MAT 101 : Nombres réels, Suites et Fonctions

Sylvain Lavau

Galatasaray Üniversitesi

Table des matières

Remerciements. *Ces notes de cours s'appuient sur plusieurs manuscrits, dont :*

— *F. Liret et D. Martinais,* Analyse 1ère année, *Dunod.*

— *R. Bartle et D. Sherbert,* Introduction to Real Analysis, *John Wiley & Sons.*

— *les cours en ligne de G. Lavau https://gerardlavau.fr/*

Le présent manuscrit contient plusieurs images tirées de ces ressources, ainsi que de Wikipedia ou du cours de J.-L. Rouget (entre autres) que je remercie ici. Il y a bien sûr encore des fautes et des approximations donc n'hésitez pas à m'en faire part.

1 Ensembles, réels et suites

1.1 Historique et définitions de base

Les mathématiques arabes et européennes se sont développées sans fondement axiomatique rigoureux des nombres réels. Ce n'est qu'au XIXème siècle que les évolutions de l'étude des fonctions réelles – qu'on appelle *analyse* en mathématiques – a exigé qu'on développe une compréhension rigoureuse des nombres réels. Cela s'accompagne de la définition axiomatique de la notion d'ensemble et le début de la topologie. Dans l'histoire des mathématiques (même encore aujourd'hui !), beaucoup de choses se font pragmatiquement, de façon pas très bien définie tant que ça fonctionne dans le cadre où on travaille, et ce n'est qu'ensuite que l'on formalise (les résultats précédents restent souvent vrais dans le nouveau formalisme plus rigoureux, à quelques exceptions près).

Apparté historique : la théorie des ensembles de Zermelo-Fraenkel, abrégée en ZF, est une axiomatisation en logique du premier ordre de la théorie des ensembles telle qu'elle avait été développée dans le dernier quart du XIXe siècle par Georg Cantor. L'axiomatisation a été élaborée au début du XXe siècle par plusieurs mathématiciens dont Ernst Zermelo et Abraham Fraenkel mais aussi Thoralf Skolem. En raison de son statut particulier, on considère en général que l'axiome du choix ne fait pas partie de la définition de ZF et on note ZFC la théorie obtenue en ajoutant celui-ci.Les mathématiques usuelles peuvent être théoriquement développées entièrement dans le cadre de la théorie ZFC, éventuellement en ajoutant des axiomes, comme les axiomes de grands cardinaux, pour certains développements (ceux de la théorie des catégories par exemple). En ce sens il s'agit d'une théorie des fondements des mathématiques. Attention si ZFC est cohérente, on ne peut pas axiomatiser ZFC avec un nombre fini d'axiomes (Montague 1961).

La théorie des ensembles de von Neumann–Bernays–Gödel, abrégée en NBG ou théorie des classes, est une théorie axiomatique essentiellement équivalente à la théorie ZFC de Zermelo-Fraenkel avec axiome du choix, mais dont le pouvoir expressif est plus riche (c'est à dire que la théorie NBG décrit des objets plus riches, mais sa restriction aux ensembles coincide avec ZFC). Elle peut s'énoncer en un nombre fini d'axiomes, au contraire de ZFC. Ceci n'est possible que grâce à une modification du langage de la théorie, qui permet de parler directement de classe, une notion par ailleurs utile en théorie des ensembles et qui apparaissait déjà, de façon assez informelle, dans les écrits de Georg Cantor dès avant 1900. La théorie des classes a été introduite en 1925 par John von Neumann, mais celui-ci avait pris des fonctions pour objets primitifs. Elle est reformulée en termes de théorie des ensembles et simplifiée par Paul Bernays vers 19293. Kurt Gödel en donne une version inspirée de celle de Bernays, pour sa preuve de cohérence relative de l'axiome du choix et de l'hypothèse du continu par les constructibles, lors de conférences à Princeton en 1937-1938 (publiées en 1940). Une théorie des classes plus forte, la théorie de Morse-Kelley, a été proposée plus tard par plusieurs mathématiciens, et apparaît pour la première fois en 1955 dans le livre de topologie générale de John L. Kelley.

A notre niveau, nous n'avons pas besoin d'étudier les fondements des mathématiques, car nous devons d'abord étudier les mathématiques dans leur grande régularité, avant de s'intéresser aux mathématiques plus exceptionnelles. D'autre part, les notions d'ensembles, d'éléments et d'appartenance à un ensemble sont des notions abstraites en logique mathématiques, qui sont très difficiles à définir et donc que nous ne définissons pas explicitement. Dans ce cadre, pour nous, un ensemble est intuitivement une collection ou famille – qui est finie ou infinie – d'objets mathématiques appelés *éléments*. Nous ne définirons pas explicitement ces termes et ne donnerons pas une liste d'axiomes pour la théorie des ensembles. Cette approche pragmatique est souvent désignée sous le nom de théorie des ensembles "naïve", et est suffisante pour travailler avec les ensembles rencontrés en analyse réelle.

Si A dénote un ensemble, et si x est un élément, nous utilisons la notation $x \in A$ pour signifier que l'élément x appartient à l'ensemble A. Si l'élément x n'appartient pas à l'ensemble A, on note $x \notin A$. L'ensemble vide est l'ensemble qui ne contient aucun élément ; il est noté $\varnothing$. Un ensemble B est un *sous-ensemble* – ou *partie* – de l'ensemble A si tout élément qui appartient à B appartient aussi à A et on note $B \subset A$. Bien entendu A est un sous-ensemble de lui-même. L'ensemble vide $\varnothing$ est aussi un sous-ensemble de A. Deux ensembles A et B sont dits égaux lorsqu'ils ont les mêmes éléments, c'est à dire lorsque $B \subset A$ et $A \subset B$. Notons que pour prouver que $A = B$ il faut en général montrer les deux inclusions $A \subset B$ et $B \subset A$. L'ensemble dont les éléments sont les sous-ensembles de A est appelé *ensemble des parties de A* et noté $\mathcal{P}(A)$.

