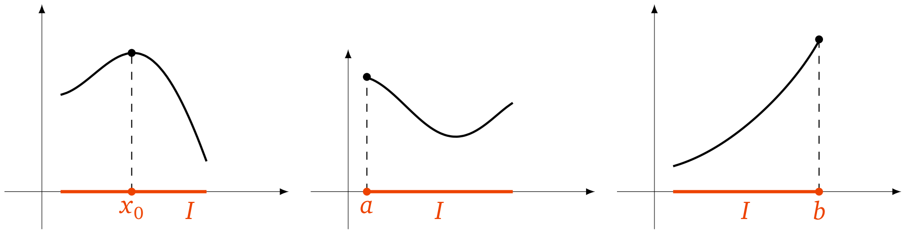
- Si $\lambda > 0$ alors $f'_\lambda(x) > 0$ et ne s'annule jamais il n'y a pas de points critiques donc pas non plus d'extremums. En anticipant sur la suite : f_λ est strictement croissante sur $\mathbb{R}$.
- Si $\lambda = 0$ alors $f'_\lambda(x) = 3x^2$. Le seul point critique est $x_0 = 0$. Mais ce n'est ni un maximum local, ni un minimum local. En effet si $x < 0$, $f_0(x) < 0 = f_0(0)$ et si $x > 0$, $f_0(x) > 0 = f_0(0)$.
- Si $\lambda < 0$ alors $f'_\lambda(x) = 3x^2 - |\lambda| = 3\left(x + \sqrt{\frac{|\lambda|}{3}}\right)\left(x - \sqrt{\frac{|\lambda|}{3}}\right)$. Il y a deux points critiques $x_1 = -\sqrt{\frac{|\lambda|}{3}}$ et $x_2 = +\sqrt{\frac{|\lambda|}{3}}$. En anticipant sur la suite : $f'_\lambda(x) > 0$ sur $]-\infty, x_1[$ et $]x_2, +\infty[$ et $f'_\lambda(x) < 0$ sur $]x_1, x_2[$; maintenant f_λ est croissante sur $]-\infty, x_1[$, puis décroissante sur $]x_1, x_2[$, donc x_1 est un maximum local. D'autre part f_λ est décroissante sur $]x_1, x_2[$ puis croissante sur $]x_2, +\infty[$ donc x_2 est un minimum local.



Remarque.

1. La réciproque du théorème 2 est fausse. Par exemple la fonction $f : \mathbb{R} \to \mathbb{R}$, définie par $f(x) = x^3$ vérifie $f'(0) = 0$ mais $x_0 = 0$ n'est ni maximum local ni un minimum local.
2. L'intervalle du théorème 2 est ouvert. Pour le cas d'un intervalle fermé, il faut faire attention aux extrémités. Par exemple si $f : [a, b] \to \mathbb{R}$ est une fonction dérivable qui admet un extremum en x_0, alors on est dans l'une des situations suivantes :
   - $x_0 = a$,
   - $x_0 = b$,
   - $x_0 \in]a, b[$ et dans ce cas on a bien $f'(x_0) = 0$ par le théorème 2.

Aux extrémités on ne peut rien dire pour $f'(a)$ et $f'(b)$, comme le montre les différents maximums sur les dessins suivants.



3. Pour déterminer $\max_{[a,b]} f$ et $\min_{[a,b]} f$ (où $f : [a,b] \to \mathbb{R}$ est une fonction dérivable) il faut comparer les valeurs de f aux différents points critiques et en a et en b.

Preuve du théorème. Supposons que x_0 soit un maximum local de f, soit donc J l'intervalle ouvert de la définition contenant x_0 tel que pour tout $x \in I \cap J$ on a $f(x) \leqslant f(x_0)$.

- Pour $x \in I \cap J$ tel que $x < x_0$ on a $f(x) - f(x_0) \leqslant 0$ et $x - x_0 < 0$ donc $\frac{f(x)-f(x_0)}{x-x_0} \geqslant 0$ et donc à la limite $\lim_{x \to x_0^-} \frac{f(x)-f(x_0)}{x-x_0} \geqslant 0$.
- Pour $x \in I \cap J$ tel que $x > x_0$ on a $f(x) - f(x_0) \leqslant 0$ et $x - x_0 > 0$ donc $\frac{f(x)-f(x_0)}{x-x_0} \leqslant 0$ et donc à la limite $\lim_{x \to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} \leqslant 0$.

Or f est dérivable en x_0 donc

$$\lim_{x \to x_0^-} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{x \to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0).$$

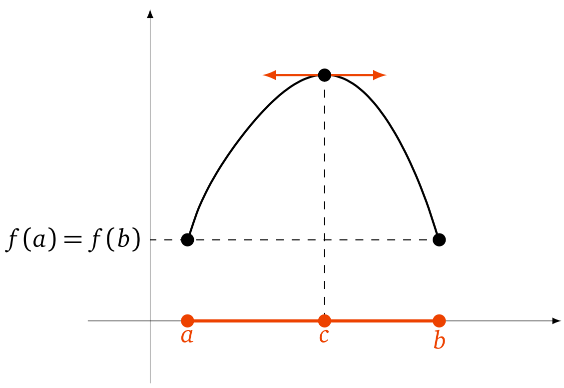
La première limite est positive, la seconde est négative, la seule possibilité est que $f'(x_0) = 0$. □

3.2. Théorème de Rolle

Théorème 3 (Théorème de Rolle).
Soit $f : [a,b] \to \mathbb{R}$ telle que

- *f est continue sur $[a,b]$,*
- *f est dérivable sur $]a,b[$,*
- *$f(a) = f(b)$.*

Alors il existe $c \in]a,b[$ tel que $f'(c) = 0$.



Interprétation géométrique : il existe au moins un point du graphe de f où la tangente est horizontale.

Démonstration. Tout d'abord, si f est constante sur $[a,b]$ alors n'importe quel $c \in]a,b[$ convient. Sinon il existe $x_0 \in [a,b]$ tel que $f(x_0) \neq f(a)$. Supposons par exemple $f(x_0) > f(a)$. Alors f est continue sur l'intervalle fermé et borné $[a,b]$, donc elle admet un maximum en un point $c \in [a,b]$. Mais $f(c) \geqslant f(x_0) > f(a)$ donc $c \neq a$. De même comme $f(a) = f(b)$ alors $c \neq b$. Ainsi $c \in]a,b[$. En c, f est donc dérivable et admet un maximum (local) donc $f'(c) = 0$. □

Exemple 8.
Soit $P(X) = (X-\alpha_1)(X-\alpha_2)\cdots(X-\alpha_n)$ un polynôme ayant n racines réelles différentes : $\alpha_1 < \alpha_2 < \cdots < \alpha_n$.

1. *Montrons que P' a $n-1$ racines distinctes.*

   On considère P comme une fonction polynomiale $x \mapsto P(x)$. P est une fonction continue et dérivable sur $\mathbb{R}$. Comme $P(\alpha_1) = 0 = P(\alpha_2)$ alors par le théorème de Rolle il existe $c_1 \in]\alpha_1, \alpha_2[$ tel que $P'(c_1) = 0$. Plus généralement, pour $1 \leqslant k \leqslant n-1$, comme $P(\alpha_k) = 0 = P(\alpha_{k+1})$ alors le théorème de Rolle implique l'existence de $c_k \in]\alpha_k, \alpha_{k+1}[$ tel que $P'(c_k) = 0$. Nous avons bien trouvé $n-1$ racines de P' : $c_1 < c_2 < \cdots < c_{n-1}$. Comme P' est un polynôme de degré $n-1$, toutes ses racines sont réelles et distinctes.

2. *Montrons que $P + P'$ a $n-1$ racines distinctes.*

   L'astuce consiste à considérer la fonction auxiliaire $f(x) = P(x)\exp x$. f est une fonction continue et dérivable sur $\mathbb{R}$. f s'annule comme P en $\alpha_1, \ldots, \alpha_n$. La dérivée de f est $f'(x) = \big(P(x) + P'(x)\big)\exp x$. Donc par le théorème de Rolle, pour chaque $1 \leqslant k \leqslant n-1$, comme $f(\alpha_k) = 0 = f(\alpha_{k+1})$ alors il existe $\gamma_k \in]\alpha_k, \alpha_{k+1}[$ tel que $f'(\gamma_k) = 0$. Mais comme la fonction exponentielle ne s'annule jamais alors $(P+P')(\gamma_k) = 0$. Nous avons bien trouvé $n-1$ racines distinctes de $P+P'$: $\gamma_1 < \gamma_2 < \cdots < \gamma_{n-1}$.

3. *Déduisons-en que $P + P'$ a toutes ses racines réelles.*

   $P + P'$ est un polynôme à coefficients réels qui admet $n-1$ racines réelles. Donc $(P+P')(X) = (X-\gamma_1)\cdots(X-\gamma_{n-1})Q(X)$ où $Q(x) = X - \gamma_n$ est un polynôme de degré 1. Comme $P + P'$ est à coefficients réels et que les γ_i sont aussi réels, ainsi $\gamma_n \in \mathbb{R}$. Ainsi on a obtenu une n-ème racine réelle γ_n (pas nécessairement distincte des autres γ_i).

Mini-exercices.

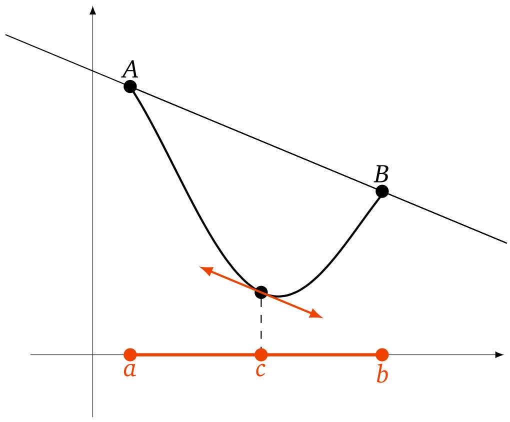
1. Dessiner le graphe de fonctions vérifiant : f_1 admet deux minimums locaux et un maximum local ; f_2 admet un minimum local qui n'est pas global et un maximum local qui est global ; f_3 admet une infinité d'extremums locaux ; f_4 n'admet aucun extremum local.
2. Calculer en quel point la fonction $f(x) = ax^2 + bx + c$ admet un extremum local.
3. Soit $f : [0,2] \to \mathbb{R}$ une fonction deux fois dérivable telle que $f(0) = f(1) = f(2) = 0$. Montrer qu'il existe c_1, c_2 tels que $f'(c_1) = 0$ et $f'(c_2) = 0$. Montrer qu'il existe c_3 tel que $f''(c_3) = 0$.
4. Montrer que chacune des trois hypothèses du théorème de Rolle est nécessaire.

4. Théorème des accroissements finis

4.1. Théorème des accroissements finis

Théorème 4 (Théorème des accroissements finis)**.**
Soit $f : [a,b] \to \mathbb{R}$ une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$. Il existe $c \in]a,b[$ tel que

$$\boxed{f(b) - f(a) = f'(c)\,(b-a)}$$



Interprétation géométrique : il existe au moins un point du graphe de f où la tangente est parallèle à la droite (AB) où $A = (a, f(a))$ et $B = (b, f(b))$.

Démonstration. Posons $\ell = \frac{f(b)-f(a)}{b-a}$ et $g(x) = f(x) - \ell \cdot (x-a)$. Alors $g(a) = f(a)$, $g(b) = f(b) - \frac{f(b)-f(a)}{b-a} \cdot (b-a) = f(a)$. Par le théorème de Rolle, il existe $c \in]a,b[$ tel que $g'(c) = 0$. Or $g'(x) = f'(x) - \ell$. Ce qui donne $f'(c) = \frac{f(b)-f(a)}{b-a}$. □

4.2. Fonction croissante et dérivée

Corollaire 2.
Soit $f : [a,b] \to \mathbb{R}$ une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$.

1. $\forall x \in]a,b[\quad f'(x) \geqslant 0 \iff f$ *est croissante ;*
2. $\forall x \in]a,b[\quad f'(x) \leqslant 0 \iff f$ *est décroissante ;*
3. $\forall x \in]a,b[\quad f'(x) = 0 \iff f$ *est constante ;*
4. $\forall x \in]a,b[\quad f'(x) > 0 \implies f$ *est strictement croissante ;*
5. $\forall x \in]a,b[\quad f'(x) < 0 \implies f$ *est strictement décroissante.*

Remarque.
La réciproque au point (4) (et aussi au (5)) est fausse. Par exemple la fonction $x \mapsto x^3$ est strictement croissante et pourtant sa dérivée s'annule en 0.

Démonstration. Prouvons par exemple (1).
Sens $\implies$. Supposons d'abord la dérivée positive. Soient $x, y \in]a,b[$ avec $x \leqslant y$. Alors par le théorème des accroissements finis, il existe $c \in]x,y[$ tel que $f(x) - f(y) = f'(c)(x-y)$. Mais $f'(c) \geqslant 0$ et $x - y \leqslant 0$ donc $f(x) - f(y) \leqslant 0$. Cela implique que $f(x) \leqslant f(y)$. Ceci étant vrai pour tout x, y alors f est croissante.

Sens $\impliedby$. Réciproquement, supposons que f est croissante. Fixons $x \in]a,b[$. Pour tout $y > x$ nous avons $y - x > 0$ et $f(y) - f(x) \geqslant 0$, ainsi le taux d'accroissement vérifie $\frac{f(y)-f(x)}{y-x} \geqslant 0$. À la limite, quand $y \to x$, ce taux d'accroissement tend vers la dérivée de f en x et donc $f'(x) \geqslant 0$. □

4.3. Inégalité des accroissements finis

Corollaire 3 (Inégalité des accroissements finis)**.**
Soit $f : I \to \mathbb{R}$ une fonction dérivable sur un intervalle I ouvert. S'il existe une constante M telle que pour tout $x \in I$, $\left|f'(x)\right| \leqslant M$ alors

$$\boxed{\forall x, y \in I \qquad \left|f(x) - f(y)\right| \leqslant M|x-y|}$$

Démonstration. Fixons $x, y \in I$, il existe alors $c \in]x, y[$ ou $]y, x[$ tel que $f(x) - f(y) = f'(c)(x - y)$ et comme $|f'(c)| \leqslant M$ alors $\left|f(x) - f(y)\right| \leqslant M|x - y|$. □

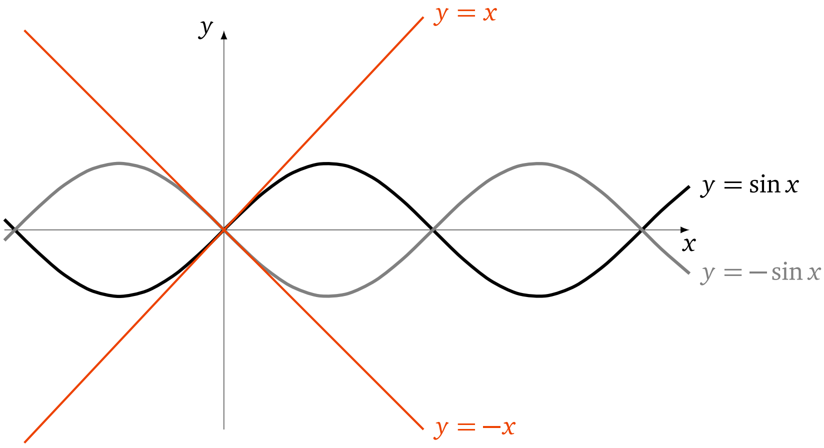
Exemple 9.
Soit $f(x) = \sin(x)$. Comme $f'(x) = \cos x$ alors $|f'(x)| \leqslant 1$ pour tout $x \in \mathbb{R}$. L'inégalité des accroissements finis s'écrit alors :

$$\text{pour tout } x, y \in \mathbb{R} \qquad |\sin x - \sin y| \leqslant |x - y|.$$

En particulier si l'on fixe $y = 0$ alors on obtient

$$\boxed{|\sin x| \leqslant |x|}$$

ce qui est particulièrement intéressant pour x proche de 0.



4.4. Règle de l'Hospital

Corollaire 4 (Règle de l'Hospital).
Soient $f, g : I \to \mathbb{R}$ deux fonctions dérivables et soit $x_0 \in I$. On suppose que
- $f(x_0) = g(x_0) = 0$,
- $\forall x \in I \setminus \{x_0\} \quad g'(x) \neq 0$.

$$\boxed{\textit{Si} \quad \lim_{x \to x_0} \frac{f'(x)}{g'(x)} = \ell \quad (\in \mathbb{R}) \quad \textit{alors} \quad \lim_{x \to x_0} \frac{f(x)}{g(x)} = \ell.}$$

Démonstration. Fixons $a \in I \setminus \{x_0\}$ avec par exemple $a < x_0$. Soit $h : I \to \mathbb{R}$ définie par $h(x) = g(a)f(x) - f(a)g(x)$. Alors
- h est continue sur $[a, x_0] \subset I$,
- h est dérivable sur $]a, x_0[$,
- $h(x_0) = h(a) = 0$.

Donc par le théorème de Rolle il existe $c_a \in]a, x_0[$ tel que $h'(c_a) = 0$. Or $h'(x) = g(a)f'(x) - f(a)g'(x)$ donc $g(a)f'(c_a) - f(a)g'(c_a) = 0$. Comme g' ne s'annule pas sur $I \setminus \{x_0\}$ cela conduit à $\frac{f(a)}{g(a)} = \frac{f'(c_a)}{g'(c_a)}$. Comme $a < c_a < x_0$ lorsque l'on fait tendre a vers x_0 on obtient $c_a \to x_0$. Cela implique

$$\lim_{a \to x_0} \frac{f(a)}{g(a)} = \lim_{a \to x_0} \frac{f'(c_a)}{g'(c_a)} = \lim_{c_a \to x_0} \frac{f'(c_a)}{g'(c_a)} = \ell.$$

□

Exemple 10.
Calculer la limite en 1 de $\frac{\ln(x^2+x-1)}{\ln(x)}$. On vérifie que :
- $f(x) = \ln(x^2 + x - 1)$, $f(1) = 0$, $f'(x) = \frac{2x+1}{x^2+x-1}$,
- $g(x) = \ln(x)$, $g(1) = 0$, $g'(x) = \frac{1}{x}$,
- Prenons $I =]0, 1]$, $x_0 = 1$, alors g' ne s'annule pas sur $I \setminus \{x_0\}$.

$$\frac{f'(x)}{g'(x)} = \frac{2x+1}{x^2+x-1} \times x = \frac{2x^2+x}{x^2+x-1} \xrightarrow[x\to 1]{} 3.$$

Donc

$$\frac{f(x)}{g(x)} \xrightarrow[x\to 1]{} 3.$$

Mini-exercices.

1. Soit $f(x) = \frac{x^3}{3} + \frac{x^2}{2} - 2x + 2$. Étudier la fonction f. Tracer son graphe. Montrer que f admet un minimum local et un maximum local.
2. Soit $f(x) = \sqrt{x}$. Appliquer le théorème des accroissements finis sur l'intervalle $[100, 101]$. En déduire l'encadrement $10 + \frac{1}{22} \leqslant \sqrt{101} \leqslant 10 + \frac{1}{20}$.
3. Appliquer le théorème des accroissements finis pour montrer que $\ln(1 + x) - \ln(x) < \frac{1}{x}$ (pour tout $x > 0$).
4. Soit $f(x) = e^x$. Que donne l'inégalité des accroissements finis sur $[0, x]$?
5. Appliquer la règle de l'Hospital pour calculer les limites suivantes (quand $x \to 0$) : $\dfrac{x}{(1+x)^n - 1}$; $\dfrac{\ln(x+1)}{\sqrt{x}}$; $\dfrac{1-\cos x}{\tan x}$; $\dfrac{x - \sin x}{x^3}$.

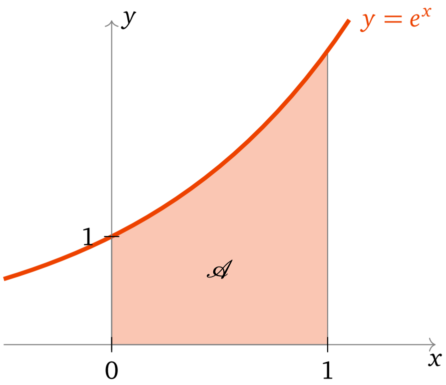
Auteurs du chapitre Arnaud Bodin, Niels Borne, Laura Desideri

Chapitre 5

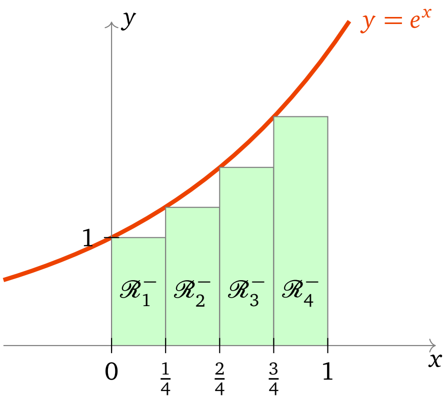
Intégrales

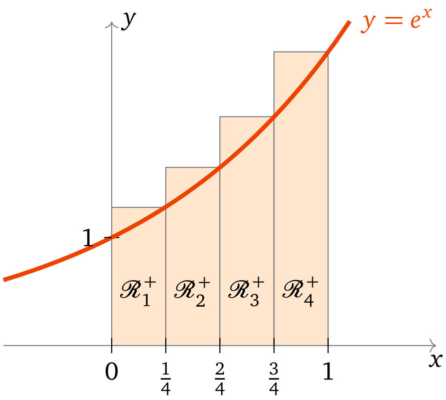
Objectif

Nous allons introduire l'intégrale à l'aide d'un exemple. Considérons la fonction exponentielle $f(x) = e^x$. On souhaite calculer l'aire $\mathscr{A}$ en-dessous du graphe de f et entre les droites d'équation $(x = 0)$, $(x = 1)$ et l'axe (Ox).