Exemple 1.1. L'ensemble des nombres premiers est inclus dans l'ensemble des nombres supérieurs ou égaux à 2.

Remarque 1.2. Si un ensemble A contient un nombre fini d'élément, on appelle *cardinal de A* le nombre d'éléments de A et on le note $\mathrm{card}(A)$ ou $|A|$. On a dans ce cas la relation suivante :

$$\mathrm{card}(\mathcal{P}(A)) = 2^{\mathrm{card}(A)}$$

Si l'ensemble A contient un nombre infini d'élément, le cardinal de A est noté avec un "alef"$\aleph$, indiquant le type d'infini du cardinal de A (ils ne sont pas tous égaux !). Par exemple $\mathrm{card}(\mathbb{N}) = \aleph_0$ ("alef 0") tandis que $\mathrm{card}(\mathbb{R}) = \aleph_1$ ("alef 1"). L'indice en bas du alpeh nous dit que le cardinal de $\mathbb{R}$ est "infiniment plus grand" que le cardinal de $\mathbb{N}$, dans un sens à préciser (voir grands cardinaux et hypothèse du continu sur Wikipedia).

Les ensembles sont notés à l'aide d'accolades. On peut avoir des ensembles de n'importe quoi, mais certains ensembles de nombres sont bien connus :

— l'ensemble des nombres *entiers naturels* $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$;
— l'ensemble des nombres *entiers relatifs* $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$;
— l'ensemble des nombres *rationnels* $\mathbb{Q} = \{\frac{p}{q} \text{ tel que } p, q \in \mathbb{Z}, q \neq 0\}$;
— l'ensemble des nombres *réels* $\mathbb{R}$ = le complété topologique de $\mathbb{Q}$, c'est à dire qu'on a complété $\mathbb{Q}$ en lui ajoutant toutes les limites des suites de nombres rationnels ;
— l'ensemble des nombres *complexes* $\mathbb{C}$ = la clôture algébrique de $\mathbb{R}$, c'est à dire qu'on a rajouté à $\mathbb{R}$ toutes les solutions des équations polynômiales à coefficients réels (en particulier la solution de $x^2 = -1$) ;
— l'ensemble des nombres *quaternions* $\mathbb{H}$ (c'est une généralisation des nombres complexes) ;
— l'ensemble des nombres *octonions* $\mathbb{O}$ (voir Wikipedia pour ces deux derniers ensembles).

Ces ensembles forment une suite de sous-ensembles au sens de l'inclusion :

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C} \subset \mathbb{H} \subset \mathbb{O}$$

Ce sont des inclusions strictes. On rajoute une étoile * aux lettres pour indiquer qu'on considère l'ensemble, mais sans le nombre 0, ainsi :

$$\mathbb{N}^* = \{1, 2, 3, \ldots\}, \quad \mathbb{Z}^* = \{\ldots, -2, -1, 1, 2, \ldots\}, \quad \mathbb{Q}^* = \left\{\frac{p}{q} \text{ tel que } p, q \in \mathbb{Z}^*\right\}$$

Souvent on décrit un ensemble à partir d'une propriété ou de plusieurs propriétés que ses éléments satisfont, par exemple : "être supérieur ou égal à 1" ou "être la solution de l'équation $x^2 - 3x + 2 = 0$". Ces deux propriétés définissent les deux ensembles suivants :

$$A = \{n \in \mathbb{N} \text{ tel que } n \geq 1\} \qquad \text{et} \qquad B = \left\{x \in \mathbb{R} \text{ tel que } x^2 - 3x + 2 = 0\right\}$$

Autrement dit, $A = \mathbb{N}^*$ et B est le sous-ensemble des nombres réels satisfaisant l'équation $x^2 - 3x + 2 = 0$. Comme les solutions de cette équation sont $x = 1$ et $x = 2$, nous avons que $B = \{1, 2\} \subset \mathbb{N}^*$. Donc en particulier B est un sous-ensemble de A. Plus généralement, on aura par exemple :

$$\begin{aligned} A &= \{x \text{ tel que } x \text{ satisfait la propriété } P\} \\ B &= \{x \text{ tel que } x \text{ satisfait les propriétés } P \text{ et } Q\} \end{aligned}$$

pour désigner un ensemble A et un sous-ensemble $B \subset A$. En effet, comme tous les éléments de B satisfont la propriété P, alors ce sont des éléments de A (c'est la définition de sous-ensemble).

Une propriété mathématique est une **proposition** – ou phrase – **logique**. Pour décrire les ensembles comme ci-dessus à l'aide de propriétés il est donc nécessaire d'étudier les propositions logiques.

Les propositions logiques et mathématiques s'expriment souvent à l'aide de *quantificateurs* logiques :

— $\forall x$ veut dire "pour tout (élément) x" (quantificateur universel)

— $\exists x$ veut dire "il existe (un élément) x" (quantificateur existentiel)

Pour dire qu'il n'existe pas d'élément x on peut écrire $\nexists x$, mais c'est un abus d'écriture, il n'a pas de valeur logique particulière. Ainsi, si on veut exprimer la phrase française "pour tout élément x dans l'ensemble A, la propriété P est vérifiée" en termes logiques, on note $\forall x \in A \quad P$. Si on veut exprimer "il existe un élément x dans l'ensemble A tel que la propriété Q est vérifiée par x" en termes logiques, on écrit "$\exists x \in A \mid Q$". La barre verticale signifie "tel que". En général, une expression $\exists x$ est suivie de la barre verticale ou des mots "tel que" et ensuite de la propriété que x satisfait. Par exemple, on a la propriété "être positif ou nul" est vraie pour tout nombre entier naturel donc on peut noter $\forall n \in \mathbb{N}$, $n \geq 0$. Par contre la propriété "être solution de l'équation $x^2 = 100$" n'est satisfaite que pour $x = 10$ et $x = -10$. L'équation a donc au moins une solution et on peut donc écrire $\exists x \in \mathbb{R} \mid x^2 = 100$.

On peut faire se suivre des quantificateurs pour former des propositions logiques plus compliquées. Par exemple la proposition "$\forall x \in \mathbb{R},\ \forall \epsilon > 0,\ \exists n \in \mathbb{N} \mid n\epsilon \geq x$" se lit "pour tout x nombre réel et pour tout ϵ strictement positif, il existe un entier naturel n tel que $n\epsilon$ est supérieur ou égal à x". On appelle cette propriété l'Archimédianité de $\mathbb{R}$. Attention l'ordre des quantificateurs est important ! Si deux quantificateurs se suivent dans l'ordre de lecture de la phrase logique, on peut les échanger :

$$\begin{array}{lcl} \forall n \in \mathbb{N},\ \forall \epsilon > 0,\ \text{blabla} & \iff & \forall \epsilon > 0,\ \forall n \in \mathbb{N},\ \text{blabla} \\ \exists x \in \mathbb{R},\ \exists \delta > 0,\ \text{blabla} & \iff & \exists \delta > 0,\ \exists x \in \mathbb{R},\ \text{blabla} \end{array}$$

Remarquons qu'en général, c'est très rare de voir deux quantificateurs existentiels $\exists$ se suivre. Par contre, si un quantificateur universel $\forall$ et un quantificateur existentiel $\exists$ se suivent, on ne peut PAS intervertir les deux. Par exemple nous avons la proposition suivante :

$$\forall n \in \mathbb{N},\ \exists m \in \mathbb{N} \mid n \leq m$$

qui est vraie, mais la proposition logique obtenue en intervertissant les deux quantificateurs est fausse :

$$\exists m \in \mathbb{N},\ \forall n \in \mathbb{N} \mid n \leq m$$

Le sens de la phrase logique dépend de l'ordre des deux quantificateurs ; il faut donc y faire bien attention.

En logique, il existe des propositions vraies et des propositions fausses. Ecrire une proposition logique n'induit en rien sa valeur de vérité.

$$P_1 :\ \forall z \in \mathbb{N},\ \exists x, y \in \mathbb{N} \mid x^2 + y^2 = z^2$$

est vraie (premiers exemples : 1800 av. J.-C. dans l'ancien Irak sur une tablette d'argile, la première preuve vient des grecs anciens (Pythagore-Euclide), c'est le Théorème de Pythagore). Par contre, la proposition logique suivante est fausse :

$$P_2 :\ \forall n \geq 3, \forall z \in \mathbb{N},\ \exists x, y \in \mathbb{N} \mid x^n + y^n = z^n$$

C'est le grand théorème de Fermat (XVIIème siècle), qui n'a été démontré qu'en 1994 par Andrew Wiles et collaborateurs (Taylor) (Wiles a prouvé un cas particulier de la conjecture de Shimura-Taniyama-Weil, dont on savait depuis quelque temps déjà, via les travaux de Yves Hellegouarch en 1971, puis de Gerhard Frey, Jean-Pierre Serre et Ken Ribet, qu'elle impliquait le théorème). Il existe aussi des propositions logiques dont on ne sait pas encore si elles sont vraies ou fausses (et certaines sont indémontrables, d'après Gödel) :

$$P_3 : \ \forall\, n \geq 2, \ \exists\, p, p' \text{ nombres premiers } \ | \ 2n = p + p'$$

Cette proposition peut se lire "Tout nombre entier pair supérieur à 3 peut s'écrire comme une somme de deux nombres premiers". C'est une conjecture (1742 par Christian Goldbach) vérifiée jusqu'à 4×10^{18} grâce à l'ordinateur, mais jamais démontrée mathématiquement.

La notion de démonstration en mathématique repose sur des *implications*, dénotées avec le symbole $\Longrightarrow$. Ainsi la phrase "si P est vraie alors Q est vraie" s'écrit $P \Longrightarrow Q$. On voit que la proposition P est une *condition suffisante* pour avoir Q : en effet il *suffit* que P soit vraie pour que Q soit vraie aussi. Attention, la notion d'implication ne permet de transférer que la vérité, par la fausseté. Par exemple, si P est vraie alors Q est vraie. Mais si P est fausse, on ne sait rien de Q, et elle peut être vraie ou fausse. Par contre si Q est fausse, P ne peut certainement pas être vraie : car autrement Q serait vraie, ce qui est contradictoire avec le fait qu'elle est fausse. Supposons qu'on ait une autre implication : $Q \Longrightarrow R$. Alors d'après l'argument précédent, pour que Q soit vraie, il *faut* nécessairement que R soit vraie, car si elle était fausse, alors Q serait fausse. La proposition R est donc une *condition nécessaire* pour avoir que la proposition Q est vraie. D'après la chaine d'implications suivantes

$$P \ \Longrightarrow \ Q \ \Longrightarrow \ R,$$

la condition suffisante est plus forte (moins générale) que la condition nécessaire. Cela peut se vérifier par exemple avec les trois propositions suivantes :

$$\begin{aligned} P &: \ \text{"}p \in \mathbb{N} \text{ est un nombre premier"} \\ Q &: \ \text{"}p \in \mathbb{N} \text{ est impair différent de 1 ou } p = 2\text{"} \\ R &: \ \text{"}p \in \mathbb{N} \text{ est supérieur ou égal à 2"} \end{aligned}$$