Nous approchons cette aire par des sommes d'aires des rectangles situés sous la courbe. Plus précisément, soit $n \geqslant 1$ un entier ; découpons notre intervalle $[0,1]$ à l'aide de la subdivision $(0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{i}{n}, \cdots, \frac{n-1}{n}, 1)$. On considère les « rectangles inférieurs » $\mathscr{R}_i^-$, chacun ayant pour base l'intervalle $\left[\frac{i-1}{n}, \frac{i}{n}\right]$ et pour hauteur $f\left(\frac{i-1}{n}\right) = e^{(i-1)/n}$. L'entier i varie de 1 à n. L'aire de $\mathscr{R}_i^-$ est « base × hauteur » : $\left(\frac{i}{n} - \frac{i-1}{n}\right) \times e^{(i-1)/n} = \frac{1}{n} e^{\frac{i-1}{n}}$.

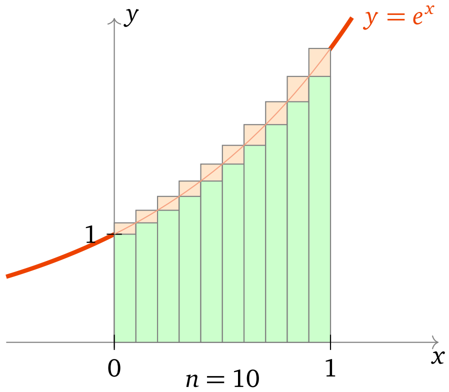




La somme des aires des $\mathscr{R}_i^-$ se calcule alors comme somme d'une suite géométrique :

$$\sum_{i=1}^{n} \frac{e^{\frac{i-1}{n}}}{n} = \frac{1}{n} \sum_{i=1}^{n} \left(e^{\frac{1}{n}}\right)^{i-1} = \frac{1}{n} \frac{1 - \left(e^{\frac{1}{n}}\right)^n}{1 - e^{\frac{1}{n}}} = \frac{\frac{1}{n}}{e^{\frac{1}{n}} - 1}(e-1) \xrightarrow[n \to +\infty]{} e - 1.$$

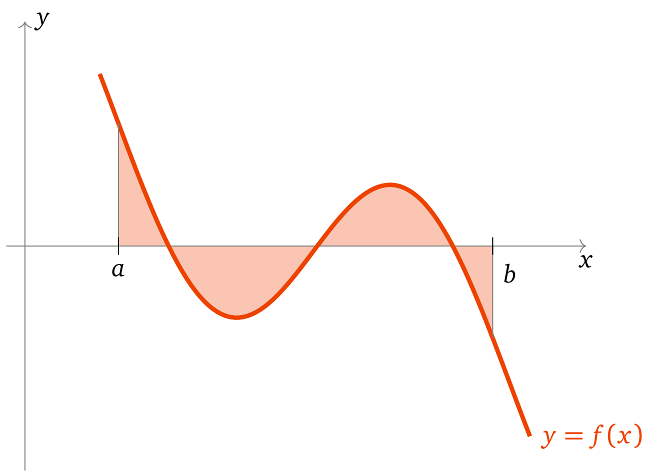
Pour la limite on a reconnu l'expression du type $\frac{e^x-1}{x} \xrightarrow[x\to 0]{} 1$ (avec ici $x = \frac{1}{n}$).
Soit maintenant les « rectangles supérieurs » $\mathscr{R}_i^+$, ayant la même base $\left[\frac{i-1}{n}, \frac{i}{n}\right]$ mais la hauteur $f\left(\frac{i}{n}\right) = e^{i/n}$.
Un calcul similaire montre que $\sum_{i=1}^{n} \frac{e^{\frac{i}{n}}}{n} \to e-1$ lorsque $n \to +\infty$.
L'aire $\mathscr{A}$ de notre région est supérieure à la somme des aires des rectangles inférieurs ; et elle est inférieure à la somme des aires des rectangles supérieurs. Lorsque l'on considère des subdivisions de plus en plus petites (c'est-à-dire lorsque l'on fait tendre n vers $+\infty$) alors on obtient à la limite que l'aire $\mathscr{A}$ de notre région est encadrée par deux aires qui tendent vers $e-1$. Donc l'aire de notre région est $\mathscr{A} = e-1$.



Voici le plan de lecture conseillé pour ce chapitre : il est tout d'abord nécessaire de bien comprendre comment est définie l'intégrale et quelles sont ses principales propriétés (parties 1 et 2). Mais il est important d'arriver rapidement à savoir calculer des intégrales : à l'aide de primitives ou par les deux outils efficaces que sont l'intégration par parties et le changement de variable.
Dans un premier temps on peut lire les sections 1.1, 1.2 puis 2.1, 2.2, 2.3, avant de s'attarder longuement sur les parties 3, 4. Lors d'une seconde lecture, revenez sur la construction de l'intégrale et les preuves.
Dans ce chapitre on s'autorisera (abusivement) une confusion entre une fonction f et son expression $f(x)$. Par exemple on écrira « *une primitive de la fonction* $\sin x$ *est* $-\cos x$ » au lieu « *une primitive de la fonction* $x \mapsto \sin x$ *est* $x \mapsto -\cos x$ ».

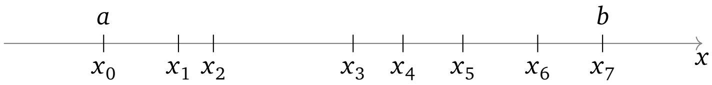
1. L'intégrale de Riemann

Nous allons reprendre la construction faite dans l'introduction pour une fonction f quelconque. Ce qui va remplacer les rectangles seront des ***fonctions en escalier***. Si la limite des aires en-dessous égale la limite des aires au-dessus on appelle cette limite commune ***l'intégrale*** de f que l'on note $\int_a^b f(x)\,dx$. Cependant il n'est pas toujours vrai que ces limites soient égales, l'intégrale n'est donc définie que pour les fonctions ***intégrables***. Heureusement nous verrons que si la fonction f est continue alors elle est intégrable.



1.1. Intégrale d'une fonction en escalier

Définition 1.
Soit $[a, b]$ un intervalle fermé borné de $\mathbb{R}$ ($-\infty < a < b < +\infty$). On appelle une ***subdivision*** de $[a, b]$ une suite finie, strictement croissante, de nombres $\mathscr{S} = (x_0, x_1, \ldots, x_n)$ telle que $x_0 = a$ et $x_n = b$. Autrement dit $a = x_0 < x_1 < \cdots < x_n = b$.

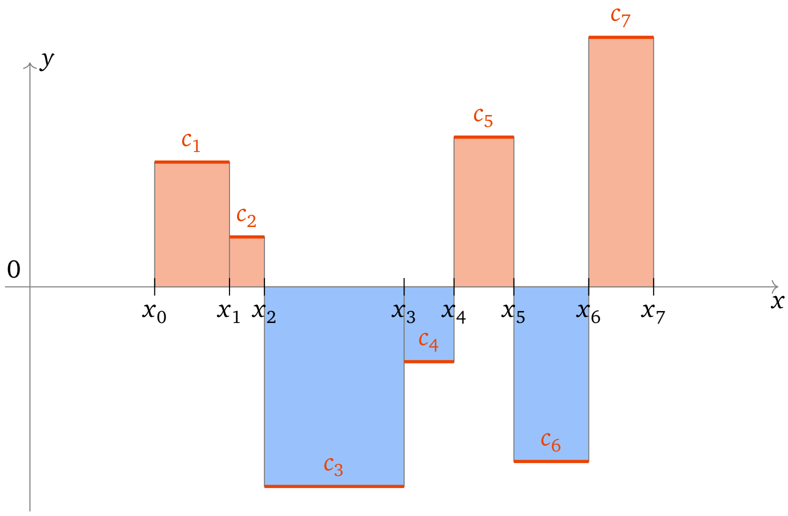


Définition 2.
Une fonction $f : [a, b] \to \mathbb{R}$ est une ***fonction en escalier*** s'il existe une subdivision $(x_0, x_1, \ldots, x_n)$ et des nombres réels $c_1, \ldots, c_n$ tels que pour tout $i \in \{1, \ldots, n\}$ on ait

$$\forall x \in]x_{i-1}, x_i[\quad f(x) = c_i$$

Autrement dit f est une fonction constante sur chacun des sous-intervalles de la subdivision.

Remarque.
La valeur de f aux points x_i de la subdivision n'est pas imposée. Elle peut être égale à celle de l'intervalle qui précède ou de celui qui suit, ou encore une autre valeur arbitraire. Cela n'a pas d'importance car l'aire ne changera pas.



Définition 3.
Pour une fonction en escalier comme ci-dessus, son ***intégrale*** est le réel $\int_a^b f(x)\,dx$ défini par

$$\boxed{\int_a^b f(x)\,dx = \sum_{i=1}^{n} c_i(x_i - x_{i-1})}$$

Remarque.
Notez que chaque terme $c_i(x_i - x_{i-1})$ est l'aire du rectangle compris entre les abscisses x_{i-1} et x_i et de hauteur c_i. Il faut juste prendre garde que l'on compte l'aire avec un signe « + » si $c_i > 0$ et un signe « − » si $c_i < 0$.
L'intégrale d'une fonction en escalier est l'aire de la partie située au-dessus de l'axe des abscisses (ici en rouge) moins l'aire de la partie située en-dessous (en bleu). L'intégrale d'une fonction en escalier est bien un nombre réel qui mesure l'aire algébrique (c'est-à-dire avec signe) entre la courbe de f et l'axe des abscisses.

1.2. Fonction intégrable

Rappelons qu'une fonction $f : [a,b] \to \mathbb{R}$ est ***bornée*** s'il existe $M \geqslant 0$ tel que :

$$\forall x \in [a,b] \quad -M \leqslant f(x) \leqslant M.$$

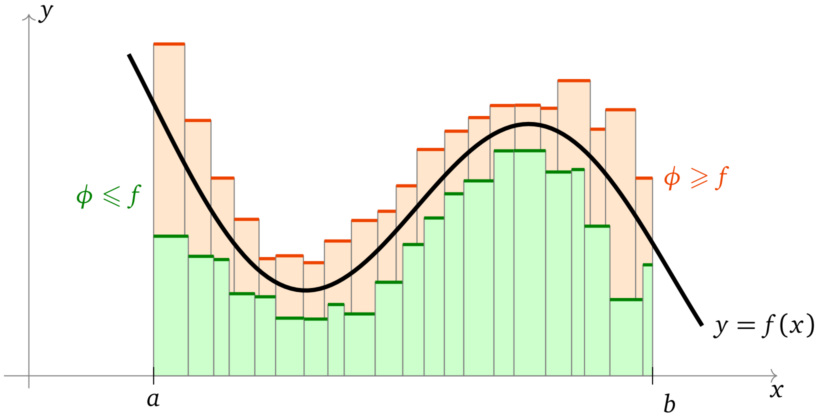
Rappelons aussi que si l'on a deux fonctions $f, g : [a,b] \to \mathbb{R}$, alors on note

$$f \leqslant g \qquad \Longleftrightarrow \qquad \forall x \in [a,b] \quad f(x) \leqslant g(x).$$

On suppose à présent que $f : [a,b] \to \mathbb{R}$ est une fonction bornée quelconque. On définit deux nombres réels :

$$I^-(f) = \sup\left\{ \int_a^b \phi(x)\,dx \mid \phi \text{ en escalier et } \phi \leqslant f \right\}$$

$$I^+(f) = \inf\left\{ \int_a^b \phi(x)\,dx \mid \phi \text{ en escalier et } \phi \geqslant f \right\}$$



Pour $I^-(f)$ on prend toutes les fonctions en escalier (avec toutes les subdivisions possibles) qui restent inférieures à f. On prend l'aire la plus grande parmi toutes ces fonctions en escalier, comme on n'est pas sûr que ce maximum existe on prend la borne supérieure. Pour $I^+(f)$ c'est le même principe mais les fonctions en escalier sont supérieures à f et on cherche l'aire la plus petite possible.

Il est intuitif que l'on a :

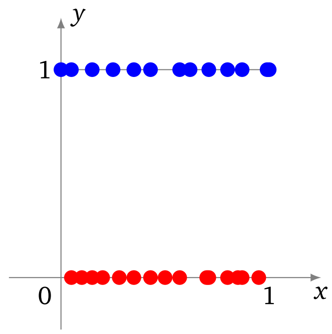
Proposition 1.
$I^-(f) \leqslant I^+(f)$.

Les preuves sont reportées en fin de section.

Définition 4.
Une fonction bornée $f : [a,b] \to \mathbb{R}$ est dite ***intégrable*** (***au sens de Riemann***) si $I^-(f) = I^+(f)$. On appelle alors ce nombre ***l'intégrale de Riemann*** de f sur $[a,b]$ et on le note $\int_a^b f(x)\,dx$.

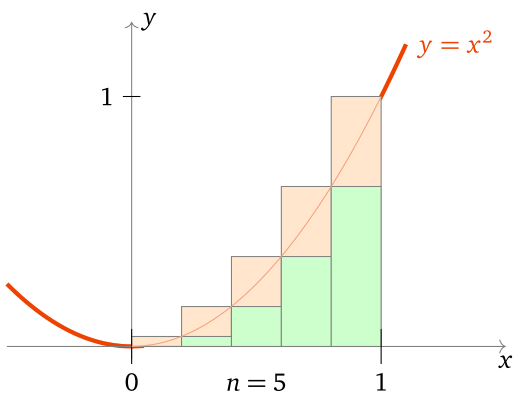
Exemple 1.

- Les fonctions en escalier sont intégrables ! En effet si f est une fonction en escalier alors la borne inférieure $I^-(f)$ et supérieure $I^+(f)$ sont atteintes avec la fonction $\phi = f$. Bien sûr l'intégrale $\int_a^b f(x)\,dx$ coïncide avec l'intégrale de la fonction en escalier définie lors du paragraphe 1.1.
- Nous verrons dans la section suivante que les fonctions continues et les fonctions monotones sont intégrables.
- Cependant toutes les fonctions ne sont pas intégrables. La fonction $f : [0,1] \to \mathbb{R}$ définie par $f(x) = 1$ si x est rationnel et $f(x) = 0$ sinon, n'est pas intégrable sur $[0,1]$. Convainquez-vous que si ϕ est une fonction en escalier avec $\phi \leqslant f$ alors $\phi \leqslant 0$ et que si $\phi \geqslant f$ alors $\phi \geqslant 1$. On en déduit que $I^-(f) = 0$ et $I^+(f) = 1$. Les bornes inférieure et supérieure ne coïncident pas, donc f n'est pas intégrable.



Il n'est pas si facile de calculer des exemples avec la définition. Nous avons vu l'exemple de la fonction exponentielle dans l'introduction où nous avions en fait montré que $\int_0^1 e^x\,dx = e-1$. Nous allons voir maintenant l'exemple de la fonction $f(x)=x^2$. Plus tard nous verrons que les primitives permettent de calculer simplement beaucoup d'intégrales.

Exemple 2.
Soit $f:[0,1]\to\mathbb{R}$, $f(x)=x^2$. Montrons qu'elle est intégrable et calculons $\int_0^1 f(x)\,dx$.



Soit $n\geqslant 1$ et considérons la subdivision régulière de $[0,1]$ suivante $\mathscr{S}=\left(0,\frac{1}{n},\frac{2}{n},\ldots,\frac{i}{n},\ldots,\frac{n-1}{n},1\right)$.
Sur l'intervalle $\left[\frac{i-1}{n},\frac{i}{n}\right]$ nous avons

$$\forall x\in\left[\tfrac{i-1}{n},\tfrac{i}{n}\right]\quad \left(\tfrac{i-1}{n}\right)^2\leqslant x^2\leqslant\left(\tfrac{i}{n}\right)^2.$$

Nous construisons une fonction en escalier ϕ^- en-dessous de f par $\phi^-(x)=\frac{(i-1)^2}{n^2}$ si $x\in\left[\frac{i-1}{n},\frac{i}{n}\right[$ (pour chaque $i=1,\ldots,n$) et $\phi^-(1)=1$. De même nous construisons une fonction en escalier ϕ^+ au-dessus de f définie par $\phi^+(x)=\frac{i^2}{n^2}$ si $x\in\left[\frac{i-1}{n},\frac{i}{n}\right[$ (pour chaque $i=1,\ldots,n$) et $\phi^+(1)=1$. ϕ^- et ϕ^+ sont des fonctions en escalier et l'on a $\phi^-\leqslant f\leqslant\phi^+$.
L'intégrale de la fonction en escalier ϕ^+ est par définition

$$\int_0^1\phi^+(x)\,dx=\sum_{i=1}^n\frac{i^2}{n^2}\left(\frac{i}{n}-\frac{i-1}{n}\right)=\sum_{i=1}^n\frac{i^2}{n^2}\frac{1}{n}=\frac{1}{n^3}\sum_{i=1}^n i^2.$$

On se souvient de la formule $\sum_{i=1}^n i^2=\frac{n(n+1)(2n+1)}{6}$, et donc

$$\int_0^1\phi^+(x)\,dx=\frac{n(n+1)(2n+1)}{6n^3}=\frac{(n+1)(2n+1)}{6n^2}.$$

De même pour la fonction ϕ^- :

$$\int_0^1\phi^-(x)\,dx=\sum_{i=1}^n\frac{(i-1)^2}{n^2}\frac{1}{n}=\frac{1}{n^3}\sum_{j=1}^{n-1}j^2=\frac{(n-1)n(2n-1)}{6n^3}=\frac{(n-1)(2n-1)}{6n^2}.$$

Maintenant $I^-(f)$ est la borne supérieure sur toutes les fonctions en escalier inférieures à f donc en particulier $I^-(f) \geqslant \int_0^1 \phi^-(x)\,dx$. De même $I^+(f) \leqslant \int_0^1 \phi^+(x)\,dx$. En résumé :

$$\frac{(n-1)(2n-1)}{6n^2} = \int_0^1 \phi^-(x)\,dx \leqslant I^-(f) \leqslant I^+(f) \leqslant \int_0^1 \phi^+(x)\,dx = \frac{(n+1)(2n+1)}{6n^2}.$$

Lorsque l'on fait tendre n vers $+\infty$ alors les deux extrémités tendent vers $\frac{1}{3}$. On en déduit que $I^-(f) = I^+(f) = \frac{1}{3}$. Ainsi f est intégrable et $\int_0^1 x^2\,dx = \frac{1}{3}$.

1.3. Premières propriétés

Proposition 2.

1. *Si $f : [a,b] \to \mathbb{R}$ est intégrable et si l'on change les valeurs de f en un nombre fini de points de $[a,b]$ alors la fonction f est toujours intégrable et la valeur de l'intégrale $\int_a^b f(x)\,dx$ ne change pas.*
2. *Si $f : [a,b] \to \mathbb{R}$ est intégrable alors la restriction de f à tout intervalle $[a',b'] \subset [a,b]$ est encore intégrable.*

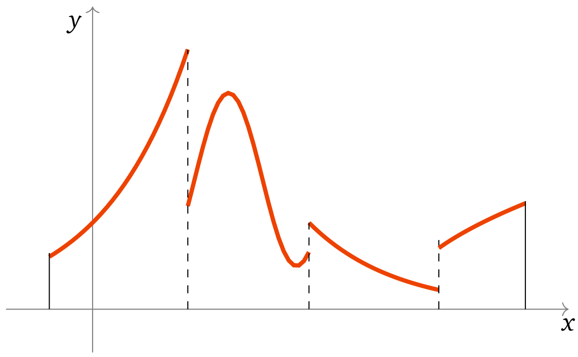
1.4. Les fonctions continues sont intégrables

Voici le résultat théorique le plus important de ce chapitre.

Théorème 1.
Si $f : [a,b] \to \mathbb{R}$ est continue alors f est intégrable.

La preuve sera vue plus loin mais l'idée est que les fonctions continues peuvent être approchées d'aussi près que l'on veut par des fonctions en escalier, tout en gardant un contrôle d'erreur uniforme sur l'intervalle.

Une fonction $f : [a,b] \to \mathbb{R}$ est dite ***continue par morceaux*** s'il existe un entier n et une subdivision $(x_0, \dots, x_n)$ telle que $f_{|]x_{i-1},x_i[}$ soit continue, admette une limite finie à droite en x_{i-1} et une limite à gauche en x_i pour tout $i \in \{1, \dots, n\}$.



Corollaire 1.
Les fonctions continues par morceaux sont intégrables.