Les conditions suffisante P et nécessaire R ont des rôles différents vis-à-vis de la valeur de vérité de la proposition Q : la condition suffisante P – si elle est vraie/satisfaite – nous permet de savoir au que Q est vraie ; tandis que la condition nécessaire R – si elle est fausse/non satisfaite – nous permet de savoir au que Q est fausse. On a donc une connaissance de la condition Q à partir des conditions suffisantes et nécessaires. On peut le résumer par le tableau suivant :

Valeur de vérité	Condition suffisante P	Condition nécessaire R
si la proposition est vraie	alors Q est vraie	On ne sait rien sur Q
si la proposition est fausse	On ne sait rien sur Q	alors Q est fausse

Maintenant si $P = R$, c'est à dire si la condition suffisante est aussi une condition nécessaire, alors on a la chaine d'implications suivante :

$$P \ \Longrightarrow \ Q \ \Longrightarrow \ P,$$

Dans ce cas on dit que P et Q sont des conditions équivalentes, ou encore que "P si et seulement si Q". Les deux conditions/propositions sont simultanément vraies ou simultanément fausses.

Maintenant intéressons nous à écrire des négations de propositions logiques/mathématiques. Notons d'abord qu'une proposition logique P contient en général dans notre cours une partie avec tous les quantificateurs, qu'on note P_{quantif}, qui introduit tous les objets mathématiques qui seront utilisés dans la deuxième partie de la proposition logique. La deuxième partie de la proposition P consiste en une ou plusieurs formules mathématiques, sans quantificateurs, où les objets mathématiques introduits P_{quantif} jouent un rôle. On note cette deuxième partie P_{formule}. Par exemple la proposition logique d'Archimédianité de $\mathbb{R}$ (qu'on peut prouver intuitivement avec un dessin, donc qui est vraie) :

$$\forall x \in \mathbb{R},\ \forall \epsilon > 0,\ \exists n \in \mathbb{N} \mid n\epsilon \leq x < (n+1)\epsilon$$

contient les deux parties suivantes :

$$P_{\text{quantif}} :\ \forall x \in \mathbb{R},\ \text{et}\ \forall \epsilon > 0,\ \exists n \in \mathbb{N} \quad \text{et} \quad P_{\text{formule}} :\ n\epsilon \leq x < (n+1)\epsilon$$

Notons que la deuxième partie s'écrit, de façon plus rigoureuse : $P_{\text{formule}} :\ n\epsilon \leq x$ et $x < (n+1)\epsilon$. La négation de la proposition P revient à écrire une proposition logique notée $non - P$, formée de la négation de la partie P_{quantif}, et de la négation de la partie P_{formule}, obtenues par les règles suivantes :

- pour écrire la négation de la partie P_{quantif}, on change tous les $\forall$ par des $\exists$, et les $\exists$ par des $\forall$, tout en conservant l'ordre dans lequel ils apparaissent ;
- pour écrire la négation de la partie P_{formule}, on change tous les $=$ par des $\neq$, les $\geq$ par des $<$, les $\leq$ par des $>$, les "ou" par des "et" et les "et" par des "ou"

Dans l'exemple ci dessus, on obtient selon les régles données :

$$\begin{aligned} non - P_{\text{quantif}} &:\ \exists x \in \mathbb{R} \text{ et } \exists \epsilon > 0 \text{ tels que } \forall n \in \mathbb{N} \\ non - P_{\text{formule}} &:\ n\epsilon > x \text{ ou } x \geq (n+1)\epsilon \end{aligned}$$

ce qui donne pour la négation de la proposition P :

$$non - P :\ \exists x \in \mathbb{R} \text{ et } \exists \epsilon > 0 \text{ tels que } \forall n \in \mathbb{N}\ n\epsilon > x \text{ ou } x \geq (n+1)\epsilon$$

Cette proposition est évidemment fausse – rappelons que la proposition P est vraie – comme cela peut se comprendre en lisant la phrase.

Parfois la partie P_{formule} contient une flèche d'implication $\Longrightarrow$, comme dans l'exemple suivant (intégrité de $\mathbb{R}$) :

$$P :\ \forall a, b \in \mathbb{R},\quad ab = 0 \implies (\ a = 0 \text{ ou } b = 0)$$

Dans ce cas on a les deux parties suivantes :

$$P_{\text{quantif}} :\ \forall a, b \in \mathbb{R} \quad \text{et} \quad P_{\text{formule}} :\ ab = 0 \implies (a = 0 \text{ ou } b = 0)$$

Pour écrire la négation de la partie formule, on admet que la négation de l'implication $Q \implies R$ est la proposition "Q et $non - R$". Ainsi la négation de P_{formule} ci dessus devient (rappelons qu'il faut changer le "ou" en "et") : $non - P_{\text{formule}} :\ ab = 0$ et $(a \neq 0$ et $b \neq 0)$. La négation de l'intégrité de $\mathbb{R}$ devient donc :

$$\exists a, b \in \mathbb{R} \mid ab = 0 \text{ et } a \neq 0 \text{ et } b \neq 0$$