Voici un résultat qui prouve que l'on peut aussi intégrer des fonctions qui ne sont pas continues à condition que la fonction soit croissante (ou décroissante).

Théorème 2.
Si $f : [a,b] \to \mathbb{R}$ est monotone alors f est intégrable.

1.5. Les preuves

Les preuves peuvent être sautées lors d'une première lecture. Les démonstrations demandent une bonne maîtrise des bornes sup et inf et donc des « epsilons ». La proposition 1 se prouve en manipulant les « epsilons ». Pour la preuve de la proposition 2 : on prouve d'abord les propriétés pour les fonctions en escalier et on en déduit qu'elles restent vraies pour les fonctions intégrables (cette technique sera développée en détails dans la partie suivante).

Le théorème 1 établit que les fonctions continues sont intégrables. Nous allons démontrer une version affaiblie de ce résultat. Rappelons que f est dite de ***classe*** $\mathscr{C}^1$ si f est continue, dérivable et f' est aussi continue.

> **Théorème 3** (Théorème 1 faible).
> *Si $f : [a,b] \to \mathbb{R}$ est de classe $\mathscr{C}^1$ alors f est intégrable.*

Démonstration. Comme f est de classe $\mathscr{C}^1$ alors f' est une fonction continue sur l'intervalle fermé et borné $[a,b]$; f' est donc une fonction bornée : il existe $M \geqslant 0$ tel que pour tout $x \in [a,b]$ on ait $|f'(x)| \leqslant M$.
Nous allons utiliser l'inégalité des accroissements finis :

$$\forall x,y \in [a,b] \quad |f(x)-f(y)| \leqslant M|x-y|. \qquad (\star)$$

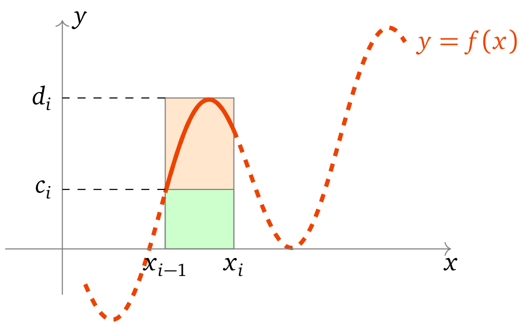
Soit $\epsilon > 0$ et soit $(x_0, x_1, \ldots, x_n)$ une subdivision de $[a,b]$ vérifiant pour tout $i = 1, \ldots, n$:

$$0 < x_i - x_{i-1} \leqslant \epsilon. \qquad (\star\star)$$

Nous allons construire deux fonctions en escalier $\phi^-, \phi^+ : [a,b] \to \mathbb{R}$ définies de la façon suivante : pour chaque $i = 1, \ldots, n$ et chaque $x \in [x_{i-1}, x_i[$ on pose

$$c_i = \phi^-(x) = \inf_{t \in [x_{i-1}, x_i[} f(t) \quad \text{et} \quad d_i = \phi^+(x) = \sup_{t \in [x_{i-1}, x_i[} f(t)$$

et aussi $\phi^-(b) = \phi^+(b) = f(b)$. ϕ^- et ϕ^+ sont bien deux fonctions en escalier (elles sont constantes sur chaque intervalle $[x_{i-1}, x_i[$).



De plus par construction on a bien $\phi^- \leqslant f \leqslant \phi^+$ et donc

$$\int_a^b \phi^-(x)\,dx \leqslant I^-(f) \leqslant I^+(f) \leqslant \int_a^b \phi^+(x)\,dx\,.$$

En utilisant la continuité de f sur l'intervalle $[x_{i-1}, x_i]$, on déduit l'existence de $a_i, b_i \in [x_{i-1}, x_i]$ tels que $f(a_i) = c_i$ et $f(b_i) = d_i$. Avec $(\star)$ et $(\star\star)$ on sait que $d_i - c_i = f(b_i) - f(a_i) \leqslant M|b_i - a_i| \leqslant M(x_i - x_{i-1}) \leqslant M\epsilon$ (pour tout $i = 1, \ldots, n$). Alors

$$\int_a^b \phi^+(x)\,dx - \int_a^b \phi^-(x)\,dx \leqslant \sum_{i=1}^n M\epsilon(x_i - x_{i-1}) = M\epsilon(b-a)$$

Ainsi $0 \leqslant I^+(f) - I^-(f) \leqslant M\epsilon(b-a)$ et lorsque l'on fait tendre $\epsilon \to 0$ on trouve $I^+(f) = I^-(f)$, ce qui prouve que f est intégrable. □

La preuve du théorème 2 est du même style et nous l'omettons.

Mini-exercices.

1. Soit $f : [1,4] \to \mathbb{R}$ définie par $f(x) = 1$ si $x \in [1,2[$, $f(x) = 3$ si $x \in [2,3[$ et $f(x) = -1$ si $x \in [3,4]$. Calculer $\int_1^2 f(x)\,dx$, $\int_1^3 f(x)\,dx$, $\int_1^4 f(x)\,dx$, $\int_1^{\frac{3}{2}} f(x)\,dx$, $\int_{\frac{3}{2}}^{\frac{7}{2}} f(x)\,dx$.
2. Montrer que $\int_0^1 x\,dx = \frac{1}{2}$ (prendre une subdivision régulière et utiliser $\sum_{i=1}^n i = \frac{n(n+1)}{2}$).
3. Montrer que si f est une fonction intégrable et *paire* sur l'intervalle $[-a,a]$ alors $\int_{-a}^a f(x)\,dx = 2\int_0^a f(x)\,dx$ (on prendra une subdivision symétrique par rapport à l'origine).
4. Montrer que si f est une fonction intégrable et *impaire* sur l'intervalle $[-a,a]$ alors $\int_{-a}^a f(x)\,dx = 0$.
5. Montrer que toute fonction monotone est intégrable en s'inspirant de la preuve du théorème 3.

2. Propriétés de l'intégrale

Les trois principales propriétés de l'intégrale sont la relation de Chasles, la positivité et la linéarité.

2.1. Relation de Chasles

Proposition 3 (Relation de Chasles)**.**
Soient $a < c < b$. Si f est intégrable sur $[a,c]$ et $[c,b]$, alors f est intégrable sur $[a,b]$. Et on a

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

Pour s'autoriser des bornes sans se préoccuper de l'ordre on définit :

$$\int_a^a f(x)\,dx = 0 \qquad \text{et pour } a < b \quad \int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.$$

Pour a, b, c quelconques la ***relation de Chasles*** devient alors

$$\boxed{\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx}$$

2.2. Positivité de l'intégrale

Proposition 4 (Positivité de l'intégrale)**.**
Soit $a \leqslant b$ deux réels et f et g deux fonctions intégrables sur $[a,b]$.

$$\boxed{\text{Si } f \leqslant g \quad \text{alors} \quad \int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx}$$

En particulier l'intégrale d'une fonction positive est positive :

$$\boxed{\text{Si} \quad f \geqslant 0 \quad \text{alors} \quad \int_a^b f(x)\,dx \geqslant 0}$$

2.3. Linéarité de l'intégrale

Proposition 5.
Soient f, g deux fonctions intégrables sur $[a, b]$.

1. *$f+g$ est une fonction intégrable et $\int_a^b (f+g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$.*
2. *Pour tout réel λ, λf est intégrable et on a $\int_a^b \lambda f(x)\,dx = \lambda \int_a^b f(x)\,dx$.*
   Par ces deux premiers points nous avons la **linéarité de l'intégrale** *: pour tous réels λ, μ*
   $$\boxed{\int_a^b \big(\lambda f(x) + \mu g(x)\big)\,dx = \lambda \int_a^b f(x)\,dx + \mu \int_a^b g(x)\,dx}$$
3. *$f \times g$ est une fonction intégrable sur $[a, b]$ mais en général $\int_a^b (fg)(x)\,dx \neq \big(\int_a^b f(x)\,dx\big)\big(\int_a^b g(x)\,dx\big)$.*
4. *$|f|$ est une fonction intégrable sur $[a, b]$ et*
   $$\boxed{\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b \big|f(x)\big|\,dx}$$

Exemple 3.

$$\int_0^1 \big(7x^2 - e^x\big)\,dx = 7\int_0^1 x^2\,dx \;-\; \int_0^1 e^x\,dx = 7\frac{1}{3} \;-\; (e-1) = \frac{10}{3} - e$$

Nous avons utilisé les calculs déjà vus : $\int_0^1 x^2\,dx = \frac{1}{3}$ et $\int_0^1 e^x\,dx = e-1$.

Exemple 4.
Soit $I_n = \int_1^n \frac{\sin(nx)}{1+x^n}\,dx$. Montrons que $I_n \to 0$ lorsque $n \to +\infty$.

$$|I_n| = \left|\int_1^n \frac{\sin(nx)}{1+x^n}\,dx\right| \leqslant \int_1^n \frac{|\sin(nx)|}{1+x^n}\,dx \leqslant \int_1^n \frac{1}{1+x^n}\,dx \leqslant \int_1^n \frac{1}{x^n}\,dx$$

Il ne reste plus qu'à calculer cette dernière intégrale (en anticipant un peu sur la suite du chapitre) :

$$\int_1^n \frac{1}{x^n}\,dx = \int_1^n x^{-n}\,dx = \left[\frac{x^{-n+1}}{-n+1}\right]_1^n = \frac{n^{-n+1}}{-n+1} - \frac{1}{-n+1} \xrightarrow[n\to+\infty]{} 0$$

(car $n^{-n+1} \to 0$ et $\frac{1}{-n+1} \to 0$).

Remarque.
Notez que même si $f \times g$ est intégrable on a en général $\int_a^b (fg)(x)\,dx \neq \big(\int_a^b f(x)\,dx\big)\big(\int_a^b g(x)\,dx\big)$. Par exemple, soit $f : [0,1] \to \mathbb{R}$ la fonction définie par $f(x) = 1$ si $x \in [0, \frac{1}{2}[$ et $f(x) = 0$ sinon. Soit $g : [0,1] \to \mathbb{R}$ la fonction définie par $g(x) = 1$ si $x \in [\frac{1}{2}, 1[$ et $g(x) = 0$ sinon. Alors $f(x) \times g(x) = 0$ pour tout $x \in [0,1]$ et donc $\int_0^1 f(x)g(x)\,dx = 0$ alors que $\int_0^1 f(x)\,dx = \frac{1}{2}$ et $\int_0^1 g(x)\,dx = \frac{1}{2}$.

2.4. Une preuve

Nous allons prouver la linéarité de l'intégrale : $\int \lambda f = \lambda \int f$ et $\int f + g = \int f + \int g$. L'idée est la suivante : il est facile de voir que pour des fonctions en escalier l'intégrale (qui est alors une somme finie) est linéaire. Comme les fonctions en escalier approchent autant qu'on le souhaite les fonctions intégrables alors cela implique la linéarité de l'intégrale.

Preuve de $\int \lambda f = \lambda \int f$. Soit $f : [a, b] \to \mathbb{R}$ une fonction intégrable et $\lambda \in \mathbb{R}$. Soit $\epsilon > 0$. Il existe ϕ^- et ϕ^+

deux fonctions en escalier approchant suffisamment f, avec $\phi^- \leqslant f \leqslant \phi^+$:

$$\int_a^b f(x)\,dx \; -\epsilon \;\leqslant \int_a^b \phi^-(x)\,dx \quad \text{et} \quad \int_a^b \phi^+(x)\,dx \;\leqslant \int_a^b f(x)\,dx \; +\epsilon \qquad (\dagger)$$

Quitte à rajouter des points, on peut supposer que la subdivision $(x_0, x_1, \ldots, x_n)$ de $[a, b]$ est suffisamment fine pour que ϕ^- et ϕ^+ soient toutes les deux constantes sur les intervalles $]x_{i-1}, x_i[$; on note c_i^- et c_i^+ leurs valeurs respectives.

Dans un premier temps on suppose $\lambda \geqslant 0$. Alors $\lambda\phi^-$ et $\lambda\phi^+$ sont encore des fonctions en escalier vérifiant $\lambda\phi^- \leqslant \lambda f \leqslant \lambda\phi^+$. De plus

$$\int_a^b \lambda\phi^-(x)\,dx = \sum_{i=1}^n \lambda c_i^-(x_i - x_{i-1}) = \lambda \sum_{i=1}^n c_i^-(x_i - x_{i-1}) = \lambda \int_a^b \phi^-(x)\,dx$$

De même pour ϕ^+. Ainsi

$$\lambda \int_a^b \phi^-(x)\,dx \leqslant I^-(\lambda f) \leqslant I^+(\lambda f) \leqslant \lambda \int_a^b \phi^+(x)\,dx$$

En utilisant les deux inégalités ($\dagger$) on obtient

$$\lambda \int_a^b f(x)\,dx \; - \lambda\epsilon \; \leqslant I^-(\lambda f) \leqslant I^+(\lambda f) \leqslant \lambda \int_a^b f(x)\,dx \; + \lambda\epsilon$$

Lorsque l'on fait tendre $\epsilon \to 0$ cela prouve que $I^-(\lambda f) = I^+(\lambda f)$, donc λf est intégrable et $\int_a^b \lambda f(x)\,dx = \lambda \int_a^b f(x)\,dx$. Si $\lambda \leqslant 0$ on a $\lambda\phi^+ \leqslant \lambda f \leqslant \lambda\phi^-$ et le raisonnement est similaire. □

Preuve de $\int f + g = \int f + \int g$. Soit $\epsilon > 0$. Soient $f, g : [a, b] \to \mathbb{R}$ deux fonctions intégrables. On définit deux fonctions en escalier ϕ^+, ϕ^- pour f et deux fonctions en escalier ψ^+, ψ^- pour g vérifiant des inégalités similaires à ($\dagger$) de la preuve au-dessus. On fixe une subdivision suffisamment fine pour toutes les fonctions $\phi^\pm, \psi^\pm$. On note $c_i^\pm, d_i^\pm$ les constantes respectives sur l'intervalle $]x_{i-1}, x_i[$. Les fonctions $\phi^- + \psi^-$ et $\phi^+ + \psi^+$ sont en escalier et vérifient $\phi^- + \psi^- \leqslant f + g \leqslant \phi^+ + \psi^+$. Nous avons aussi que

$$\int_a^b (\phi^- + \psi^-)(x)\,dx = \sum_{i=1}^n (c_i^- + d_i^-)(x_i - x_{i-1}) = \int_a^b \phi^-(x)\,dx + \int_a^b \psi^-(x)\,dx$$

De même pour $\phi^+ + \psi^+$. Ainsi

$$\int_a^b \phi^-(x)\,dx + \int_a^b \psi^-(x)\,dx \leqslant I^-(f+g) \leqslant I^+(f+g) \leqslant \int_a^b \phi^+(x)\,dx + \int_a^b \psi^+(x)\,dx$$

Les conditions du type ($\dagger$) impliquent alors

$$\int_a^b f(x)\,dx + \int_a^b g(x)\,dx \; -2\epsilon \; \leqslant I^-(f+g) \leqslant I^+(f+g) \leqslant \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \; +2\epsilon$$

Lorsque $\epsilon \to 0$ on déduit $I^-(f+g) = I^+(f+g)$, donc $f+g$ est intégrable et $\int_a^b \big(f(x) + g(x)\big)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$. □

Mini-exercices.

1. En admettant que $\int_0^1 x^n\,dx = \frac{1}{n+1}$. Calculer l'intégrale $\int_0^1 P(x)\,dx$ où $P(x) = a_n x^n + \cdots + a_1 x + a_0$. Trouver un polynôme $P(x)$ non nul de degré 2 dont l'intégrale est nulle : $\int_0^1 P(x)\,dx = 0$.
2. A-t-on $\int_a^b f(x)^2\,dx = \left(\int_a^b f(x)\,dx\right)^2$; $\int_a^b \sqrt{f(x)}\,dx = \sqrt{\int_a^b f(x)\,dx}$; $\int_a^b |f(x)|\,dx = \left|\int_a^b f(x)\,dx\right|$; $\int |f(x) + g(x)|\,dx = \left|\int_a^b f(x)\,dx\right| + \left|\int_a^b g(x)\,dx\right|$?
3. Peut-on trouver $a < b$ tels que $\int_a^b x\,dx = -1$; $\int_a^b x\,dx = 0$; $\int_a^b x\,dx = +1$? Mêmes questions avec $\int_a^b x^2\,dx$.

4. Montrer que $0 \leqslant \int_1^2 \sin^2 x \, dx \leqslant 1$ et $\left|\int_a^b \cos^3 x \, dx\right| \leqslant |b-a|$.

3. Primitive d'une fonction

3.1. Définition

Définition 5.
Soit $f : I \to \mathbb{R}$ une fonction définie sur un intervalle I quelconque. On dit que $F : I \to \mathbb{R}$ est une ***primitive*** de f sur I si F est une fonction dérivable sur I vérifiant $F'(x) = f(x)$ pour tout $x \in I$.

Trouver une primitive est donc l'opération inverse de calculer la fonction dérivée.

Exemple 5.

1. Soit $I = \mathbb{R}$ et $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$. Alors $F : \mathbb{R} \to \mathbb{R}$ définie par $F(x) = \frac{x^3}{3}$ est une primitive de f. La fonction définie par $F(x) = \frac{x^3}{3} + 1$ est aussi une primitive de f.
2. Soit $I = [0, +\infty[$ et $g : I \to \mathbb{R}$ définie par $g(x) = \sqrt{x}$. Alors $G : I \to \mathbb{R}$ définie par $G(x) = \frac{2}{3}x^{\frac{3}{2}}$ est une primitive de g sur I. Pour tout $c \in \mathbb{R}$, la fonction $G + c$ est aussi une primitive de g.

Nous allons voir que trouver une primitive permet de les trouver toutes.

Proposition 6.
Soit $f : I \to \mathbb{R}$ une fonction et soit $F : I \to \mathbb{R}$ une primitive de f. Toute primitive de f s'écrit $G = F + c$ où $c \in \mathbb{R}$.

Démonstration. Notons tout d'abord que si l'on note G la fonction définie par $G(x) = F(x) + c$ alors $G'(x) = F'(x)$ mais comme $F'(x) = f(x)$ alors $G'(x) = f(x)$ et G est bien une primitive de f.
Pour la réciproque supposons que G soit une primitive quelconque de f. Alors $(G-F)'(x) = G'(x) - F'(x) = f(x) - f(x) = 0$, ainsi la fonction $G - F$ a une dérivée nulle sur un intervalle, c'est donc une fonction constante ! Il existe donc $c \in \mathbb{R}$ tel que $(G-F)(x) = c$. Autrement dit $G(x) = F(x) + c$ (pour tout $x \in I$). $\square$

Notations. On notera une primitive de f par $\int f(t)\,dt$ ou $\int f(x)\,dx$ ou $\int f(u)\,du$ (les lettres $t, x, u, \ldots$ sont des lettres dites *muettes*, c'est-à-dire interchangeables). On peut même noter une primitive simplement par $\int f$.
La proposition 6 nous dit que si F est une primitive de f alors il existe un réel c, tel que $F = \int f(t)\,dt + c$.
Attention : $\int f(t)\,dt$ désigne une fonction de I dans $\mathbb{R}$ alors que l'intégrale $\int_a^b f(t)\,dt$ désigne un nombre réel. Plus précisément nous verrons que si F est une primitive de f alors $\int_a^b f(t)\,dt = F(b) - F(a)$.

Par dérivation on prouve facilement le résultat suivant :

Proposition 7.
Soient F une primitive de f et G une primitive de g. Alors $F + G$ est une primitive de $f + g$. Et si $\lambda \in \mathbb{R}$ alors λF est une primitive de λf.

Une autre formulation est de dire que pour tous réels λ, μ on a

$$\boxed{\int \big(\lambda f(t) + \mu g(t)\big)\,dt = \lambda \int f(t)\,dt + \mu \int g(t)\,dt}$$

3.2. Primitives des fonctions usuelles

$\int e^x\,dx = e^x + c \quad$ sur $\mathbb{R}$
$\int \cos x\,dx = \sin x + c \quad$ sur $\mathbb{R}$
$\int \sin x\,dx = -\cos x + c \quad$ sur $\mathbb{R}$
$\int x^n\,dx = \frac{x^{n+1}}{n+1} + c \quad (n \in \mathbb{N}) \quad$ sur $\mathbb{R}$
$\int x^\alpha\,dx = \frac{x^{\alpha+1}}{\alpha+1} + c \quad (\alpha \in \mathbb{R} \setminus \{-1\})$ sur $]0,+\infty[$
$\int \frac{1}{x}\,dx = \ln\lvert x\rvert + c \quad$ sur $]0,+\infty[$ ou $]-\infty,0[$
$\int \operatorname{sh} x\,dx = \operatorname{ch} x + c$, $\int \operatorname{ch} x\,dx = \operatorname{sh} x + c \quad$ sur $\mathbb{R}$
$\int \frac{dx}{1+x^2} = \arctan x + c \quad$ sur $\mathbb{R}$
$\int \frac{dx}{\sqrt{1-x^2}} = \begin{cases} \arcsin x + c \\ \frac{\pi}{2} - \arccos x + c \end{cases} \quad$ sur $]-1,1[$
$\int \frac{dx}{\sqrt{x^2+1}} = \begin{cases} \operatorname{Argsh} x + c \\ \ln\left(x + \sqrt{x^2+1}\right) + c \end{cases} \quad$ sur $\mathbb{R}$
$\int \frac{dx}{\sqrt{x^2-1}} = \begin{cases} \operatorname{Argch} x + c \\ \ln\left(x + \sqrt{x^2-1}\right) + c \end{cases} \quad$ sur $x \in]1,+\infty[$

Remarque.
Ces primitives sont à connaître par cœur.

1. Voici comment lire ce tableau. Si f est la fonction définie sur $\mathbb{R}$ par $f(x) = x^n$ alors la fonction : $x \mapsto \frac{x^{n+1}}{n+1}$ est une primitive de f sur $\mathbb{R}$. Les primitives de f sont les fonctions définies par $x \mapsto \frac{x^{n+1}}{n+1} + c$ (pour c une constante réelle quelconque). Et on écrit $\int x^n\,dx = \frac{x^{n+1}}{n+1} + c$, où $c \in \mathbb{R}$.
2. Souvenez vous que la variable sous le symbole intégrale est une variable muette. On peut aussi bien écrire $\int t^n\,dt = \frac{x^{n+1}}{n+1} + c$.
3. La constante est définie pour un intervalle. Si l'on a deux intervalles, il y a deux constantes qui peuvent être différentes. Par exemple pour $\int \frac{1}{x}\,dx$ nous avons deux domaines de validité : $I_1 =]0,+\infty[$ et $I_2 =]-\infty,0[$. Donc $\int \frac{1}{x}\,dx = \ln x + c_1$ si $x > 0$ et $\int \frac{1}{x}\,dx = \ln\lvert x\rvert + c_2 = \ln(-x) + c_2$ si $x < 0$.
4. On peut trouver des primitives aux allures très différentes par exemple $x \mapsto \arcsin x$ et $x \mapsto \frac{\pi}{2} - \arccos x$ sont deux primitives de la même fonction $x \mapsto \frac{1}{\sqrt{1-x^2}}$. Mais bien sûr on sait que $\arcsin x + \arccos x = \frac{\pi}{2}$, donc les primitives diffèrent bien d'une constante !

3.3. Relation primitive-intégrale

Théorème 4.
Soit $f : [a,b] \to \mathbb{R}$ une fonction continue. La fonction $F : I \to \mathbb{R}$ définie par

$$\boxed{F(x) = \int_a^x f(t)\,dt}$$

est une primitive de f, c'est-à-dire F est dérivable et $F'(x) = f(x)$.
Par conséquent pour une primitive F quelconque de f :

$$\boxed{\int_a^b f(t)\,dt = F(b) - F(a)}$$

Notation. On note $\left[F(x)\right]_a^b = F(b) - F(a)$.

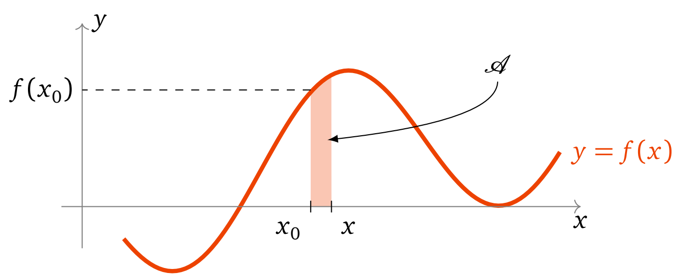
Exemple 6.
Nous allons pouvoir calculer plein d'intégrales. Recalculons d'abord les intégrales déjà rencontrées.

1. Pour $f(x) = e^x$ une primitive est $F(x) = e^x$ donc
$$\int_0^1 e^x\,dx = \left[e^x\right]_0^1 = e^1 - e^0 = e - 1.$$
2. Pour $g(x) = x^2$ une primitive est $G(x) = \frac{x^3}{3}$ donc
$$\int_0^1 x^2\,dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}.$$
3. $\int_a^x \cos t\,dt = \left[\sin t\right]_{t=a}^{t=x} = \sin x - \sin a$ est une primitive de $\cos x$.
4. Si f est impaire alors ses primitives sont paires (le montrer). En déduire que $\int_{-a}^a f(t)\,dt = 0$.

Remarque.

1. $F(x) = \int_a^x f(t)\,dt$ est même ***l'unique primitive de f qui s'annule en a***.
2. En particulier si F est une fonction de classe $\mathscr{C}^1$ alors $\boxed{\int_a^b F'(t)\,dt = F(b) - F(a)}$.
3. On évitera la notation $\int_a^x f(x)\,dx$ où la variable x est présente à la fois aux bornes et à l'intérieur de l'intégrale. Mieux vaut utiliser la notation $\int_a^x f(t)\,dt$ ou $\int_a^x f(u)\,du$ pour éviter toute confusion.
4. Une fonction intégrable n'admet pas forcément une primitive. Considérer par exemple $f : [0,1] \to \mathbb{R}$ définie par $f(x) = 0$ si $x \in [0, \frac{1}{2}[$ et $f(x) = 1$ si $x \in [\frac{1}{2}, 1]$. f est intégrable sur $[0,1]$ mais elle n'admet pas de primitive sur $[0,1]$. En effet par l'absurde si F était une primitive de f, par exemple la primitive qui vérifie $F(0) = 0$. Alors $F(x) = 0$ pour $x \in [0, \frac{1}{2}[$ et $F(x) = x - \frac{1}{2}$ pour $x \in [\frac{1}{2}, 1]$. Mais alors nous obtenons une contradiction car F n'est pas dérivable en $\frac{1}{2}$ alors que par définition une primitive doit être dérivable.

Démonstration. Essayons de visualiser tout d'abord pourquoi la fonction F est dérivable et $F'(x) = f(x)$.



Fixons $x_0 \in [a, b]$. Par la relation de Chasles nous savons :

$$F(x) - F(x_0) = \int_a^x f(t)\,dt - \int_a^{x_0} f(t)\,dt = \int_{x_0}^a f(t)\,dt + \int_a^x f(t)\,dt = \int_{x_0}^x f(t)\,dt$$

Donc le taux d'accroissement

$$\frac{F(x) - F(x_0)}{x - x_0} = \frac{1}{x - x_0}\int_{x_0}^x f(t)\,dt = \frac{\mathscr{A}}{x - x_0}$$

où $\mathscr{A}$ est l'aire hachurée (en rouge). Mais cette aire est presque un rectangle, si x est suffisamment proche de x_0, donc l'aire $\mathscr{A}$ vaut environ $(x - x_0) \times f(x_0)$; lorsque $x \to x_0$ le taux d'accroissement tend donc vers $f(x_0)$. Autrement dit $F'(x_0) = f(x_0)$.

Passons à la preuve rigoureuse. Comme $f(x_0)$ est une constante alors $\int_{x_0}^x f(x_0)\,dt = (x - x_0)f(x_0)$, donc

$$\begin{aligned} \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) &= \frac{1}{x - x_0}\int_{x_0}^x f(t)\,dt - \frac{1}{x - x_0}\int_{x_0}^x f(x_0)\,dt \\ &= \frac{1}{x - x_0}\int_{x_0}^x \big(f(t) - f(x_0)\big)\,dt \end{aligned}$$

Fixons $\epsilon > 0$. Puisque f est continue en x_0, il existe $\delta > 0$ tel que $(|t - x_0| < \delta \implies |f(t) - f(x_0)| < \epsilon)$. Donc :

$$\begin{aligned} \left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| &= \left|\frac{1}{x - x_0}\int_{x_0}^x \big(f(t) - f(x_0)\big)\,dt\right| \\ &\leqslant \frac{1}{|x - x_0|}\left|\int_{x_0}^x \big|f(t) - f(x_0)\big|\,dt\right| \\ &\leqslant \frac{1}{|x - x_0|}\left|\int_{x_0}^x \epsilon\,dt\right| = \epsilon \end{aligned}$$

Ce qui prouve que F est dérivable en x_0 et $F'(x_0) = f(x_0)$.

Maintenant on sait que F est une primitive de f, F est même la primitive qui s'annule en a car $F(a) = \int_a^a f(t)\,dt = 0$. Si G est une autre primitive on sait $F = G + c$. Ainsi

$$G(b) - G(a) = F(b) + c - \big(F(a) + c\big) = F(b) - F(a) = F(b) = \int_a^b f(t)\,dt.$$

□

3.4. Sommes de Riemann

L'intégrale est définie à partir de limites de sommes. Mais maintenant que nous savons calculer des intégrales sans utiliser ces sommes on peut faire le cheminement inverse : calculer des limites de sommes à partir d'intégrales.

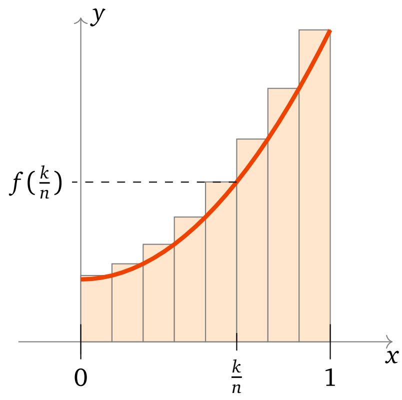
Théorème 5.
Soit $f : [a, b] \to \mathbb{R}$ une fonction intégrable, alors

$$\boxed{S_n = \frac{b-a}{n}\sum_{k=1}^n f\Big(a + k\frac{b-a}{n}\Big) \xrightarrow[n \to +\infty]{} \int_a^b f(x)\,dx}$$

La somme S_n s'appelle la *somme de Riemann* associée à l'intégrale et correspond à une subdivision régulière de l'intervalle $[a, b]$ en n petits intervalles. La hauteur de chaque rectangle étant évaluée à son extrémité droite.

Le cas le plus utile est le cas où $a = 0$, $b = 1$ alors $\frac{b-a}{n} = \frac{1}{n}$ et $f\left(a + k\frac{b-a}{n}\right) = f\left(\frac{k}{n}\right)$ et ainsi

$$S_n = \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) \xrightarrow[n\to+\infty]{} \int_0^1 f(x)\,dx$$



Exemple 7.
Calculer la limite de la somme $S_n = \sum_{k=1}^{n} \frac{1}{n+k}$.
On a $S_1 = \frac{1}{2}$, $S_2 = \frac{1}{3} + \frac{1}{4}$, $S_3 = \frac{1}{4} + \frac{1}{5} + \frac{1}{6}$, $S_4 = \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}, \ldots$
La somme S_n s'écrit aussi $S_n = \frac{1}{n}\sum_{k=1}^{n} \frac{1}{1+\frac{k}{n}}$. En posant $f(x) = \frac{1}{1+x}$, $a = 0$ et $b = 1$, on reconnaît que S_n est une somme de Riemann. Donc

$$S_n = \frac{1}{n}\sum_{k=1}^{n} \frac{1}{1+\frac{k}{n}} = \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)$$

$$\xrightarrow[n\to+\infty]{} \int_a^b f(x)\,dx = \int_0^1 \frac{1}{1+x}\,dx = \Big[\ln|1+x|\Big]_0^1 = \ln 2 - \ln 1 = \ln 2.$$

Ainsi $S_n \to \ln 2$ (lorsque $n \to +\infty$).

Mini-exercices.

1. Trouver les primitives des fonctions : $x^3 - x^7$, $\cos x + \exp x$, $\sin(2x)$, $1 + \sqrt{x} + x$, $\frac{1}{\sqrt{x}}$, $\sqrt[3]{x}$, $\frac{1}{x+1}$.
2. Trouver les primitives des fonctions : $\mathrm{ch}(x) - \mathrm{sh}(\frac{x}{2})$, $\frac{1}{1+4x^2}$, $\frac{1}{\sqrt{1+x^2}} - \frac{1}{\sqrt{1-x^2}}$.
3. Trouver une primitive de $x^2 e^x$ sous la forme $(ax^2 + bx + c)e^x$.
4. Trouver toutes les primitives de $x \mapsto \frac{1}{x^2}$ (préciser les intervalles et les constantes).
5. Calculer les intégrales $\int_0^1 x^n\,dx$, $\int_0^{\frac{\pi}{4}} \frac{dx}{1+x^2}$, $\int_1^e \frac{1-x}{x^2}\,dx$, $\int_0^{\frac{1}{2}} \frac{dx}{x^2-1}$.
6. Calculer la limite (lorsque $n \to +\infty$) de la somme $S_n = \sum_{k=1}^{n} \frac{e^{k/n}}{n}$. Idem avec $S'_n = \sum_{k=1}^{n} \frac{n}{(n+k)^2}$.

4. Intégration par parties – Changement de variable

Pour trouver une primitive d'une fonction f on peut avoir la chance de reconnaître que f est la dérivée d'une fonction bien connue. C'est malheureusement très rarement le cas, et on ne connaît pas les primitives de la plupart des fonctions. Cependant nous allons voir deux techniques qui permettent des calculer des intégrales et des primitives : l'intégration par parties et le changement de variable.

4.1. Intégration par parties

Théorème 6.
Soient u et v deux fonctions de classe $\mathscr{C}^1$ sur un intervalle $[a,b]$.

$$\boxed{\int_a^b u(x)\,v'(x)\,dx = \left[uv\right]_a^b - \int_a^b u'(x)\,v(x)\,dx}$$

Notation. Le crochet $\left[F\right]_a^b$ est par définition $\left[F\right]_a^b = F(b) - F(a)$. Donc $\left[uv\right]_a^b = u(b)v(b) - u(a)v(a)$. Si l'on omet les bornes alors $\left[F\right]$ désigne la fonction $F + c$ où c est une constante quelconque.
La formule d'intégration par parties pour les primitives est la même mais sans les bornes :

$$\int u(x)v'(x)\,dx = \left[uv\right] - \int u'(x)v(x)\,dx.$$

La preuve est très simple :

Démonstration. On a $(uv)' = u'v + uv'$. Donc $\int_a^b(u'v + uv') = \int_a^b(uv)' = \left[uv\right]_a^b$. D'où $\int_a^b uv' = \left[uv\right]_a^b - \int_a^b u'v$. □

L'utilisation de l'intégration par parties repose sur l'idée suivante : on ne sait pas calculer directement l'intégrale d'une fonction f s'écrivant comme un produit $f(x) = u(x)v'(x)$ mais si l'on sait calculer l'intégrale de $g(x) = u'(x)v(x)$ (que l'on espère plus simple) alors par la formule d'intégration par parties on aura l'intégrale de f.

Exemple 8.

1. Calcul de $\int_0^1 xe^x\,dx$. On pose $u(x) = x$ et $v'(x) = e^x$. Nous aurons besoin de savoir que $u'(x) = 1$ et qu'une primitive de v' est simplement $v(x) = e^x$. La formule d'intégration par parties donne :

$$\begin{aligned}
\int_0^1 xe^x\,dx &= \int_0^1 u(x)v'(x)\,dx\\
&= \left[u(x)v(x)\right]_0^1 - \int_0^1 u'(x)v(x)\,dx\\
&= \left[xe^x\right]_0^1 - \int_0^1 1\cdot e^x\,dx\\
&= \left(1\cdot e^1 - 0\cdot e^0\right) - \left[e^x\right]_0^1\\
&= e - (e^1 - e^0)\\
&= 1
\end{aligned}$$

2. Calcul de $\int_1^e x\ln x\,dx$.

   On pose cette fois $u = \ln x$ et $v' = x$. Ainsi $u' = \frac{1}{x}$ et $v = \frac{x^2}{2}$. Alors

$$\int_1^e \ln x\cdot x\,dx = \int_1^e uv' = \left[uv\right]_1^e - \int_1^e u'v = \left[\ln x\cdot\tfrac{x^2}{2}\right]_1^e - \int_1^e \tfrac{1}{x}\tfrac{x^2}{2}\,dx$$
$$= \left(\ln e\tfrac{e^2}{2} - \ln 1\tfrac{1^2}{2}\right) - \tfrac{1}{2}\int_1^e x\,dx = \tfrac{e^2}{2} - \frac{1}{2}\left[\tfrac{x^2}{2}\right]_1^e = \tfrac{e^2}{2} - \tfrac{e^2}{4} + \tfrac{1}{4} = \tfrac{e^2+1}{4}$$

3. Calcul de $\int \arcsin x\,dx$.

   Pour déterminer une primitive de $\arcsin x$, nous faisons artificiellement apparaître un produit en écrivant $\arcsin x = 1\cdot\arcsin x$ pour appliquer la formule d'intégration par parties. On pose $u = \arcsin x$, $v' = 1$

(et donc $u' = \frac{1}{\sqrt{1-x^2}}$ et $v = x$) alors

$$\begin{aligned}\int 1 \cdot \arcsin x \, dx &= \big[x \arcsin x\big] - \int \frac{x}{\sqrt{1-x^2}} \, dx \\ &= \big[x \arcsin x\big] - \big[-\sqrt{1-x^2}\big] \\ &= x \arcsin x + \sqrt{1-x^2} + c\end{aligned}$$

4. Calcul de $\int x^2 e^x \, dx$. On pose $u = x^2$ et $v' = e^x$ pour obtenir :

$$\int x^2 e^x \, dx = \big[x^2 e^x\big] - 2 \int x e^x \, dx$$

On refait une deuxième intégration par parties pour calculer

$$\int x e^x \, dx = \big[x e^x\big] - \int e^x \, dx = (x-1)e^x + c$$

D'où

$$\int x^2 e^x \, dx = (x^2 - 2x + 2)e^x + c.$$

Exemple 9.

Nous allons étudier les intégrales définies par $I_n = \displaystyle\int_0^1 \frac{\sin(\pi x)}{x+n} \, dx$, pour tout entier $n > 0$.

1. Montrer que $0 \leqslant I_{n+1} \leqslant I_n$.
   Pour $0 \leqslant x \leqslant 1$, on a $0 < x + n \leqslant x + n + 1$ et $\sin(\pi x) \geqslant 0$, donc $0 \leqslant \frac{\sin(\pi x)}{x+n+1} \leqslant \frac{\sin(\pi x)}{x+n}$. D'où $0 \leqslant I_{n+1} \leqslant I_n$ par la positivité de l'intégrale.
2. Montrer que $I_n \leqslant \ln \frac{n+1}{n}$. En déduire $\lim_{n\to+\infty} I_n$.
   De $0 \leqslant \sin(\pi x) \leqslant 1$, on a $\frac{\sin(\pi x)}{x+n} \leqslant \frac{1}{x+n}$. D'où $0 \leqslant I_n \leqslant \int_0^1 \frac{1}{x+n} \, dx = \big[\ln(x+n)\big]_0^1 = \ln \frac{n+1}{n} \to 0$.
3. Calculer $\lim_{n\to+\infty} nI_n$.
   Nous allons faire une intégration par parties avec $u = \frac{1}{x+n}$ et $v' = \sin(\pi x)$ (et donc $u' = -\frac{1}{(x+n)^2}$ et $v = -\frac{1}{\pi}\cos(\pi x)$) :

$$\begin{aligned}nI_n &= n \int_0^1 \frac{1}{x+n} \sin(\pi x) \, dx \\ &= -\frac{n}{\pi}\left[\frac{1}{x+n}\cos(\pi x)\right]_0^1 - \frac{n}{\pi}\int_0^1 \frac{1}{(x+n)^2}\cos(\pi x) \, dx \\ &= \frac{n}{\pi(n+1)} + \frac{1}{\pi} - \frac{n}{\pi} J_n\end{aligned}$$

   Il nous reste à évaluer $J_n = \int_0^1 \frac{\cos(\pi x)}{(x+n)^2} \, dx$.

$$\begin{aligned}\left|\frac{n}{\pi} J_n\right| \leqslant \frac{n}{\pi}\int_0^1 \frac{|\cos(\pi x)|}{(x+n)^2} \, dx \leqslant \frac{n}{\pi}\int_0^1 \frac{1}{(x+n)^2} \, dx \\ = \frac{n}{\pi}\left[-\frac{1}{x+n}\right]_0^1 = \frac{n}{\pi}\left(-\frac{1}{1+n} + \frac{1}{n}\right) = \frac{1}{\pi}\frac{1}{n+1} \to 0.\end{aligned}$$

   Donc $\lim_{n\to+\infty} nI_n = \lim_{n\to+\infty} \frac{n}{\pi(n+1)} + \frac{1}{\pi} - \frac{n}{\pi} J_n = \frac{2}{\pi}$.

4.2. Changement de variable

Théorème 7.

Soit f une fonction définie sur un intervalle I et $\varphi : J \to I$ une bijection de classe $\mathscr{C}^1$. Pour tout $a, b \in J$

$$\boxed{\int_{\varphi(a)}^{\varphi(b)} f(x)\,dx = \int_a^b f\big(\varphi(t)\big)\cdot\varphi'(t)\,dt}$$

Si F est une primitive de f alors $F\circ\varphi$ est une primitive de $\big(f\circ\varphi\big)\cdot\varphi'$.

Voici un moyen simple de s'en souvenir. En effet si l'on note $x=\varphi(t)$ alors par dérivation on obtient $\frac{dx}{dt}=\varphi'(t)$ donc $dx=\varphi'(t)\,dt$. D'où la substitution $\int_{\varphi(a)}^{\varphi(b)} f(x)\,dx=\int_a^b f(\varphi(t))\,\varphi'(t)\,dt$.