Comme $\mathbb{R}$ est intègre, la proposition P est vraie et sa négation est fausse (cela se voit). La négation d'une implication $Q \implies R$ consiste à établir que l'implication ne tient pas. En termes logiques, cela revient à avoir Q et $non - R$ en même temps. En effet, si jamais ces deux propositions Q et $non - R$ sont vraies en même temps, alors l'implication originelle $Q \implies R$ n'est pas vraie. On a donc bien sa négation, c'est à dire qu'on a l'équivalence logique suivante :

$$non - (Q \implies R) \quad \Longleftrightarrow \quad Q \text{ et } non - R$$

La notion de *contraposée* transforme le rôle des conditions dans une implication pour obtenir une nouvelle implication complètement équivalente à la première (ce n'est donc pas une négation d'implication comme ci-dessus). Imaginons qu'on ait $P \Longrightarrow Q$, alors du fait que Q est une condition nécessaire à P, on sait que si on n'a pas Q (i.e. si Q est fausse) alors nécessairement on n'a pas P (i.e. P est fausse). Autrement dit, et de façon tout à fait équivalente, si la négation de la condition/proposition Q est vraie – on la note $non - Q$ – alors la négation de la condition/proposition P est vraie – on la note $non - P$. On a donc par ce raisonnement l'implication

$$non - Q \implies non - P$$

Cette implication est appelée la *contraposée* de l'implication originale $P \Longrightarrow Q$. Il est importante de savoir que démontrer l'une ou l'autre est équivalent logiquement, c'est à dire que :

$$(P \Longrightarrow Q) \quad \iff \quad (non - Q \Longrightarrow non - P)$$

Par exemple, on a vu ci dessus que si $a, b \in \mathbb{R}$, alors on a l'implication $ab = 0 \implies (a = 0$ ou $b = 0)$. La contraposée de cette formule logique est $(a \neq 0$ et $b \neq 0) \implies ab \neq 0$. Attention nous rappelons que la contraposée n'est pas la négation de la proposition logique d'origine, il y a donc stricte équivalence entre les deux propositions suivantes :

$$P : \forall a, b \in \mathbb{R},\ \ ab = 0 \implies (\ a = 0 \text{ ou } b = 0) \quad \text{et} \quad P' : \forall a, b \in \mathbb{R},\ \ (a \neq 0 \text{ et } b \neq 0) \implies ab \neq 0$$

Le raisonnement par l'absurde se base sur l'idée que $non - (non - P) = P$. Autrement dit "si $non - P$ est fausse, alors P est vraie". Pour utiliser la proposition P, on peut donc supposer que sa négation $non - P$ est vraie, et si après un raisonnement on trouve une contradiction, cela veut dire que $non - P$ est fausse, donc P est vraie. Cela fonctionne bien pour l'irrationalité de la racine carrée de 2 par exemple. On pose P : "$\sqrt{2}$ est irrationnelle". Supposons que la racine carrée de 2 est rationnelle – c'est à dire on suppose que $non - P$ est vraie. Dans ce cas, $\sqrt{2}$ s'écrit p/q où $p \in \mathbb{Z}$ et $q \in \mathbb{Z}^*$ sont des entiers naturels qui n'ont aucun facteurs en commun (on peut le supposer). Alors $p^2 = 2q^2$ est pair, mais si un carré est pair alors p est pair et s'écrit $p = 2n$ avec un certain $n \in \mathbb{Z}$. Mais alors $p^2 = 4n^2$ donc $2q^2 = 4n^2$ donc $q^2 = 2n^2$ donc q est un nombre pair aussi et s'écrit donc $q = 2m$ avec un certain $m \in \mathbb{Z}$. Mais dans ce cas $p = 2n$ et $q = 2m$ ont un facteur commun : c'est 2. C'est une contradiction avec notre hypothèse. Donc $non - P$ est fausse donc P est vraie. Attention, en général on n'a pas besoin d'utiliser le raisonnement par l'absurde, mais les élèves le font par défaut. Ce n'est que très rarement qu'on doive démontrer un résultat par l'absurde, en général pour montrer des résultats très puissants (par exemple l'argument diagonal de Cantor).

De façon plus abstraite, les démonstrations d'inclusions d'ensembles utilisent beaucoup les implications. Pour démontrer qu'on a une inclusion $B \subset A$ (deux ensembles définis par certaines conditions), on raisonne comme suit : on prend un élément $x \in B$ (quelconque, donc abstrait, qui satisfait la condition définissant l'ensemble B), on fait un raisonnement logique, et on arrive à déduire qu'il est dans A aussi (c'est à dire qu'il satisfait la condition définissant A). Par exemple soit A l'ensemble des nombres entiers supérieurs à 2 et soit B l'ensemble des nombres premiers. Soit $p \in B$ un nombre premier, alors nécessairement p est impair strictement plus grand que 1 ou $p = 2$. Dans tous les cas $p \geq 2$ donc $p \in A$. On a donc montré que $B \subset A$. Pour montrer l'égalité entre deux ensembles A et B, il **faut** montrer les deux inclusions $B \subset A$ et $A \subset B$. Cela se traduit logiquement par la proposition suivante :

$$P : \ A = B \iff B \subset A \text{ et } A \subset B$$

Par contre pour prouver que $A \neq B$, il **suffit** de trouver UN élément de A qui n'est pas dans B ou inversement, c'est à dire montrer $non - P$:

$$non - P : \ A \neq B \iff \exists\, x \in B \text{ tel que } x \notin A \text{ ou } x \in A \text{ tel que } x \notin B$$

Par exemple, soit A l'ensemble des nombres entiers supérieurs à 2, soit B l'ensemble des nombres premiers, et soit C l'ensemble des nombres entiers naturels impairs. Alors on a $B \subset A$ mais $B \neq A$ car $1 \in A$ mais $1 \notin B$. D'autre part on a $B \neq C$, et B n'est pas inclus dans C car $2 \in B$ mais $2 \notin C$.