Démonstration. Comme F est une primitive de f alors $F'(x)=f(x)$ et par la formule de la dérivation de la composition $F\circ\varphi$ on a

$$(F\circ\varphi)'(t)=F'(\varphi(t))\varphi'(t)=f(\varphi(t))\varphi'(t).$$

Donc $F\circ\varphi$ est une primitive de $f(\varphi(t))\varphi'(t)$.

Pour les intégrales : $\displaystyle\int_a^b f(\varphi(t))\varphi'(t)\,dt=\big[F\circ\varphi\big]_a^b=F\big(\varphi(b)\big)-F\big(\varphi(a)\big)=\big[F\big]_{\varphi(a)}^{\varphi(b)}=\int_{\varphi(a)}^{\varphi(b)} f(x)\,dx.$ □

Remarque.
Comme φ est une bijection de J sur $\varphi(J)$, sa réciproque φ^{-1} existe et est dérivable sauf quand φ s'annule. Si φ ne s'annule pas, on peut écrire $t=\varphi^{-1}(x)$ et faire un changement de variable en sens inverse.

Exemple 10.
Calculons la primitive $F=\int \tan t\,dt$.

$$F=\int \tan t\,dt=\int \frac{\sin t}{\cos t}\,dt\ .$$

On reconnaît ici une forme $\frac{u'}{u}$ (avec $u=\cos t$ et $u'=-\sin t$) dont une primitive est $\ln|u|$. Donc $F=\int -\frac{u'}{u}=-\big[\ln|u|\big]=-\ln|u|+c=-\ln|\cos t|+c$.

Nous allons reformuler tout cela en terme de changement de variable. Notons $\varphi(t)=\cos t$ alors $\varphi'(t)=-\sin t$, donc

$$F=\int -\frac{\varphi'(t)}{\varphi(t)}\,dt$$

Si f désigne la fonction définie par $f(x)=\frac{1}{x}$, qui est bijective tant que $x\neq 0$; alors $F=-\int \varphi'(t)f(\varphi(t))\,dt$. En posant $x=\varphi(t)$ et donc $dx=\varphi'(t)dt$, on reconnaît la formule du changement de variable, par conséquent

$$F\circ\varphi^{-1}=-\int f(x)\,dx=-\int \frac{1}{x}\,dx=-\ln|x|+c\ .$$

Comme $x=\varphi(t)=\cos t$, on retrouve bien $F(t)=-\ln|\cos t|+c$.
Remarque : pour que l'intégrale soit bien définie il faut que $\tan t$ soit définie, donc $t\not\equiv\frac{\pi}{2}\mod\pi$. La restriction d'une primitive à un intervalle $]-\frac{\pi}{2}+k\pi,\frac{\pi}{2}+k\pi[$ est donc de la forme $-\ln|\cos t|+c$. Mais la constante c peut être différente sur un intervalle différent.

Exemple 11.
Calcul de $\int_0^{1/2}\frac{x}{(1-x^2)^{3/2}}\,dx$.
Soit le changement de variable $u=\varphi(x)=1-x^2$. Alors $du=\varphi'(x)\,dx=-2x\,dx$. Pour $x=0$ on a $u=\varphi(0)=1$ et pour $x=\frac12$ on a $u=\varphi(\frac12)=\frac34$. Comme $\varphi'(x)=-2x$, φ est une bijection de $[0,\frac12]$ sur $[1,\frac34]$. Alors

$$\begin{aligned}\int_0^{1/2}\frac{x\,dx}{(1-x^2)^{3/2}} &= \int_1^{3/4}\frac{\frac{-du}{2}}{u^{3/2}}=-\frac12\int_1^{3/4}u^{-3/2}\,du\\ &= -\frac12\big[-2u^{-1/2}\big]_1^{3/4}=\big[\frac{1}{\sqrt{u}}\big]_1^{3/4}=\frac{1}{\sqrt{\frac34}}-1=\frac{2}{\sqrt3}-1.\end{aligned}$$

Exemple 12.
Calcul de $\int_0^{1/2} \frac{1}{(1-x^2)^{3/2}}\,dx$.
On effectue le changement de variable $x = \varphi(t) = \sin t$ et $dx = \cos t\,dt$. De plus $t = \arcsin x$ donc pour $x = 0$ on a $t = \arcsin(0) = 0$ et pour $x = \frac{1}{2}$ on a $t = \arcsin(\frac{1}{2}) = \frac{\pi}{6}$. Comme φ est une bijection de $[0, \frac{\pi}{6}]$ sur $[0, \frac{1}{2}]$,

$$\begin{aligned}\int_0^{1/2} \frac{dx}{(1-x^2)^{3/2}} &= \int_0^{\pi/6} \frac{\cos t\,dt}{(1-\sin^2 t)^{3/2}} = \int_0^{\pi/6} \frac{\cos t\,dt}{(\cos^2 t)^{3/2}} \\ &= \int_0^{\pi/6} \frac{\cos t}{\cos^3 t}\,dt = \int_0^{\pi/6} \frac{1}{\cos^2 t}\,dt = \big[\tan t\big]_0^{\pi/6} = \frac{1}{\sqrt{3}}\,.\end{aligned}$$

Exemple 13.
Calcul de $\int \frac{1}{(1+x^2)^{3/2}}\,dx$.
Soit le changement de variable $x = \tan t$ donc $t = \arctan x$ et $dx = \frac{dt}{\cos^2 t}$ (la fonction tangente établit une bijection de $]-\frac{\pi}{2}, +\frac{\pi}{2}[$ sur $\mathbb{R}$). Donc

$$\begin{aligned}F = \int \frac{1}{(1+x^2)^{3/2}}\,dx &= \int \frac{1}{(1+\tan^2 t)^{3/2}} \frac{dt}{\cos^2 t} \\ &= \int (\cos^2 t)^{3/2} \frac{dt}{\cos^2 t} \qquad \text{car } 1+\tan^2 t = \frac{1}{\cos^2 t} \\ &= \int \cos t\,dt = \big[\sin t\big] = \sin t + c = \sin(\arctan x) + c\end{aligned}$$

Donc

$$\int \frac{1}{(1+x^2)^{3/2}}\,dx = \sin(\arctan x) + c.$$

En manipulant un peu les fonctions on trouverait que la primitive s'écrit aussi $F(x) = \frac{x}{\sqrt{1+x^2}} + c$.

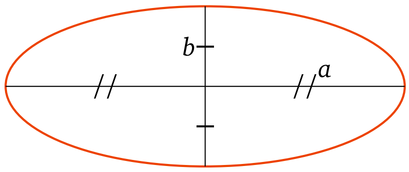
Mini-exercices.

1. Calculer les intégrales à l'aide d'intégrations par parties : $\int_0^{\pi/2} t \sin t\,dt$, $\int_0^{\pi/2} t^2 \sin t\,dt$, puis par récurrence $\int_0^{\pi/2} t^n \sin t\,dt$.
2. Déterminer les primitives à l'aide d'intégrations par parties : $\int t \operatorname{sh} t\,dt$, $\int t^2 \operatorname{sh} t\,dt$, puis par récurrence $\int t^n \operatorname{sh} t\,dt$.
3. Calculer les intégrales à l'aide de changements de variable : $\int_0^a \sqrt{a^2-t^2}\,dt$; $\int_{-\pi}^{\pi} \sqrt{1+\cos t}\,dt$ (pour ce dernier poser deux changements de variables : $u = \cos t$, puis $v = 1-u$).
4. Déterminer les primitives suivantes à l'aide de changements de variable : $\int \operatorname{th} t\,dt$ où $\operatorname{th} t = \frac{\operatorname{sh} t}{\operatorname{ch} t}$, $\int e^{\sqrt{t}}\,dt$.

5. Intégration des fractions rationnelles

Nous savons intégrer beaucoup de fonctions simples. Par exemple toutes les fonctions polynomiales : si $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ alors $\int f(x)\,dx = a_0 x + a_1 \frac{x^2}{2} + a_2 \frac{x^3}{3} + \cdots + a_n \frac{x^{n+1}}{n+1} + c$.

Il faut être conscient cependant que beaucoup de fonctions ne s'intègrent pas à l'aide de fonctions simples. Par exemple si $f(t) = \sqrt{a^2 \cos^2 t + b^2 \sin^2 t}$ alors l'intégrale $\int_0^{2\pi} f(t)\,dt$ ne peut pas s'exprimer comme somme, produit, inverse ou composition de fonctions que vous connaissez. En fait cette intégrale vaut la longueur d'une ellipse d'équation paramétrique $(a \cos t, b \sin t)$; il n'y a donc pas de formule pour le périmètre d'une ellipse (sauf si $a = b$ auquel cas l'ellipse est un cercle !).



Mais de façon remarquable, il y a toute une famille de fonctions que l'on saura intégrer : les fractions rationnelles.

5.1. Trois situations de base

On souhaite d'abord intégrer les fractions rationnelles $f(x) = \frac{\alpha x+\beta}{ax^2+bx+c}$ avec $\alpha, \beta, a, b, c \in \mathbb{R}$, $a \neq 0$ et $(\alpha, \beta) \neq (0,0)$.

Premier cas. Le dénominateur $ax^2 + bx + c$ possède deux racines réelles distinctes $x_1, x_2 \in \mathbb{R}$.
Alors $f(x)$ s'écrit aussi $f(x) = \frac{\alpha x+\beta}{a(x-x_1)(x-x_2)}$ et il existe des nombres $A, B \in \mathbb{R}$ tels que $f(x) = \frac{A}{x-x_1} + \frac{B}{x-x_2}$.
On a donc

$$\int f(x)\,dx = A\ln|x - x_1| + B\ln|x - x_2| + c$$

sur chacun des intervalles $]-\infty, x_1[$, $]x_1, x_2[$, $]x_2, +\infty[$ (si $x_1 < x_2$).

Deuxième cas. Le dénominateur $ax^2 + bx + c$ possède une racine double $x_0 \in \mathbb{R}$.
Alors $f(x) = \frac{\alpha x+\beta}{a(x-x_0)^2}$ et il existe des nombres $A, B \in \mathbb{R}$ tels que $f(x) = \frac{A}{(x-x_0)^2} + \frac{B}{x-x_0}$. On a alors

$$\int f(x)\,dx = -\frac{A}{x - x_0} + B\ln|x - x_0| + c$$

sur chacun des intervalles $]-\infty, x_0[$, $]x_0, +\infty[$.

Troisième cas. Le dénominateur $ax^2 + bx + c$ ne possède pas de racine réelle. Voyons comment faire sur un exemple.

Exemple 14.
Soit $f(x) = \frac{x+1}{2x^2+x+1}$. Dans un premier temps on fait apparaître une fraction du type $\frac{u'}{u}$ (que l'on sait intégrer en $\ln|u|$).

$$f(x) = \frac{(4x+1)\frac{1}{4} - \frac{1}{4} + 1}{2x^2 + x + 1} = \frac{1}{4} \cdot \frac{4x+1}{2x^2+x+1} + \frac{3}{4} \cdot \frac{1}{2x^2+x+1}$$

On peut intégrer la fraction $\frac{4x+1}{2x^2+x+1}$:

$$\int \frac{4x+1}{2x^2+x+1}\,dx = \int \frac{u'(x)}{u(x)}\,dx = \ln\left|2x^2 + x + 1\right| + c$$

Occupons nous de l'autre partie $\frac{1}{2x^2+x+1}$, nous allons l'écrire sous la forme $\frac{1}{u^2+1}$ (dont une primitive est $\arctan u$).

$$\begin{aligned}
\frac{1}{2x^2+x+1} &= \frac{1}{2(x+\frac{1}{4})^2 - \frac{1}{8} + 1} = \frac{1}{2(x+\frac{1}{4})^2 + \frac{7}{8}} \\
&= \frac{8}{7} \cdot \frac{1}{\frac{8}{7}2(x+\frac{1}{4})^2 + 1} = \frac{8}{7} \cdot \frac{1}{\left(\frac{4}{\sqrt{7}}(x+\frac{1}{4})\right)^2 + 1}
\end{aligned}$$

On pose le changement de variable $u = \frac{4}{\sqrt{7}}(x + \frac{1}{4})$ (et donc $du = \frac{4}{\sqrt{7}}dx$) pour trouver

$$\begin{aligned}
\int \frac{dx}{2x^2+x+1} &= \int \frac{8}{7} \frac{dx}{\left(\frac{4}{\sqrt{7}}(x+\frac{1}{4})\right)^2 + 1} = \frac{8}{7} \int \frac{du}{u^2+1} \cdot \frac{\sqrt{7}}{4} \\
&= \frac{2}{\sqrt{7}} \arctan u + c = \frac{2}{\sqrt{7}} \arctan\left(\frac{4}{\sqrt{7}}\left(x + \frac{1}{4}\right)\right) + c\,.
\end{aligned}$$

Finalement :

$$\int f(x)\,dx = \frac{1}{4}\ln\left(2x^2+x+1\right)+\frac{3}{2\sqrt{7}}\arctan\left(\frac{4}{\sqrt{7}}\left(x+\frac{1}{4}\right)\right)+c$$

5.2. Intégration des éléments simples

Soit $\frac{P(x)}{Q(x)}$ une fraction rationnelle, où $P(x), Q(x)$ sont des polynômes à coefficients réels. Alors la fraction $\frac{P(x)}{Q(x)}$ s'écrit comme somme d'un polynôme $E(x) \in \mathbb{R}[x]$ (la partie entière) et d'éléments simples d'une des formes suivantes :

$$\frac{\gamma}{(x-x_0)^k} \quad \text{ou} \quad \frac{\alpha x+\beta}{(ax^2+bx+c)^k} \text{ avec } b^2-4ac<0$$

où $\alpha,\beta,\gamma,a,b,c \in \mathbb{R}$ et $k \in \mathbb{N}\setminus\{0\}$.

1. On sait intégrer le polynôme $E(x)$.
2. Intégration de l'élément simple $\frac{\gamma}{(x-x_0)^k}$.
   (a) Si $k=1$ alors $\int \frac{\gamma\,dx}{x-x_0} = \gamma\ln|x-x_0|+c_0$ (sur $]-\infty,x_0[$ ou $]x_0,+\infty[$).
   (b) Si $k \geqslant 2$ alors $\int \frac{\gamma\,dx}{(x-x_0)^k} = \gamma\int (x-x_0)^{-k}\,dx = \frac{\gamma}{-k+1}(x-x_0)^{-k+1}+c_0$ (sur $]-\infty,x_0[$ ou $]x_0,+\infty[$).
3. Intégration de l'élément simple $\frac{\alpha x+\beta}{(ax^2+bx+c)^k}$. On écrit cette fraction sous la forme

$$\frac{\alpha x+\beta}{(ax^2+bx+c)^k} = \gamma\frac{2ax+b}{(ax^2+bx+c)^k}+\delta\frac{1}{(ax^2+bx+c)^k}$$

   (a) Si $k=1$, calcul de $\int \frac{2ax+b}{ax^2+bx+c}\,dx = \int \frac{u'(x)}{u(x)}\,dx = \ln|u(x)|+c_0 = \ln|ax^2+bx+c|+c_0$.
   (b) Si $k \geqslant 2$, calcul de $\int \frac{2ax+b}{(ax^2+bx+c)^k}\,dx = \int \frac{u'(x)}{u(x)^k}\,dx = \frac{1}{-k+1}u(x)^{-k+1}+c_0 = \frac{1}{-k+1}(ax^2+bx+c)^{-k+1}+c_0$.
   (c) Si $k=1$, calcul de $\int \frac{1}{ax^2+bx+c}\,dx$. Par un changement de variable $u=px+q$ on se ramène à calculer une primitive du type $\int \frac{du}{u^2+1} = \arctan u + c_0$.
   (d) Si $k \geqslant 2$, calcul de $\int \frac{1}{(ax^2+bx+c)^k}\,dx$. On effectue le changement de variable $u=px+q$ pour se ramener au calcul de $I_k = \int \frac{du}{(u^2+1)^k}$. Une intégration par parties permet de passer de I_k à I_{k-1}.
   Par exemple calculons I_2. Partant de $I_1 = \int \frac{du}{u^2+1}$ on pose $f=\frac{1}{u^2+1}$ et $g'=1$. La formule d'intégration par parties $\int fg' = [fg]-\int f'g$ donne (avec $f' = -\frac{2u}{(u^2+1)^2}$ et $g=u$)

$$\begin{aligned} I_1 &= \int \frac{du}{u^2+1} = \left[\frac{u}{u^2+1}\right]+\int\frac{2u^2\,du}{(u^2+1)^2} = \left[\frac{u}{u^2+1}\right]+2\int\frac{u^2+1-1}{(u^2+1)^2}du \\ &= \left[\frac{u}{u^2+1}\right]+2\int\frac{du}{u^2+1}-2\int\frac{du}{(u^2+1)^2} = \left[\frac{u}{u^2+1}\right]+2I_1-2I_2 \end{aligned}$$

   On en déduit $I_2 = \frac{1}{2}I_1+\frac{1}{2}\frac{u}{u^2+1}+c_0$. Mais comme $I_1 = \arctan u$ alors

$$I_2 = \int\frac{du}{(u^2+1)^2} = \frac{1}{2}\arctan u+\frac{1}{2}\frac{u}{u^2+1}+c_0.$$

5.3. Intégration des fonctions trigonométriques

On peut aussi calculer les primitives de la forme $\int P(\cos x,\sin x)\,dx$ ou de la forme $\int \frac{P(\cos x,\sin x)}{Q(\cos x,\sin x)}\,dx$ quand P et Q sont des polynômes, en se ramenant à intégrer une fraction rationnelle.
Il existe deux méthodes :

- les règles de Bioche sont assez efficaces mais ne fonctionnent pas toujours ;
- le changement de variable $t=\tan\frac{x}{2}$ fonctionne tout le temps mais conduit à davantage de calculs.

Les règles de Bioche. On note $\omega(x) = f(x)\,dx$. On a alors $\omega(-x) = f(-x)\,d(-x) = -f(-x)\,dx$ et $\omega(\pi-x) = f(\pi-x)\,d(\pi-x) = -f(\pi-x)\,dx$.

- Si $\omega(-x)=\omega(x)$ alors on effectue le changement de variable $u=\cos x$.
- Si $\omega(\pi-x)=\omega(x)$ alors on effectue le changement de variable $u=\sin x$.
- Si $\omega(\pi+x)=\omega(x)$ alors on effectue le changement de variable $u=\tan x$.

Exemple 15.
Calcul de la primitive $\int \frac{\cos x\, dx}{2-\cos^2 x}$
On note $\omega(x)=\frac{\cos x\, dx}{2-\cos^2 x}$. Comme $\omega(\pi-x)=\frac{\cos(\pi-x)\, d(\pi-x)}{2-\cos^2(\pi-x)}=\frac{(-\cos x)\,(-dx)}{2-\cos^2 x}=\omega(x)$ alors le changement de variable qui convient est $u=\sin x$ pour lequel $du=\cos x\, dx$. Ainsi :

$$\int \frac{\cos x\, dx}{2-\cos^2 x}=\int \frac{\cos x\, dx}{2-(1-\sin^2 x)}=\int \frac{du}{1+u^2}=\big[\arctan u\big]=\arctan(\sin x)+c\ .$$

Le changement de variable $t=\tan\frac{x}{2}$.
Les formules de la « tangente de l'arc moitié » permettent d'exprimer sinus, cosinus et tangente en fonction de $\tan\frac{x}{2}$.

$$\boxed{\begin{array}{c}\text{Avec}\quad t=\tan\dfrac{x}{2}\quad\text{on a}\\ \cos x=\dfrac{1-t^2}{1+t^2}\qquad \sin x=\dfrac{2t}{1+t^2}\qquad \tan x=\dfrac{2t}{1-t^2}\\ \text{et}\qquad dx=\dfrac{2\,dt}{1+t^2}.\end{array}}$$

Exemple 16.
Calcul de l'intégrale $\int_{-\pi/2}^{0}\frac{dx}{1-\sin x}$.
Le changement de variable $t=\tan\frac{x}{2}$ définit une bijection de $[-\frac{\pi}{2},0]$ vers $[-1,0]$ (pour $x=-\frac{\pi}{2}$, $t=-1$ et pour $x=0$, $t=0$). De plus on a $\sin x=\frac{2t}{1+t^2}$ et $dx=\frac{2\,dt}{1+t^2}$.

$$\begin{aligned}\int_{-\frac{\pi}{2}}^{0}\frac{dx}{1-\sin x} &= \int_{-1}^{0}\frac{\frac{2\,dt}{1+t^2}}{1-\frac{2t}{1+t^2}}=2\int_{-1}^{0}\frac{dt}{1+t^2-2t}\\ &= 2\int_{-1}^{0}\frac{dt}{(1-t)^2}=2\left[\frac{1}{1-t}\right]_{-1}^{0}=2\big(1-\frac{1}{2}\big)=1\end{aligned}$$

Mini-exercices.