Maintenant définissons quelques notions ensemblistes assez transparentes :

— La réunion des ensembles A et B est l'ensemble noté $A \cup B$ défini par la condition

$$x \in A \cup B \iff x \in A \text{ ou } x \in B$$

— L'intersection des ensembles A et B est l'ensemble noté $A \cap B$ défini par la condition

$$x \in A \cap B \iff x \in A \text{ et } x \in B$$

— L'ensemble "A privé de B" est le sous-ensemble de A noté $A \backslash B$ défini par la condition

$$x \in A \backslash B \iff x \in A \text{ et } x \notin B$$

En particulier, si $B \subset A$, l'ensemble $A \backslash B$ est appelé "sous-ensemble complémentaire de B (par rapport à l'ensemble A)" et souvent noté B^c

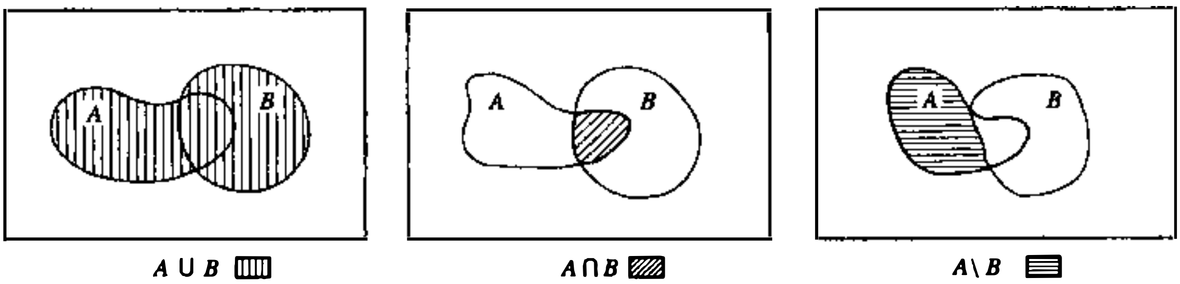
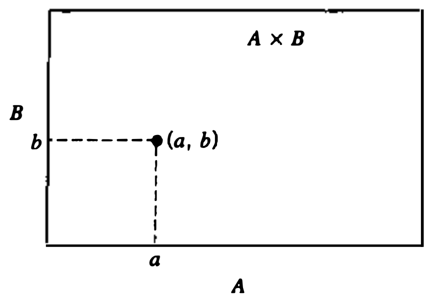


Figure 1.1.1 (a) $A \cup B$ (b) $A \cap B$ (c) $A \backslash B$

Enfin, le produit (cartésien) des ensembles A et B est un ensemble formé des couples/paires d'éléments de A et de B. On le dénote $A \times B$ et on le définit comme :

$$A \times B = \{(a, b) \mid a \in A \text{ et } b \in B\}$$



Une *application* $f : A \to B$ à partir d'un ensemble A vers un ensemble B établit une relation entre tout ou partie des éléments de A et ceux de B. L'application f associe un sous-ensemble de A appelé *domaine de définition de* f à un sous-ensemble de B appelé *image de* f. Plus précisément, chaque élément x du domaine de définition de f est associé par l'application f à un unique élément de B, noté $f(x)$: $f(x)$ est l'image de x par f. On appelle $f(x)$ *l'image de* x *par* f. Le domaine de définition de f est défini grâce à cette notion :

$$D_f = \{x \in A \mid f(x) \text{ existe}\}$$

tandis que *l'image de* f est le sous-ensemble de B formé de tous les éléments du type $f(x)$ c'est à dire :

$$\mathrm{Im}(f) = \{f(x) \mid x \in D_f\} \subset B$$

Maintenant soit $y \in B$, et supposons qu'il existe $x \in A$ tel que son image par f est y, c'est à dire tel que $y = f(x)$. On dit que x est UN *antécédent* – ou UNE *préimage* – de y par f (car en effet l'élément y peut admettre plusieurs antécédents).

Remarque 1.3. Lorsque $A = \mathbb{R}$ ou $\mathbb{C}$ (ou tout intervalle de $\mathbb{R}$, domaine de $\mathbb{C}$), et $B = \mathbb{R}$ ou $\mathbb{C}$, on parle de *fonction* plutôt que d'application.

Exemple 1.4. Prenons par exemple la fonction $f : \mathbb{R} \to \mathbb{R}, x \mapsto \frac{1}{x^2}$. Le domaine de définition de f est $D_f = \mathbb{R}^*$ car la fonction n'est pas définie en 0. L'image de f est $\mathrm{Im}(f) = \mathbb{R}_+^*$ c'est à dire l'ensemble des nombres réels strictement positifs. Les nombres strictement négatifs et 0 ne sont pas atteignables : il n'existe aucun réel non-nul $x \in \mathbb{R}$ tel que $\frac{1}{x^2} \leq 0$. Prenons $y > 0$, les antécédents de y sont au nombre de deux : $\sqrt{y}$ et $-\sqrt{y}$.

Si $U \subset A$ est un sous-ensemble de A, $f(U)$ est le sous ensemble de $\mathrm{Im}(f) \subset B$ formé par toutes les images des points de U ; c'est *l'image directe* de U :

$$f(U) = \{y \in B \mid \exists x \in U \text{ tel que } y = f(x)\} \subset \mathrm{Im}(f) \subset B$$

Si $V \subset B$ est un sous-ensemble de B, alors on définit $f^{-1}(V)$ comme étant le sous-ensemble de $D_f \subset A$ formé de toutes les préimages/antécédents des éléments de V, c'est *l'image réciproque* de V :

$$f^{-1}(V) = \{x \in A \mid f(x) \in V\} \subset D_f \subset A$$

Attention, le f^{-1} dans le membre de gauche ici est une **notation**, qui n'est pas détachable de V dans ce cadre précis où on étudie l'image réciproque. Ce n'est que si $f : A \to B$ est bijective (voir ci-dessous), qu'on peut définir une fonction inverse avec une notation indépendante $f^{-1} : B \to A$.