1. Calculer les primitives $\int \frac{4x+5}{x^2+x-2}\,dx$, $\int \frac{6-x}{x^2-4x+4}\,dx$, $\int \frac{2x-4}{(x-2)^2+1}\,dx$, $\int \frac{1}{(x-2)^2+1}\,dx$.
2. Calculer les primitives $I_k=\int \frac{dx}{(x-1)^k}$ pour tout $k\geqslant 1$. Idem avec $J_k=\int \frac{x\,dx}{(x^2+1)^k}$.
3. Calculer les intégrales suivantes : $\int_0^1 \frac{dx}{x^2+x+1}$, $\int_0^1 \frac{x\,dx}{x^2+x+1}$, $\int_0^1 \frac{x\,dx}{(x^2+x+1)^2}$, $\int_0^1 \frac{dx}{(x^2+x+1)^2}$.
4. Calculer les intégrales suivantes : $\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2 x\cos^3 x\,dx$, $\int_0^{\frac{\pi}{2}} \cos^4 x\,dx$, $\int_0^{2\pi} \frac{dx}{2+\sin x}$.

Auteurs du chapitre
Rédaction : Arnaud Bodin
Basé sur des cours de Guoting Chen et Marc Bourdon
Relecture : Pascal Romon
Dessins : Benjamin Boutin

Chapitre 6

Matrices

Vidéo ■ partie 1. Définition
Vidéo ■ partie 2. Multiplication de matrices
Vidéo ■ partie 3. Inverse d'une matrice : définition
Vidéo ■ partie 4. Inverse d'une matrice : calcul
Vidéo ■ partie 5. Inverse d'une matrice : systèmes linéaires et matrices élémentaires
Vidéo ■ partie 6. Matrices triangulaires, transposition, trace, matrices symétriques
Fiche d'exercices ♦ Calculs sur les matrices

Les matrices sont des tableaux de nombres. La résolution d'un certain nombre de problèmes d'algèbre linéaire se ramène à des manipulations sur les matrices. Ceci est vrai en particulier pour la résolution des systèmes linéaires.

Dans ce chapitre, $\mathbb{K}$ désigne un corps. On peut penser à $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.

1. Définition

1.1. Définition

Définition 1.

- Une ***matrice*** A est un tableau rectangulaire d'éléments de $\mathbb{K}$.
- Elle est dite de ***taille*** $n \times p$ si le tableau possède n lignes et p colonnes.
- Les nombres du tableau sont appelés les ***coefficients*** de A.
- Le coefficient situé à la i-ème ligne et à la j-ème colonne est noté $a_{i,j}$.

Un tel tableau est représenté de la manière suivante :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,j} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,j} & \dots & a_{2,p} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i,1} & a_{i,2} & \dots & a_{i,j} & \dots & a_{i,p} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,j} & \dots & a_{n,p} \end{pmatrix} \quad \text{ou} \quad A = \left(a_{i,j}\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}} \quad \text{ou} \quad \left(a_{i,j}\right).$$

Exemple 1.

$$A = \begin{pmatrix} 1 & -2 & 5 \\ 0 & 3 & 7 \end{pmatrix}$$

est une matrice 2×3 avec, par exemple, $a_{1,1} = 1$ et $a_{2,3} = 7$.

Encore quelques définitions :

Définition 2.

- Deux matrices sont ***égales*** lorsqu'elles ont la même taille et que les coefficients correspondants sont égaux.
- L'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est noté $M_{n,p}(\mathbb{K})$. Les éléments de $M_{n,p}(\mathbb{R})$ sont appelés ***matrices réelles***.

1.2. Matrices particulières

Voici quelques types de matrices intéressantes :

- Si $n = p$ (même nombre de lignes que de colonnes), la matrice est dite ***matrice carrée***. On note $M_n(\mathbb{K})$ au lieu de $M_{n,n}(\mathbb{K})$.

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

  Les éléments $a_{1,1}, a_{2,2}, \dots, a_{n,n}$ forment la ***diagonale principale*** de la matrice.
- Une matrice qui n'a qu'une seule ligne ($n = 1$) est appelée ***matrice ligne*** ou ***vecteur ligne***. On la note

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \end{pmatrix}.$$

- De même, une matrice qui n'a qu'une seule colonne ($p = 1$) est appelée ***matrice colonne*** ou ***vecteur colonne***. On la note

$$A = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{pmatrix}.$$

- La matrice (de taille $n \times p$) dont tous les coefficients sont des zéros est appelée la ***matrice nulle*** et est notée $0_{n,p}$ ou plus simplement 0. Dans le calcul matriciel, la matrice nulle joue le rôle du nombre 0 pour les réels.

1.3. Addition de matrices

Définition 3 (Somme de deux matrices)**.**
Soient A et B deux matrices ayant la même taille $n \times p$. Leur ***somme*** $C = A + B$ est la matrice de taille $n \times p$ définie par

$$c_{ij} = a_{ij} + b_{ij}.$$

En d'autres termes, on somme coefficients par coefficients. Remarque : on note indifféremment a_{ij} où $a_{i,j}$ pour les coefficients de la matrice A.

Exemple 2.

$$\text{Si} \quad A = \begin{pmatrix} 3 & -2 \\ 1 & 7 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 5 \\ 2 & -1 \end{pmatrix} \quad \text{alors} \quad A + B = \begin{pmatrix} 3 & 3 \\ 3 & 6 \end{pmatrix}.$$

$$\text{Par contre si} \quad B' = \begin{pmatrix} -2 \\ 8 \end{pmatrix} \quad \text{alors} \quad A + B' \quad \text{n'est pas définie.}$$

Définition 4 (Produit d'une matrice par un scalaire)**.**
Le produit d'une matrice $A = \left(a_{ij}\right)$ de $M_{n,p}(\mathbb{K})$ par un scalaire $\alpha \in \mathbb{K}$ est la matrice $\left(\alpha a_{ij}\right)$ formée en multipliant chaque coefficient de A par α. Elle est notée $\alpha \cdot A$ (ou simplement αA).

Exemple 3.

$$\text{Si} \qquad A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix} \qquad \text{et} \qquad \alpha = 2 \qquad \text{alors} \qquad \alpha A = \begin{pmatrix} 2 & 4 & 6 \\ 0 & 2 & 0 \end{pmatrix}.$$

La matrice $(-1)A$ est l'***opposée*** de A et est notée $-A$. La ***différence*** $A - B$ est définie par $A + (-B)$.

Exemple 4.

$$\text{Si} \quad A = \begin{pmatrix} 2 & -1 & 0 \\ 4 & -5 & 2 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} -1 & 4 & 2 \\ 7 & -5 & 3 \end{pmatrix} \quad \text{alors} \quad A - B = \begin{pmatrix} 3 & -5 & -2 \\ -3 & 0 & -1 \end{pmatrix}.$$

L'addition et la multiplication par un scalaire se comportent sans surprises :

Proposition 1.
Soient A, B et C trois matrices appartenant à $M_{n,p}(\mathbb{K})$. Soient $\alpha \in \mathbb{K}$ et $\beta \in \mathbb{K}$ deux scalaires.

1. $A + B = B + A$ *: la somme est commutative,*
2. $A + (B + C) = (A + B) + C$ *: la somme est associative,*
3. $A + 0 = A$ *: la matrice nulle est l'élément neutre de l'addition,*
4. $(\alpha + \beta)A = \alpha A + \beta A$*,*
5. $\alpha(A + B) = \alpha A + \alpha B$*.*

Démonstration. Prouvons par exemple le quatrième point. Le terme général de $(\alpha + \beta)A$ est égal à $(\alpha + \beta)a_{ij}$. D'après les règles de calcul dans $\mathbb{K}$, $(\alpha + \beta)a_{ij}$ est égal à $\alpha a_{ij} + \beta a_{ij}$ qui est le terme général de la matrice $\alpha A + \beta A$. □

Mini-exercices.

1. Soient $A = \begin{pmatrix} -7 & 2 \\ 0 & -1 \\ 1 & -4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 21 & -6 \\ 0 & 3 \\ -3 & 12 \end{pmatrix}$, $D = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$, $E = \begin{pmatrix} 1 & 2 \\ -3 & 0 \\ -8 & 6 \end{pmatrix}$. Calculer toutes les sommes possibles de deux de ces matrices. Calculer $3A + 2C$ et $5B - 4D$. Trouver α tel que $A - \alpha C$ soit la matrice nulle.
2. Montrer que si $A + B = A$, alors B est la matrice nulle.
3. Que vaut $0 \cdot A$? et $1 \cdot A$? Justifier l'affirmation : $\alpha(\beta A) = (\alpha\beta)A$. Idem avec $nA = A + A + \cdots + A$ (n occurrences de A).

2. Multiplication de matrices

2.1. Définition du produit

Le produit AB de deux matrices A et B est défini si et seulement si le nombre de colonnes de A est égal au nombre de lignes de B.

Définition 5 (Produit de deux matrices)**.**
Soient $A = (a_{ij})$ une matrice $n \times p$ et $B = (b_{ij})$ une matrice $p \times q$. Alors le produit $C = AB$ est une matrice $n \times q$ dont les coefficients c_{ij} sont définis par :

$$\boxed{c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}}$$

On peut écrire le coefficient de façon plus développée, à savoir :

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} + \cdots + a_{ip}b_{pj}.$$

Il est commode de disposer les calculs de la façon suivante.

$$\begin{array}{ccc}
 & \begin{pmatrix} & & \times & \\ & & \times & \\ & & \times & \\ & & \times & \end{pmatrix} & \leftarrow B \\
A \rightarrow \begin{pmatrix} & & & \\ & & & \\ \times & \times & \times & \times \end{pmatrix} & \begin{pmatrix} & & | & \\ & & | & \\ - & - & - & c_{ij} & \end{pmatrix} & \leftarrow AB
\end{array}$$

Avec cette disposition, on considère d'abord la ligne de la matrice A située à gauche du coefficient que l'on veut calculer (ligne représentée par des $\times$ dans A) et aussi la colonne de la matrice B située au-dessus du coefficient que l'on veut calculer (colonne représentée par des $\times$ dans B). On calcule le produit du premier coefficient de la ligne par le premier coefficient de la colonne ($a_{i1} \times b_{1j}$), que l'on ajoute au produit du deuxième coefficient de la ligne par le deuxième coefficient de la colonne ($a_{i2} \times b_{2j}$), que l'on ajoute au produit du troisième...

2.2. Exemples

Exemple 5.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix}$$

On dispose d'abord le produit correctement (à gauche) : la matrice obtenue est de taille 2×2. Puis on calcule chacun des coefficients, en commençant par le premier coefficient $c_{11} = 1 \times 1 + 2 \times (-1) + 3 \times 1 = 2$ (au milieu), puis les autres (à droite).

$$\begin{array}{cc}
 & \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \\
\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} & \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}
\end{array}
\qquad
\begin{array}{cc}
 & \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \\
\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} & \begin{pmatrix} 2 & c_{12} \\ c_{21} & c_{22} \end{pmatrix}
\end{array}
\qquad
\begin{array}{cc}
 & \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \\
\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} & \begin{pmatrix} 2 & 7 \\ 3 & 11 \end{pmatrix}
\end{array}$$

Un exemple intéressant est le produit d'un vecteur ligne par un vecteur colonne :

$$u = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \qquad v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Alors $u \times v$ est une matrice de taille 1×1 dont l'unique coefficient est $a_1b_1 + a_2b_2 + \cdots + a_nb_n$. Ce nombre s'appelle le ***produit scalaire*** des vecteurs u et v.

Calculer le coefficient c_{ij} dans le produit $A \times B$ revient donc à calculer le produit scalaire des vecteurs formés par la i-ème ligne de A et la j-ème colonne de B.

2.3. Pièges à éviter

Premier piège. Le produit de matrices n'est pas commutatif en général.
En effet, il se peut que AB soit défini mais pas BA, ou que AB et BA soient tous deux définis mais pas de la même taille. Mais même dans le cas où AB et BA sont définis et de la même taille, on a en général $AB \neq BA$.

Exemple 6.

$$\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 14 & 3 \\ -2 & -6 \end{pmatrix} \qquad \text{mais} \qquad \begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 29 & -2 \end{pmatrix}.$$

Deuxième piège. $AB = 0$ n'implique pas $A = 0$ ou $B = 0$.
Il peut arriver que le produit de deux matrices non nulles soit nul. En d'autres termes, on peut avoir $A \neq 0$ et $B \neq 0$ mais $AB = 0$.

Exemple 7.

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 5 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix} \qquad \text{et} \qquad AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Troisième piège. $AB = AC$ n'implique pas $B = C$. On peut avoir $AB = AC$ et $B \neq C$.

Exemple 8.

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 3 \end{pmatrix} \qquad B = \begin{pmatrix} 4 & -1 \\ 5 & 4 \end{pmatrix} \qquad C = \begin{pmatrix} 2 & 5 \\ 5 & 4 \end{pmatrix} \quad \text{et} \quad AB = AC = \begin{pmatrix} -5 & -4 \\ 15 & 12 \end{pmatrix}.$$

2.4. Propriétés du produit de matrices

Malgré les difficultés soulevées au-dessus, le produit vérifie les propriétés suivantes :

Proposition 2.

1. *$A(BC) = (AB)C$: associativité du produit,*
2. *$A(B+C) = AB + AC$ et $(B+C)A = BA + CA$: distributivité du produit par rapport à la somme,*
3. *$A \cdot 0 = 0$ et $0 \cdot A = 0$.*

Démonstration. Posons $A = (a_{ij}) \in M_{n,p}(\mathbb{K})$, $B = (b_{ij}) \in M_{p,q}(\mathbb{K})$ et $C = (c_{ij}) \in M_{q,r}(\mathbb{K})$. Prouvons que $A(BC) = (AB)C$ en montrant que les matrices $A(BC)$ et $(AB)C$ ont les mêmes coefficients.

Le terme d'indice (i,k) de la matrice AB est $x_{ik} = \sum_{\ell=1}^{p} a_{i\ell} b_{\ell k}$. Le terme d'indice (i,j) de la matrice $(AB)C$ est donc

$$\sum_{k=1}^{q} x_{ik} c_{kj} = \sum_{k=1}^{q} \left(\sum_{\ell=1}^{p} a_{i\ell} b_{\ell k} \right) c_{kj}.$$

Le terme d'indice (ℓ, j) de la matrice BC est $y_{\ell j} = \sum_{k=1}^{q} b_{\ell k} c_{kj}$. Le terme d'indice (i,j) de la matrice $A(BC)$ est donc

$$\sum_{\ell=1}^{p} a_{i\ell} \left(\sum_{k=1}^{q} b_{\ell k} c_{kj} \right).$$

Comme dans $\mathbb{K}$ la multiplication est distributive et associative, les coefficients de $(AB)C$ et $A(BC)$ coïncident. Les autres démonstrations se font comme celle de l'associativité. $\square$

2.5. La matrice identité

La matrice carrée suivante s'appelle la ***matrice identité*** :

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Ses éléments diagonaux sont égaux à 1 et tous ses autres éléments sont égaux à 0. Elle se note I_n ou simplement I. Dans le calcul matriciel, la matrice identité joue un rôle analogue à celui du nombre 1 pour les réels. C'est l'élément neutre pour la multiplication. En d'autres termes :

Proposition 3.
Si A est une matrice $n \times p$, alors

$$I_n \cdot A = A \qquad \textit{et} \qquad A \cdot I_p = A.$$

Démonstration. Nous allons détailler la preuve. Soit $A \in M_{n,p}(\mathbb{K})$ de terme général a_{ij}. La matrice unité d'ordre p est telle que tous les éléments de la diagonale principale sont égaux à 1, les autres étant tous nuls. On peut formaliser cela en introduisant le symbole de Kronecker. Si i et j sont deux entiers, on appelle ***symbole de Kronecker***, et on note $\delta_{i,j}$, le réel qui vaut 0 si i est différent de j, et 1 si i est égal à j. Donc

$$\delta_{i,j} = \begin{cases} 0 & \text{si} \quad i \neq j \\ 1 & \text{si} \quad i = j. \end{cases}$$

Alors le terme général de la matrice identité I_p est $\delta_{i,j}$ avec i et j entiers, compris entre 1 et p.
La matrice produit AI_p est une matrice appartenant à $M_{n,p}(\mathbb{K})$ dont le terme général c_{ij} est donné par la formule $c_{ij} = \sum_{k=1}^{p} a_{ik}\delta_{kj}$. Dans cette somme, i et j sont fixés et k prend toutes les valeurs comprises entre 1 et p. Si $k \neq j$ alors $\delta_{kj} = 0$, et si $k = j$ alors $\delta_{kj} = 1$. Donc dans la somme qui définit c_{ij}, tous les termes correspondant à des valeurs de k différentes de j sont nuls et il reste donc $c_{ij} = a_{ij}\delta_{jj} = a_{ij}1 = a_{ij}$. Donc les matrices AI_p et A ont le même terme général et sont donc égales. L'égalité $I_nA = A$ se démontre de la même façon. $\square$

2.6. Puissance d'une matrice

Dans l'ensemble $M_n(\mathbb{K})$ des matrices carrées de taille $n \times n$ à coefficients dans $\mathbb{K}$, la multiplication des matrices est une opération interne : si $A, B \in M_n(\mathbb{K})$ alors $AB \in M_n(\mathbb{K})$.
En particulier, on peut multiplier une matrice carrée par elle-même : on note $A^2 = A \times A$, $A^3 = A \times A \times A$.
On peut ainsi définir les puissances successives d'une matrice :

Définition 6.
Pour tout $A \in M_n(\mathbb{K})$, on définit les puissances successives de A par $A^0 = I_n$ et $A^{p+1} = A^p \times A$ pour tout $p \in \mathbb{N}$. Autrement dit, $A^p = \underbrace{A \times A \times \cdots \times A}_{p \text{ facteurs}}$.

Exemple 9.

On cherche à calculer A^p avec $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. On calcule A^2, A^3 et A^4 et on obtient :

$$A^2 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \qquad A^3 = A^2 \times A = \begin{pmatrix} 1 & 0 & 7 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{pmatrix} \qquad A^4 = A^3 \times A = \begin{pmatrix} 1 & 0 & 15 \\ 0 & 1 & 0 \\ 0 & 0 & 16 \end{pmatrix}.$$

L'observation de ces premières puissances permet de penser que la formule est : $A^p = \begin{pmatrix} 1 & 0 & 2^p - 1 \\ 0 & (-1)^p & 0 \\ 0 & 0 & 2^p \end{pmatrix}$.

Démontrons ce résultat par récurrence.
Il est vrai pour $p = 0$ (on trouve l'identité). On le suppose vrai pour un entier p et on va le démontrer pour $p+1$. On a, d'après la définition,

$$A^{p+1} = A^p \times A = \begin{pmatrix} 1 & 0 & 2^p - 1 \\ 0 & (-1)^p & 0 \\ 0 & 0 & 2^p \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2^{p+1} - 1 \\ 0 & (-1)^{p+1} & 0 \\ 0 & 0 & 2^{p+1} \end{pmatrix}.$$

Donc la propriété est démontrée.

2.7. Formule du binôme

Comme la multiplication n'est pas commutative, les identités binomiales usuelles sont fausses. En particulier, $(A+B)^2$ ne vaut en général pas $A^2 + 2AB + B^2$, mais on sait seulement que

$$(A+B)^2 = A^2 + AB + BA + B^2.$$

Proposition 4 (Calcul de $(A+B)^p$ lorsque $AB = BA$).
Soient A et B deux éléments de $M_n(\mathbb{K})$ qui **commutent**, *c'est-à-dire tels que $AB = BA$. Alors, pour tout entier $p \geqslant 0$, on a la formule*

$$(A+B)^p = \sum_{k=0}^{p} \binom{p}{k} A^{p-k} B^k$$

où $\binom{p}{k}$ désigne le coefficient du binôme.

La démonstration est similaire à celle de la formule du binôme pour $(a+b)^p$, avec $a, b \in \mathbb{R}$.

Exemple 10.

Soit $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. On pose $N = A - I = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. La matrice N est nilpotente (c'est-à-dire il existe $k \in \mathbb{N}$ tel que $N^k = 0$) comme le montrent les calculs suivants :

$$N^2 = \begin{pmatrix} 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad N^3 = \begin{pmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad N^4 = 0.$$

Comme on a $A = I + N$ et les matrices N et I commutent (la matrice identité commute avec toutes les matrices), on peut appliquer la formule du binôme de Newton. On utilise que $I^k = I$ pour tout k et surtout que $N^k = 0$ si $k \geqslant 4$. On obtient

$$A^p = \sum_{k=0}^{p} \binom{p}{k} N^k I^{p-k} = \sum_{k=0}^{3} \binom{p}{k} N^k = I + pN + \tfrac{p(p-1)}{2!} N^2 + \tfrac{p(p-1)(p-2)}{3!} N^3.$$

D'où
$$A^p = \begin{pmatrix} 1 & p & p^2 & p(p^2-p+1) \\ 0 & 1 & 2p & p(3p-2) \\ 0 & 0 & 1 & 3p \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Mini-exercices.

1. Soient $A = \left(\begin{smallmatrix} 0 & 2 & -2 \\ 6 & -4 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & -2 & -3 \end{smallmatrix}\right)$, $C = \left(\begin{smallmatrix} 8 & 2 \\ -3 & 2 \\ -5 & 5 \end{smallmatrix}\right)$, $D = \left(\begin{smallmatrix} 5 \\ 2 \\ -1 \end{smallmatrix}\right)$, $E = \begin{pmatrix} x & y & z \end{pmatrix}$. Quels produits sont possibles ? Les calculer !
2. Soient $A = \left(\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \end{smallmatrix}\right)$. Calculer A^2, B^2, AB et BA.
3. Soient $A = \left(\begin{smallmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 1 & 0 \end{smallmatrix}\right)$. Calculer A^p et B^p pour tout $p \geqslant 0$. Montrer que $AB = BA$. Calculer $(A+B)^p$.

3. Inverse d'une matrice : définition

3.1. Définition

Définition 7 (Matrice inverse)**.**
Soit A une matrice carrée de taille $n \times n$. S'il existe une matrice carrée B de taille $n \times n$ telle que
$$AB = I \qquad \text{et} \qquad BA = I,$$
on dit que A est ***inversible***. On appelle B l'***inverse de A*** et on la note A^{-1}.

On verra plus tard qu'il suffit en fait de vérifier une seule des conditions $AB = I$ ou bien $BA = I$.

- Plus généralement, quand A est inversible, pour tout $p \in \mathbb{N}$, on note :
$$A^{-p} = (A^{-1})^p = \underbrace{A^{-1}A^{-1}\cdots A^{-1}}_{p \text{ facteurs}}.$$
- L'ensemble des matrices inversibles de $M_n(\mathbb{K})$ est noté $GL_n(\mathbb{K})$.

3.2. Exemples

Exemple 11.
Soit $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 3 \end{smallmatrix}\right)$. Étudier si A est inversible, c'est étudier l'existence d'une matrice $B = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ à coefficients dans $\mathbb{K}$, telle que $AB = I$ et $BA = I$. Or $AB = I$ équivaut à :
$$AB = I \iff \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \iff \begin{pmatrix} a+2c & b+2d \\ 3c & 3d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
Cette égalité équivaut au système :
$$\begin{cases} a+2c = 1 \\ b+2d = 0 \\ 3c = 0 \\ 3d = 1 \end{cases}$$
Sa résolution est immédiate : $a = 1$, $b = -\frac{2}{3}$, $c = 0$, $d = \frac{1}{3}$. Il n'y a donc qu'une seule matrice possible, à savoir $B = \left(\begin{smallmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{smallmatrix}\right)$. Pour prouver qu'elle convient, il faut aussi montrer l'égalité $BA = I$, dont la vérification est laissée au lecteur. La matrice A est donc inversible et $A^{-1} = \begin{pmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{pmatrix}$.

Exemple 12.

La matrice $A = \left(\begin{smallmatrix} 3 & 0 \\ 5 & 0 \end{smallmatrix}\right)$ n'est pas inversible. En effet, soit $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice quelconque. Alors le produit

$$BA = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 5 & 0 \end{pmatrix} = \begin{pmatrix} 3a+5b & 0 \\ 3c+5d & 0 \end{pmatrix}$$

ne peut jamais être égal à la matrice identité.

Exemple 13.

- Soit I_n la matrice carrée identité de taille $n \times n$. C'est une matrice inversible, et son inverse est elle-même par l'égalité $I_n I_n = I_n$.
- La matrice nulle 0_n de taille $n \times n$ n'est pas inversible. En effet on sait que, pour toute matrice B de $M_n(\mathbb{K})$, on a $B0_n = 0_n$, qui ne peut jamais être la matrice identité.

3.3. Propriétés

Unicité

> **Proposition 5.**
> *Si A est inversible, alors son inverse est unique.*

Démonstration. La méthode classique pour mener à bien une telle démonstration est de supposer l'existence de deux matrices B_1 et B_2 satisfaisant aux conditions imposées et de démontrer que $B_1 = B_2$.
Soient donc B_1 telle que $AB_1 = B_1A = I_n$ et B_2 telle que $AB_2 = B_2A = I_n$. Calculons $B_2(AB_1)$. D'une part, comme $AB_1 = I_n$, on a $B_2(AB_1) = B_2$. D'autre part, comme le produit des matrices est associatif, on a $B_2(AB_1) = (B_2A)B_1 = I_nB_1 = B_1$. Donc $B_1 = B_2$. $\square$

Inverse de l'inverse

> **Proposition 6.**
> *Soit A une matrice inversible. Alors A^{-1} est aussi inversible et on a :*
>
> $$\boxed{(A^{-1})^{-1} = A}$$

Inverse d'un produit

> **Proposition 7.**
> *Soient A et B deux matrices inversibles de même taille. Alors AB est inversible et*
>
> $$\boxed{(AB)^{-1} = B^{-1}A^{-1}}$$

Il faut bien faire attention à l'inversion de l'ordre !

Démonstration. Il suffit de montrer $(B^{-1}A^{-1})(AB) = I$ et $(AB)(B^{-1}A^{-1}) = I$. Cela suit de

$$\begin{aligned} & (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I, \\ \text{et} \quad & (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I. \end{aligned}$$

$\square$

De façon analogue, on montre que si $A_1, \ldots, A_m$ sont inversibles, alors

$$(A_1A_2 \cdots A_m)^{-1} = A_m^{-1}A_{m-1}^{-1} \cdots A_1^{-1}.$$

Simplification par une matrice inversible

Si C est une matrice quelconque de $M_n(\mathbb{K})$, nous avons vu que la relation $AC = BC$ où A et B sont des éléments de $M_n(\mathbb{K})$ n'entraîne pas forcément l'égalité $A = B$. En revanche, si C est une matrice inversible, on a la proposition suivante :

Proposition 8.
Soient A et B deux matrices de $M_n(\mathbb{K})$ et C une matrice inversible de $M_n(\mathbb{K})$. Alors l'égalité $AC = BC$ implique l'égalité $A = B$.

Démonstration. Ce résultat est immédiat : si on multiplie à droite l'égalité $AC = BC$ par C^{-1}, on obtient l'égalité : $(AC)C^{-1} = (BC)C^{-1}$. En utilisant l'associativité du produit des matrices on a $A(CC^{-1}) = B(CC^{-1})$, ce qui donne d'après la définition de l'inverse $AI = BI$, d'où $A = B$. □

Mini-exercices.

1. Soient $A = \begin{pmatrix} -1 & -2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$. Calculer A^{-1}, B^{-1} , $(AB)^{-1}$, $(BA)^{-1}$, A^{-2}.
2. Calculer l'inverse de $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$.
3. Soit $A = \begin{pmatrix} -1 & -2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Calculer $2A - A^2$. Sans calculs, en déduire A^{-1}.

4. Inverse d'une matrice : calcul

Nous allons voir une méthode pour calculer l'inverse d'une matrice quelconque de manière efficace. Cette méthode est une reformulation de la méthode du pivot de Gauss pour les systèmes linéaires. Auparavant, nous commençons par une formule directe dans le cas simple des matrices 2×2.

4.1. Matrices 2×2

Considérons la matrice 2×2 : $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Proposition 9.
Si $ad - bc \neq 0$, alors A est inversible et

$$\boxed{A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$$

Démonstration. On vérifie que si $B = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ alors $AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Idem pour BA. □

4.2. Méthode de Gauss pour inverser les matrices

La méthode pour inverser une matrice A consiste à faire des opérations élémentaires sur les lignes de la matrice A jusqu'à la transformer en la matrice identité I. On fait simultanément les mêmes opérations élémentaires en partant de la matrice I. On aboutit alors à une matrice qui est A^{-1}. La preuve sera vue dans la section suivante.

En pratique, on fait les deux opérations en même temps en adoptant la disposition suivante : à côté de la matrice A que l'on veut inverser, on rajoute la matrice identité pour former un tableau $(A \mid I)$. Sur les lignes de cette matrice augmentée, on effectue des opérations élémentaires jusqu'à obtenir le tableau $(I \mid B)$. Et alors $B = A^{-1}$.

Ces opérations élémentaires sur les lignes sont :

1. $L_i \leftarrow \lambda L_i$ avec $\lambda \neq 0$: on peut multiplier une ligne par un réel non nul (ou un élément de $\mathbb{K} \setminus \{0\}$).
2. $L_i \leftarrow L_i + \lambda L_j$ avec $\lambda \in \mathbb{K}$ (et $j \neq i$) : on peut ajouter à la ligne L_i un multiple d'une autre ligne L_j.
3. $L_i \leftrightarrow L_j$: on peut échanger deux lignes.

N'oubliez pas : tout ce que vous faites sur la partie gauche de la matrice augmentée, vous devez aussi le faire sur la partie droite.

4.3. Un exemple

Calculons l'inverse de $A = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 0 & -1 \\ -1 & 2 & 2 \end{pmatrix}$.

Voici la matrice augmentée, avec les lignes numérotées :

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 4 & 0 & -1 & 0 & 1 & 0 \\ -1 & 2 & 2 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} L_1 \\ L_2 \\ L_3 \end{array}$$

On applique la méthode de Gauss pour faire apparaître des 0 sur la première colonne, d'abord sur la deuxième ligne par l'opération élémentaire $L_2 \leftarrow L_2 - 4L_1$ qui conduit à la matrice augmentée :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -8 & -5 & -4 & 1 & 0 \\ -1 & 2 & 2 & 0 & 0 & 1 \end{array}\right) \quad L_2 \leftarrow L_2 - 4L_1$$

Puis un 0 sur la première colonne, à la troisième ligne, avec $L_3 \leftarrow L_3 + L_1$:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -8 & -5 & -4 & 1 & 0 \\ 0 & 4 & 3 & 1 & 0 & 1 \end{array}\right) \quad L_3 \leftarrow L_3 + L_1$$

On multiplie la ligne L_2 afin qu'elle commence par 1 :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 4 & 3 & 1 & 0 & 1 \end{array}\right) \quad L_2 \leftarrow -\tfrac{1}{8} L_2$$

On continue afin de faire apparaître des 0 partout sous la diagonale, et on multiplie la ligne L_3. Ce qui termine la première partie de la méthode de Gauss :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{2} & -1 & \frac{1}{2} & 1 \end{array}\right) \quad L_3 \leftarrow L_3 - 4L_2$$

puis

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) \quad L_3 \leftarrow 2L_3$$

Il ne reste plus qu'à « remonter » pour faire apparaître des zéros au-dessus de la diagonale :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) \quad L_2 \leftarrow L_2 - \tfrac{5}{8} L_3$$

puis

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) \quad L_1 \leftarrow L_1 - 2L_2 - L_3$$

Ainsi l'inverse de A est la matrice obtenue à droite et après avoir factorisé tous les coefficients par $\frac{1}{4}$, on a obtenu :

$$A^{-1} = \frac{1}{4}\begin{pmatrix} -2 & 2 & 2 \\ 7 & -3 & -5 \\ -8 & 4 & 8 \end{pmatrix}$$

Pour se rassurer sur ses calculs, on n'oublie pas de vérifier rapidement que $A \times A^{-1} = I$.

Mini-exercices.

1. Si possible calculer l'inverse des matrices : $\begin{pmatrix} 3 & 1 \\ 7 & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & -3 \\ -5 & 4 \end{pmatrix}$, $\begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix}$, $\begin{pmatrix} \alpha+1 & 1 \\ 2 & \alpha \end{pmatrix}$.
2. Soit $A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Calculer $A(\theta)^{-1}$.
3. Calculer l'inverse des matrices : $\begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & -1 \\ -2 & 1 & 1 \end{pmatrix}$, $\begin{pmatrix} 2 & -2 & 1 \\ 3 & 0 & 5 \\ 1 & 1 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -2 & 0 \\ -1 & 2 & 0 & 1 \\ 0 & 2 & 1 & 3 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 1 & 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ -1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 5 & 3 \end{pmatrix}$.

5. Inverse d'une matrice : systèmes linéaires et matrices élémentaires

5.1. Matrices et systèmes linéaires

Le système linéaire

$$\begin{cases} a_{11}\,x_1 & + & a_{12}\,x_2 & + & \cdots & + & a_{1p}\,x_p & = & b_1 \\ a_{21}\,x_1 & + & a_{22}\,x_2 & + & \cdots & + & a_{2p}\,x_p & = & b_2 \\ & & \dots & & & & & & \\ a_{n1}\,x_1 & + & a_{n2}\,x_2 & + & \cdots & + & a_{np}\,x_p & = & b_n \end{cases}$$

peut s'écrire sous forme matricielle :

$$\underbrace{\begin{pmatrix} a_{11} & \dots & a_{1p} \\ a_{21} & \dots & a_{2p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}_{B}.$$

On appelle $A \in M_{n,p}(\mathbb{K})$ la matrice des coefficients du système. $B \in M_{n,1}(\mathbb{K})$ est le vecteur du second membre. Le vecteur $X \in M_{p,1}(\mathbb{K})$ est une solution du système si et seulement si

$$AX = B.$$

Nous savons que :

Théorème 1.
Un système d'équations linéaires n'a soit aucune solution, soit une seule solution, soit une infinité de solutions.

5.2. Matrices inversibles et systèmes linéaires

Considérons le cas où le nombre d'équations égale le nombre d'inconnues :

$$\underbrace{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}_{B}.$$

Alors $A \in M_n(\mathbb{K})$ est une matrice carrée et B un vecteur de $M_{n,1}(\mathbb{K})$. Pour tout second membre, nous pouvons utiliser les matrices pour trouver la solution du système linéaire.

Proposition 10.
Si la matrice A est inversible, alors la solution du système $AX = B$ est unique et est :

$$\boxed{X = A^{-1}B.}$$

La preuve est juste de vérifier que si $X = A^{-1}B$, alors $AX = A\big(A^{-1}B\big) = \big(AA^{-1}\big)B = I \cdot B = B$. Réciproquement si $AX = B$, alors nécessairement $X = A^{-1}B$. Nous verrons bientôt que si la matrice n'est pas inversible, alors soit il n'y a pas de solution, soit une infinité.

5.3. Les matrices élémentaires

Pour calculer l'inverse d'une matrice A, et aussi pour résoudre des systèmes linéaires, nous avons utilisé trois opérations élémentaires sur les lignes qui sont :

1. $L_i \leftarrow \lambda L_i$ avec $\lambda \neq 0$: on peut multiplier une ligne par un réel non nul (ou un élément de $\mathbb{K} \setminus \{0\}$).
2. $L_i \leftarrow L_i + \lambda L_j$ avec $\lambda \in \mathbb{K}$ (et $j \neq i$) : on peut ajouter à la ligne L_i un multiple d'une autre ligne L_j.
3. $L_i \leftrightarrow L_j$: on peut échanger deux lignes.

Nous allons définir trois matrices élémentaires $E_{L_i \leftarrow \lambda L_i}$, $E_{L_i \leftarrow L_i + \lambda L_j}$, $E_{L_i \leftrightarrow L_j}$ correspondant à ces opérations. Plus précisément, le produit $E \times A$ correspondra à l'opération élémentaire sur A. Voici les définitions accompagnées d'exemples.

1. La matrice $E_{L_i \leftarrow \lambda L_i}$ est la matrice obtenue en multipliant par λ la i-ème ligne de la matrice identité I_n, où λ est un nombre réel non nul.

$$E_{L_2 \leftarrow 5L_2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2. La matrice $E_{L_i \leftarrow L_i + \lambda L_j}$ est la matrice obtenue en ajoutant λ fois la j-ème ligne de I_n à la i-ème ligne de I_n.

$$E_{L_2 \leftarrow L_2 - 3L_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3. La matrice $E_{L_i \leftrightarrow L_j}$ est la matrice obtenue en permutant les i-ème et j-ème lignes de I_n.

$$E_{L_2 \leftrightarrow L_4} = E_{L_4 \leftrightarrow L_2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Les opérations élémentaires sur les lignes sont réversibles, ce qui entraîne l'inversibilité des matrices élémentaires.

Le résultat de la multiplication d'un matrice élémentaire E par A est la matrice obtenue en effectuant l'opération élémentaire correspondante sur A. Ainsi :

1. La matrice $E_{L_i \leftarrow \lambda L_i} \times A$ est la matrice obtenue en multipliant par λ la i-ème ligne de A.
2. La matrice $E_{L_i \leftarrow L_i + \lambda L_j} \times A$ est la matrice obtenue en ajoutant λ fois la j-ème ligne de A à la i-ème ligne de A.
3. La matrice $E_{L_i \leftrightarrow L_j} \times A$ est la matrice obtenue en permutant les i-ème et j-ème lignes de A.

Exemple 14.

1. $$E_{L_2 \leftarrow \frac{1}{3} L_2} \times A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \\ \frac{1}{3} y_1 & \frac{1}{3} y_2 & \frac{1}{3} y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$$

2. $$E_{L_1 \leftarrow L_1 - 7 L_3} \times A = \begin{pmatrix} 1 & 0 & -7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} x_1 - 7z_1 & x_2 - 7z_2 & x_3 - 7z_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$$

3. $$E_{L_2 \leftrightarrow L_3} \times A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \\ z_1 & z_2 & z_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

5.4. Équivalence à une matrice échelonnée

Définition 8.
Deux matrices A et B sont dites ***équivalentes par lignes*** si l'une peut être obtenue à partir de l'autre par une suite d'opérations élémentaires sur les lignes. On note $A \sim B$.

Définition 9.
Une matrice est ***échelonnée*** si :

- le nombre de zéros commençant une ligne croît strictement ligne par ligne jusqu'à ce qu'il ne reste plus que des zéros.

Elle est ***échelonnée réduite*** si en plus :

- le premier coefficient non nul d'une ligne (non nulle) vaut 1 ;
- et c'est le seul élément non nul de sa colonne.

Exemple d'une matrice échelonnée (à gauche) et échelonnée réduite (à droite) ; les $*$ désignent des coefficients quelconques, les + des coefficients non nuls :

$$\begin{pmatrix} + & * & * & * & * & * & * \\ 0 & 0 & + & * & * & * & * \\ 0 & 0 & 0 & + & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & + \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & * & 0 & 0 & * & * & 0 \\ 0 & 0 & 1 & 0 & * & * & 0 \\ 0 & 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Théorème 2.
Étant donnée une matrice $A \in M_{n,p}(\mathbb{K})$, il existe une unique matrice échelonnée réduite U obtenue à partir de A par des opérations élémentaires sur les lignes.

Ce théorème permet donc de se ramener par des opérations élémentaires à des matrices dont la structure est beaucoup plus simple : les matrices échelonnées réduites.

Démonstration. Nous admettons l'unicité.
L'existence se démontre grâce à l'algorithme de Gauss. L'idée générale consiste à utiliser des substitutions de lignes pour placer des zéros là où il faut de façon à créer d'abord une forme échelonnée, puis une forme échelonnée réduite.

Soit A une matrice $n \times p$ quelconque.