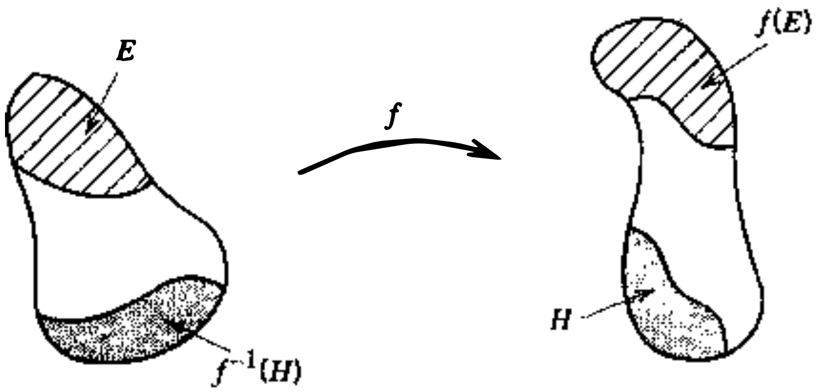


Figure 1.1.7 Direct and inverse images

Exemple 1.5. A la suite de l'exemple 1.4, où nous avions $A = B = \mathbb{R}$, et $f : x \mapsto \frac{1}{x^2}$, de façon que $D_f = \mathbb{R}^*$ et $\mathrm{Im}(f) =]0, +\infty[$. Soit $U = [-1, 1] \subset A$ et $V =]-1, 2[\subset B$. L'ensemble $f(U) \subset \mathrm{Im}(f)$ est formé de tous les éléments y dont une préimage au moins est dans U ; dans ce cas on a que $f(U) = [-1, 0[\cup]0, 1]$. L'ensemble $f^{-1}(V) \subset D_f$ est formé de tous les éléments $x \in D_f$ tels que $f(x) \in V$. Dans notre cas, comme $\mathrm{Im}(f) \cap V =]0, 2[$ donc $f^{-1}(V) = f^{-1}(]0, 2[)$, ce qui nous donne $f^{-1}(V) =]-\infty, \sqrt{2}[\cup]\sqrt{2}, +\infty[$.

On note $f|_U : U \to A$ la *restriction* de l'application f au sous-ensemble U. C'est juste la même application, mais on a arbitrairement restreint son ensemble de départ, donc son image sera certainement plus petite aussi. En particulier, on a les résultats suivants :

$$D_{f|_U} = U \cap D_f \quad \text{et} \quad \mathrm{Im}\left(f|_U\right) = f(U)$$

L'application obtenue en restreignant la portée de f au seul sous-ensemble V est appelée la *corestriction* de f au sous-ensemble V, et est notée $f|^V : A \to V$. C'est l'unique fonction induite par f dont le domaine de définition ne contient que les éléments de $D_f \subset A$ qui sont envoyés dans V. Donc on a en particulier :

$$D_{f|^V} = f^{-1}(V) \cap D_f \quad \text{et} \quad \mathrm{Im}\left(f|^V\right) = \mathrm{Im}(f) \cap V$$

Pour finir, on note B^A l'ensemble des applications de A dans B (quelles qu'elles soient).

Une application $f : A \to B$ est *surjective* si $B = f(A)$, c'est à dire si $\mathrm{Im}(f) = B$, c'est à dire si tous les points de B sont atteints par l'application. Une application $f : A \to B$ est *injective* si tout point de l'image de f n'admet qu'au plus un seul antécédent, ce qu'on peut écrire comme ceci :

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

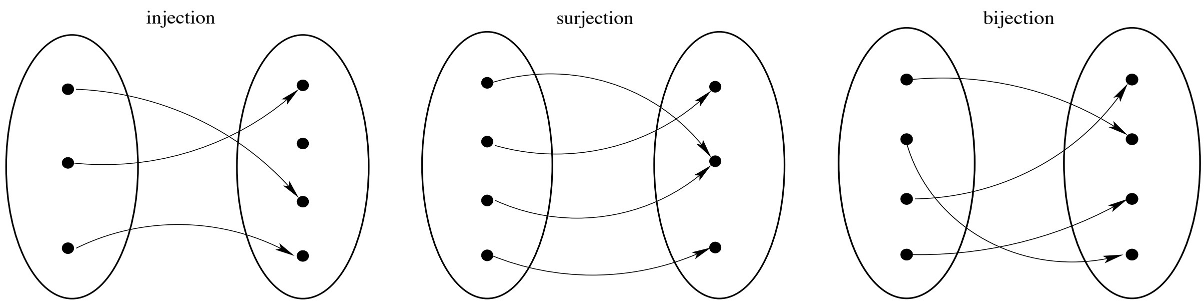
La contraposée de cette implication se lit donc :

$$\forall\, x, x' \in D_f, x \neq x' \implies f(x) \neq f(x')$$

Autrement dit l'application ne perd aucune information de D_f. Une application $f : A \to B$ est *bijective* si elle est à la fois injective et surjective. Dans ce cas, tout point de A est envoyé dans B et tout point de B est l'image par f d'un unique antécédent. On peut alors définir une application de B vers A, appelée *inverse de* f, et notée $f^{-1} : B \to A$, telle que :

$$\forall\, x \in A,\ f^{-1}(f(x)) = a \quad \text{et} \quad \forall\, y \in B,\ f(f^{-1}(y)) = y$$

Ces deux identités justifient le nom d'*inverse*. D'autre part on voit que si $f : A \to B$ est bijective, l'application $f^{-1} : B \to A$ est elle aussi bijective, et $\left(f^{-1}\right)^{-1} = f$.



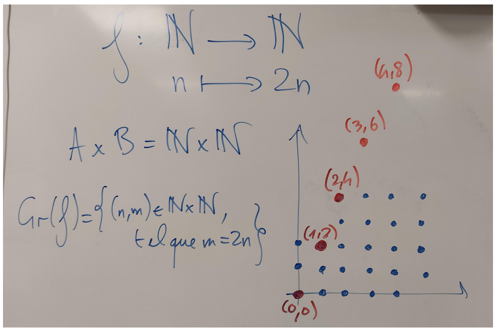
Exemple 1.6. La fonction $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^3 - x$ est surjective mais pas injective, car 0 a trois antécédents. La fonction tangente est une bijection entre $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ et $\mathbb{R}$.