Partie A. Passage à une forme échelonnée.
Étape A.1. *Choix du pivot.*
On commence par inspecter la première colonne. Soit elle ne contient que des zéros, auquel cas on passe directement à l'étape A.3, soit elle contient au moins un terme non nul. On choisit alors un tel terme, que l'on appelle le ***pivot***. Si c'est le terme a_{11}, on passe directement à l'étape A.2 ; si c'est un terme a_{i1} avec $i \neq 1$, on échange les lignes 1 et i ($L_1 \longleftrightarrow L_i$) et on passe à l'étape A.2.
Au terme de l'étape A.1, soit la matrice A a sa première colonne nulle (à gauche) ou bien on obtient une matrice équivalente dont le premier coefficient a'_{11} est non nul (à droite) :

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1j} & \cdots & a_{1p} \\ 0 & a_{22} & \cdots & a_{2j} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{i2} & \cdots & a_{ij} & \cdots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{n2} & \cdots & a_{nj} & \cdots & a_{np} \end{pmatrix} = A \quad \text{ou} \quad \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1j} & \cdots & a'_{1p} \\ a'_{21} & a'_{22} & \cdots & a'_{2j} & \cdots & a'_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a'_{i1} & a'_{i2} & \cdots & a'_{ij} & \cdots & a'_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a'_{n1} & a'_{n2} & \cdots & a'_{nj} & \cdots & a'_{np} \end{pmatrix} \sim A.$$

Étape A.2. *Élimination.*
On ne touche plus à la ligne 1, et on se sert du pivot a'_{11} pour éliminer tous les termes a'_{i1} (avec $i \geqslant 2$) situés sous le pivot. Pour cela, il suffit de remplacer la ligne i par elle-même moins $\frac{a'_{i1}}{a'_{11}} \times$ la ligne 1, ceci pour $i = 2, \ldots, n$: $L_2 \leftarrow L_2 - \frac{a'_{21}}{a'_{11}} L_1$, $L_3 \leftarrow L_3 - \frac{a'_{31}}{a'_{11}} L_1$,...
Au terme de l'étape A.2, on a obtenu une matrice de la forme

$$\begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1j} & \cdots & a'_{1p} \\ 0 & a''_{22} & \cdots & a''_{2j} & \cdots & a''_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a''_{i2} & \cdots & a''_{ij} & \cdots & a''_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a''_{n2} & \cdots & a''_{nj} & \cdots & a''_{np} \end{pmatrix} \sim A.$$

Étape A.3. *Boucle.*
Au début de l'étape A.3, on a obtenu dans tous les cas de figure une matrice de la forme

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1j}^1 & \cdots & a_{1p}^1 \\ 0 & a_{22}^1 & \cdots & a_{2j}^1 & \cdots & a_{2p}^1 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{i2}^1 & \cdots & a_{ij}^1 & \cdots & a_{ip}^1 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{n2}^1 & \cdots & a_{nj}^1 & \cdots & a_{np}^1 \end{pmatrix} \sim A$$

dont la première colonne est bien celle d'une matrice échelonnée. On va donc conserver cette première colonne. Si $a_{11}^1 \neq 0$, on conserve aussi la première ligne, et l'on repart avec l'étape A.1 en l'appliquant cette fois à la sous-matrice $(n-1)\times(p-1)$ (ci-dessous à gauche : on « oublie » la première ligne et la première colonne de A) ; si $a_{11}^1 = 0$, on repart avec l'étape A.1 en l'appliquant à la sous-matrice $n \times (p-1)$ (à droite, on « oublie » la première colonne) :

$$\begin{pmatrix} a_{22}^1 & \cdots & a_{2j}^1 & \cdots & a_{2p}^1 \\ \vdots & & \vdots & & \vdots \\ a_{i2}^1 & \cdots & a_{ij}^1 & \cdots & a_{ip}^1 \\ \vdots & & \vdots & & \vdots \\ a_{n2}^1 & \cdots & a_{nj}^1 & \cdots & a_{np}^1 \end{pmatrix} \qquad \begin{pmatrix} a_{12}^1 & \cdots & a_{1j}^1 & \cdots & a_{1p}^1 \\ a_{22}^1 & \cdots & a_{2j}^1 & \cdots & a_{2p}^1 \\ \vdots & & \vdots & & \vdots \\ a_{i2}^1 & \cdots & a_{ij}^1 & \cdots & a_{ip}^1 \\ \vdots & & \vdots & & \vdots \\ a_{n2}^1 & \cdots & a_{nj}^1 & \cdots & a_{np}^1 \end{pmatrix}$$

Au terme de cette deuxième itération de la boucle, on aura obtenu une matrice de la forme

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1j}^1 & \cdots & a_{1p}^1 \\ 0 & a_{22}^2 & \cdots & a_{2j}^2 & \cdots & a_{2p}^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{ij}^2 & \cdots & a_{ip}^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nj}^2 & \cdots & a_{np}^2 \end{pmatrix} \sim A,$$

et ainsi de suite.
Comme chaque itération de la boucle travaille sur une matrice qui a une colonne de moins que la précédente, alors au bout d'au plus $p-1$ itérations de la boucle, on aura obtenu une matrice échelonnée.

Partie B. Passage à une forme échelonnée réduite.
Étape B.1. *Homothéties.*
On repère le premier élément non nul de chaque ligne non nulle, et on multiplie cette ligne par l'inverse de cet élément. Exemple : si le premier élément non nul de la ligne i est $\alpha \neq 0$, alors on effectue $L_i \leftarrow \frac{1}{\alpha}L_i$. Ceci crée une matrice échelonnée avec des 1 en position de pivots.

Étape B.2. *Élimination.*
On élimine les termes situés au-dessus des positions de pivot comme précédemment, en procédant à partir du bas à droite de la matrice. Ceci ne modifie pas la structure échelonnée de la matrice en raison de la disposition des zéros dont on part.

□

Exemple 15.
Soit

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ -1 & 0 & 1 & 0 \end{pmatrix}.$$

A. Passage à une forme échelonnée.
Première itération de la boucle, étape A.1. Le choix du pivot est tout fait, on garde $a_{11}^1 = 1$.
Première itération de la boucle, étape A.2. On ne fait rien sur la ligne 2 qui contient déjà un zéro en bonne position et on remplace la ligne 3 par $L_3 \leftarrow L_3 + L_1$. On obtient

$$A \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ 0 & 2 & 4 & 4 \end{pmatrix}.$$

Deuxième itération de la boucle, étape A.1. Le choix du pivot est tout fait, on garde $a_{22}^2 = 2$.
Deuxième itération de la boucle, étape A.2. On remplace la ligne 3 avec l'opération $L_3 \leftarrow L_3 - L_2$. On obtient

$$A \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Cette matrice est échelonnée.

B. Passage à une forme échelonnée réduite.
Étape B.1, homothéties. On multiplie la ligne 2 par $\frac{1}{2}$ et la ligne 3 par $-\frac{1}{2}$ et l'on obtient

$$A \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Étape B.2, première itération. On ne touche plus à la ligne 3 et on remplace la ligne 2 par $L_2 \leftarrow L_2 - 3L_3$ et $L_1 \leftarrow L_1 - 4L_3$. On obtient

$$A \sim \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Étape B.2, deuxième itération. On ne touche plus à la ligne 2 et on remplace la ligne 1 par $L_1 \leftarrow L_1 - 2L_2$. On obtient

$$A \sim \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

qui est bien échelonnée et réduite.

5.5. Matrices élémentaires et inverse d'une matrice

> **Théorème 3.**
> *Soit $A \in M_n(\mathbb{K})$. La matrice A est inversible si et seulement si sa forme échelonnée réduite est la matrice identité I_n.*

Démonstration. Notons U la forme échelonnée réduite de A. Et notons E le produit de matrices élémentaires tel que $EA = U$.

$\Longleftarrow$ Si $U = I_n$ alors $EA = I_n$. Ainsi par définition, A est inversible et $A^{-1} = E$.

$\Longrightarrow$ Nous allons montrer que si $U \neq I_n$, alors A n'est pas inversible.

- Supposons $U \neq I_n$. Alors la dernière ligne de U est nulle (sinon il y aurait un pivot sur chaque ligne donc ce serait I_n).
- Cela entraîne que U n'est pas inversible : en effet, pour tout matrice carrée V, la dernière ligne de UV est nulle ; on n'aura donc jamais $UV = I_n$.
- Alors, A n'est pas inversible non plus : en effet, si A était inversible, on aurait $U = EA$ et U serait inversible comme produit de matrices inversibles (E est inversible car c'est un produit de matrices élémentaires qui sont inversibles).

□

Remarque.
Justifions maintenant notre méthode pour calculer A^{-1}.
Nous partons de $(A|I)$ pour arriver par des opérations élémentaires sur les lignes à $(I|B)$. Montrons que $B = A^{-1}$. Faire une opération élémentaire signifie multiplier à gauche par une des matrices élémentaires. Notons E le produit de ces matrices élémentaires. Dire que l'on arrive à la fin du processus à I signifie $EA = I$. Donc $A^{-1} = E$. Comme on fait les mêmes opérations sur la partie droite du tableau, alors on obtient $EI = B$. Donc $B = E$. Conséquence : $B = A^{-1}$.

Corollaire 1.
Les assertions suivantes sont équivalentes :
(i) La matrice A est inversible.
(ii) Le système linéaire $AX = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ a une unique solution $X = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.
(iii) Pour tout second membre B, le système linéaire $AX = B$ a une unique solution X.

Démonstration. Nous avons déjà vu $(i) \implies (ii)$ et $(i) \implies (iii)$.
Nous allons seulement montrer $(ii) \implies (i)$. Nous raisonnons par contraposée : nous allons montrer la proposition équivalente non$(i) \implies$ non(ii). Si A n'est pas inversible, alors sa forme échelonnée réduite U contient un premier zéro sur sa diagonale, disons à la place ℓ. Alors U à la forme suivante

$$\begin{pmatrix} 1 & 0 & \cdots & c_1 & * & \cdots & * \\ 0 & \ddots & 0 & \vdots & & \cdots & * \\ 0 & 0 & 1 & c_{\ell-1} & & \cdots & * \\ 0 & \cdots & 0 & \color{red}{0} & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \cdots & 0 & \ddots & \vdots \\ 0 & \cdots & & & \cdots & 0 & * \end{pmatrix}. \qquad \text{On note} \qquad X = \begin{pmatrix} -c_1 \\ \vdots \\ -c_{\ell-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Alors X n'est pas le vecteur nul, mais UX est le vecteur nul. Comme $A = E^{-1}U$, alors AX est le vecteur nul. Nous avons donc trouvé un vecteur non nul X tel que $AX = 0$. □

Mini-exercices.

1. Exprimer les systèmes linéaires suivants sous forme matricielle et les résoudre en inversant la matrice :
$$\begin{cases} 2x + 4y = 7 \\ -2x + 3y = -14 \end{cases}, \quad \begin{cases} x + z = 1 \\ -2y + 3z = 1 \\ x + z = 1 \end{cases}, \quad \begin{cases} x + t = \alpha \\ x - 2y = \beta \\ x + y + t = 2 \\ y + t = 4 \end{cases}.$$
2. Écrire les matrices 4×4 correspondant aux opérations élémentaires : $L_2 \leftarrow \frac{1}{3}L_2$, $L_3 \leftarrow L_3 - \frac{1}{4}L_2$, $L_1 \leftrightarrow L_4$. Sans calculs, écrire leurs inverses. Écrire la matrice 4×4 de l'opération $L_1 \leftarrow L_1 - 2L_3 + 3L_4$.
3. Écrire les matrices suivantes sous forme échelonnée, puis échelonnée réduite : $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 0 \\ -2 & -2 & -3 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \\ 2 & -2 & 3 \end{pmatrix}$, $\begin{pmatrix} 2 & 0 & -2 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & -2 & 1 & 4 \\ -1 & 2 & -1 & -2 \end{pmatrix}$.

6. Matrices triangulaires, transposition, trace, matrices symétriques

6.1. Matrices triangulaires, matrices diagonales

Soit A une matrice de taille $n \times n$. On dit que A est ***triangulaire inférieure*** si ses éléments au-dessus de la diagonale sont nuls, autrement dit :

$$i < j \implies a_{ij} = 0.$$

Une matrice triangulaire inférieure a la forme suivante :

$$\begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{pmatrix}$$

On dit que A est ***triangulaire supérieure*** si ses éléments en-dessous de la diagonale sont nuls, autrement dit :

$$i > j \implies a_{ij} = 0.$$

Une matrice triangulaire supérieure a la forme suivante :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & \dots & a_{1n} \\ 0 & a_{22} & \dots & \dots & \dots & a_{2n} \\ \vdots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & a_{nn} \end{pmatrix}$$

Exemple 16.
Deux matrices triangulaires inférieures (à gauche), une matrice triangulaire supérieure (à droite) :

$$\begin{pmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 3 & -2 & 3 \end{pmatrix} \qquad \begin{pmatrix} 5 & 0 \\ 1 & -2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix}$$

Une matrice qui est triangulaire inférieure ***et*** triangulaire supérieure est dite ***diagonale***. Autrement dit : $i \neq j \implies a_{ij} = 0$.

Exemple 17.
Exemples de matrices diagonales :

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

Exemple 18 (Puissances d'une matrice diagonale)**.**
Si D est une matrice diagonale, il est très facile de calculer ses puissances D^p (par récurrence sur p) :

$$D = \begin{pmatrix} \alpha_1 & 0 & \dots & \dots & 0 \\ 0 & \alpha_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{n-1} & 0 \\ 0 & \dots & \dots & 0 & \alpha_n \end{pmatrix} \implies D^p = \begin{pmatrix} \alpha_1^p & 0 & \dots & \dots & 0 \\ 0 & \alpha_2^p & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{n-1}^p & 0 \\ 0 & \dots & \dots & 0 & \alpha_n^p \end{pmatrix}$$

Théorème 4.
Une matrice A de taille $n \times n$, triangulaire, est inversible si et seulement si ses éléments diagonaux sont tous non nuls.

Démonstration. Supposons que A soit triangulaire supérieure.

- Si les éléments de la diagonale sont tous non nuls, alors la matrice A est déjà sous la forme échelonnée. En multipliant chaque ligne i par l'inverse de l'élément diagonal a_{ii}, on obtient des 1 sur la diagonale. De ce fait, la forme échelonnée réduite de A sera la matrice identité. Le théorème 3 permet de conclure que A est inversible.
- Inversement, supposons qu'au moins l'un des éléments diagonaux soit nul et notons $a_{\ell\ell}$ le premier élément nul de la diagonale. En multipliant les lignes 1 à $\ell - 1$ par l'inverse de leur élément diagonal, on obtient une matrice de la forme
$$\begin{pmatrix} 1 & * & \cdots & & & \cdots & * \\ 0 & \ddots & * & \cdots & & \cdots & * \\ 0 & 0 & 1 & * & & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \cdots & 0 & \ddots & \vdots \\ 0 & \cdots & & & \cdots & 0 & * \end{pmatrix}.$$
Il est alors clair que la colonne numéro ℓ de la forme échelonnée réduite ne contiendra pas de 1 comme pivot. La forme échelonnée réduite de A ne peut donc pas être I_n et par le théorème 3, A n'est pas inversible.

Dans le cas d'une matrice triangulaire inférieure, on utilise la transposition (qui fait l'objet de la section suivante) et on obtient une matrice triangulaire supérieure. On applique alors la démonstration ci-dessus. □

6.2. La transposition

Soit A la matrice de taille $n \times p$
$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{pmatrix}.$$

Définition 10.
On appelle ***matrice transposée*** de A la matrice A^T de taille $p \times n$ définie par :
$$A^T = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1p} & a_{2p} & \ldots & a_{np} \end{pmatrix}.$$

Autrement dit : le coefficient à la place (i, j) de A^T est a_{ji}. Ou encore la i-ème ligne de A devient la i-ème colonne de A^T (et réciproquement la j-ème colonne de A^T est la j-ème ligne de A).

Notation : La transposée de la matrice A se note aussi souvent ${}^tA$.

Exemple 19.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & -6 \\ -7 & 8 & 9 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 & -7 \\ 2 & 5 & 8 \\ 3 & -6 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 3 \\ 1 & -5 \\ -1 & 2 \end{pmatrix}^T = \begin{pmatrix} 0 & 1 & -1 \\ 3 & -5 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & -2 & 5 \end{pmatrix}^T = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}$$

L'opération de transposition obéit aux règles suivantes :

Théorème 5.

1. $(A+B)^T = A^T + B^T$
2. $(\alpha A)^T = \alpha A^T$
3. $(A^T)^T = A$
4. $\boxed{(AB)^T = B^T A^T}$
5. *Si A est inversible, alors A^T l'est aussi et on a $(A^T)^{-1} = (A^{-1})^T$.*

Notez bien l'inversion : $(AB)^T = B^T A^T$, comme pour $(AB)^{-1} = B^{-1}A^{-1}$.

6.3. La trace

Dans le cas d'une matrice carrée de taille $n \times n$, les éléments $a_{11}, a_{22}, \ldots, a_{nn}$ sont appelés les ***éléments diagonaux***.

Sa ***diagonale principale*** est la diagonale $(a_{11}, a_{22}, \ldots, a_{nn})$.

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

Définition 11.

La ***trace*** de la matrice A est le nombre obtenu en additionnant les éléments diagonaux de A. Autrement dit,

$$\boxed{\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}.}$$

Exemple 20.

- Si $A = \begin{pmatrix} 2 & 1 \\ 0 & 5 \end{pmatrix}$, alors $\operatorname{tr} A = 2 + 5 = 7$.
- Pour $B = \begin{pmatrix} 1 & 1 & 2 \\ 5 & 2 & 8 \\ 11 & 0 & -10 \end{pmatrix}$, $\operatorname{tr} B = 1 + 2 - 10 = -7$.

Théorème 6.

Soient A et B deux matrices $n \times n$. Alors :

1. $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$,
2. $\operatorname{tr}(\alpha A) = \alpha \operatorname{tr} A$ *pour tout* $\alpha \in \mathbb{K}$,
3. $\operatorname{tr}(A^T) = \operatorname{tr} A$,
4. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Démonstration.

1. Pour tout $1 \leqslant i \leqslant n$, le coefficient (i,i) de $A+B$ est $a_{ii}+b_{ii}$. Ainsi, on a bien $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$.
2. On a $\operatorname{tr}(\alpha A) = \alpha a_{11} + \cdots + \alpha a_{nn} = \alpha(a_{11} + \cdots + a_{nn}) = \alpha \operatorname{tr} A$.
3. Étant donné que la transposition ne change pas les éléments diagonaux, la trace de A est égale à la trace de A^T.
4. Notons c_{ij} les coefficients de AB. Alors par définition
$$c_{ii} = a_{i1}b_{1i} + a_{i2}b_{2i} + \cdots + a_{in}b_{ni}.$$
Ainsi,
$$\begin{array}{rcllll} \operatorname{tr}(AB) & = & a_{11}b_{11} & +a_{12}b_{21} & +\cdots & +a_{1n}b_{n1} \\ & & +a_{21}b_{12} & +a_{22}b_{22} & +\cdots & +a_{2n}b_{n2} \\ & & \vdots & & & \\ & & +a_{n1}b_{1n} & +a_{n2}b_{2n} & +\cdots & +a_{nn}b_{nn}. \end{array}$$
On peut réarranger les termes pour obtenir
$$\begin{array}{rcllll} \operatorname{tr}(AB) & = & a_{11}b_{11} & +a_{21}b_{12} & +\cdots & +a_{n1}b_{1n} \\ & & +a_{12}b_{21} & +a_{22}b_{22} & +\cdots & +a_{n2}b_{2n} \\ & & \vdots & & & \\ & & +a_{1n}b_{n1} & +a_{2n}b_{n2} & +\cdots & +a_{nn}b_{nn}. \end{array}$$
En utilisant la commutativité de la multiplication dans $\mathbb{K}$, la première ligne devient
$$b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1n}a_{n1}$$
qui vaut le coefficient $(1,1)$ de BA. On note d_{ij} les coefficients de BA. En faisant de même avec les autres lignes, on voit finalement que
$$\operatorname{tr}(AB) = d_{11} + \cdots + d_{nn} = \operatorname{tr}(BA).$$

□

6.4. Matrices symétriques

Définition 12.
Une matrice A de taille $n \times n$ est ***symétrique*** si elle est égale à sa transposée, c'est-à-dire si
$$A = A^T,$$
ou encore si $a_{ij} = a_{ji}$ pour tout $i,j = 1,\ldots,n$. Les coefficients sont donc symétriques par rapport à la diagonale.

Exemple 21.
Les matrices suivantes sont symétriques :
$$\begin{pmatrix} 0 & 2 \\ 2 & 4 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 & 5 \\ 0 & 2 & -1 \\ 5 & -1 & 0 \end{pmatrix}$$

Exemple 22.
Pour une matrice B quelconque, les matrices $B \cdot B^T$ et $B^T \cdot B$ sont symétriques.
Preuve : $(BB^T)^T = (B^T)^T B^T = BB^T$. Idem pour $B^T B$.

6.5. Matrices antisymétriques

Définition 13.
Une matrice A de taille $n \times n$ est ***antisymétrique*** si

$$A^T = -A,$$

c'est-à-dire si $a_{ij} = -a_{ji}$ pour tout $i, j = 1, \ldots, n$.

Exemple 23.

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 4 & 2 \\ -4 & 0 & -5 \\ -2 & 5 & 0 \end{pmatrix}$$

Remarquons que les éléments diagonaux d'une matrice antisymétrique sont toujours tous nuls.

Exemple 24.
Toute matrice est la somme d'une matrice symétrique et d'une matrice antisymétrique.

Preuve : Soit A une matrice. Définissons $B = \frac{1}{2}(A + A^T)$ et $C = \frac{1}{2}(A - A^T)$. Alors d'une part $A = B + C$; d'autre part B est symétrique, car $B^T = \frac{1}{2}(A^T + (A^T)^T) = \frac{1}{2}(A^T + A) = B$; et enfin C est antisymétrique, car $C^T = \frac{1}{2}(A^T - (A^T)^T) = -C$.
Exemple :

$$\text{Pour} \quad A = \begin{pmatrix} 2 & 10 \\ 8 & -3 \end{pmatrix} \qquad \text{alors} \qquad A = \underbrace{\begin{pmatrix} 2 & 9 \\ 9 & -3 \end{pmatrix}}_{\text{symétrique}} + \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{\text{antisymétrique}}.$$

Mini-exercices.

1. Montrer que la somme de deux matrices triangulaires supérieures reste triangulaire supérieure. Montrer que c'est aussi valable pour le produit.
2. Montrer que si A est triangulaire supérieure, alors A^T est triangulaire inférieure. Et si A est diagonale ?
3. Soit $A = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Calculer $A^T \cdot A$, puis $A \cdot A^T$.
4. Soit $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Calculer $\operatorname{tr}(A \cdot A^T)$.
5. Soit A une matrice de taille 2×2 inversible. Montrer que si A est symétrique, alors A^{-1} aussi. Et si A est antisymétrique ?
6. Montrer que la décomposition d'une matrice sous la forme « symétrique + antisymétrique » est unique.

Auteurs du chapitre

- D'après un cours de Eva Bayer-Fluckiger, Philippe Chabloz, Lara Thomas de l'École Polytechnique Fédérale de Lausanne,
- et un cours de Sophie Chemla de l'université Pierre et Marie Curie, reprenant des parties de cours de H. Ledret et d'une équipe de l'université de Bordeaux animée par J. Queyrut,
- mixés et révisés par Arnaud Bodin, relu par Vianney Combet.