Une autre façon de voir l'application $f : A \to B$ est de remarquer que f est la donnée d'un sous-ensemble $G_f \subset A \times B$ de l'ensemble produit $A \times B$. On appelle G_f le *graphe de* f, et il est défini par :

$$G_f = \{(x, y) \in A \times B \mid y = f(x)\}$$

C'est à dire que G_f est l'ensemble de toutes les paires d'éléments $(x, y) \in A \times B$ tels que x est un antécédent de y. Maintenant on voit que les sous-ensembles $G_f \subset A \times B$ du type "graphe d'une application f" sont d'un type vraiment particulier. Il existe plein de sous-ensemble de $A \times B$ qui ne *peuvent pas* s'écrire comme le graphe G_f d'une application f donnée. L'exemple le plus simple est de prendre $x \in A$ fixé, alors le sous-ensemble $G = \{x\} \times B$ est un sous-ensemble de $A \times B$ mais il n'existe pas d'application $f : A \to B$ tel que $G = G_f$, c'est parce qu'il est impossible d'atteindre l'ensemble de B à partir du seul élément $x \in A$, quelle que soit la fonction choisie.

Exemple 1.7. Par exemple, si $f : \mathbb{R} \to \mathbb{R}$ est une fonction réelle, alors le graphe de f est la courbe de $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ telle que $y = f(x)$; c'est donc bien le graphe de f au sens usuel !
Si $\varphi : \mathbb{N} \to \mathbb{N}$ est la fonction qui envoie n sur $2n$, alors le graphe de φ est le sous-ensemble de $\mathbb{R}^2$ formé des points à coordonnées entières positives ou nulles localisés sur la droite d'équation $y = 2x$ (ce sont les points rouges sur l'image).



Pour finir ce chapitre, nous nous intéressons à des sous-ensembles particuliers de $A \times A$, c'est à dire que nous allons nous limiter au cas $B = A$. Dans ce cas particulier, une application naturelle est l'application identité $\mathrm{id}_A : A \to A, x \mapsto x$. Le graphe de cette application est un sous-ensemble de $A \times A$, qu'on appelle la *diagonale de* A et qu'on note Δ_A, autrement dit $\Delta_A = G_{\mathrm{id}_A}$. Le nom vient du fait que si $A = \mathbb{R}$, alors le graphe de la fonction identité $\mathrm{id}_{\mathbb{R}} : x \mapsto x$ est le sous-ensemble de $\mathbb{R}^2$ formé des points (x, y) tels que $y = x$. Autrement dit, c'est la droite d'équation $y = x$: la diagonale dans $\mathbb{R}^2$. Nous avons donc ici une caractérisation géométrique d'une notion ensembliste (la fonction identité). En voici une autre : commençons par dire qu'un sous-ensemble $G \subset A \times A$ est *symétrique par rapport à la diagonale* Δ_A si pour tout $x, y \in A$, on a $(x, y) \in G \implies (y, x) \in G$. Lorsque $A = \mathbb{R}$, un sous-ensemble $G \subset \mathbb{R}^2$ qui est symétrique dans ce sens, est réellement symétrique par rapport à la droite d'équation $y = x$. Nous voyons que nous pouvons donner des interprétations géométriques à des notions analytiques, algébriques ou ensemblistes, ce qui offre des perspectives différentes, ouvre des portes, etc. C'est une illustration des liens forts entre différents domaines de mathématiques.

Tournons nous maintenant vers un autre type de sous ensemble spécifique de $A \times A$. On appelle *relation binaire sur* A une propriété portant sur les couples d'éléments de A. On note en général une relation binaire $\mathcal{R}$, et on écrit $x \,\mathcal{R}\, y$ pour dire que la paire d'éléments (x, y) a la propriété définie par la relation binaire $\mathcal{R}$. Nous connaissons déjà des relations binaires, par exemple "être égal à", "être inférieur ou égal à", etc. En effet lorsque $A = \mathbb{R}$: une paire d'éléments (x, y) satisfait la propriété "être égal à" si l'élément de gauche x est égal à celui de droite y, et on note $x = y$. Une paire d'éléments (x, y) satisfait la propriété "être inférieur ou égal à" si l'élément de gauche x est inférieur ou égal à celui de droite y, et on symbolise cette relation par $x \leq y$. Lorsque $A = \mathbb{N}^*$ la division euclidienne est une relation binaire car la paire d'entiers naturels (n, m) satisfait la propriété "divise" si l'entier de gauche n divise l'entier de droite m. On note cette relation $n|m$ (avec une barre verticale).

Vu depuis la perspective des paires d'éléments, nous voyons qu'une relation binaire $\mathcal{R}$ sur A définit un sous-ensemble du produit cartésien $A \times A$, qu'on notera $G_{\mathcal{R}}$. Ce sous-ensemble est formé de toutes les paires $(x, y) \in A \times A$ telles que $x\mathcal{R}y$, c'est à dire :

$$G_{\mathcal{R}} = \{(x, y) \in A \times A \text{ tel que } x\mathcal{R}y\}$$

De façon réciproque, tout sous-ensemble G du produit cartésien $A \times A$ définit une relation binaire $\mathcal{R}_G$, par la convention qu'une paire (x, y) de G peut aussi s'écrire $x \,\mathcal{R}_G\, y$, où $\mathcal{R}_G$ traduit la notion de "relation" entre x et y, celle que $(x, y) \in G$:

$$x\mathcal{R}_G y \text{ est définie si et seulement si } (x, y) \in G$$

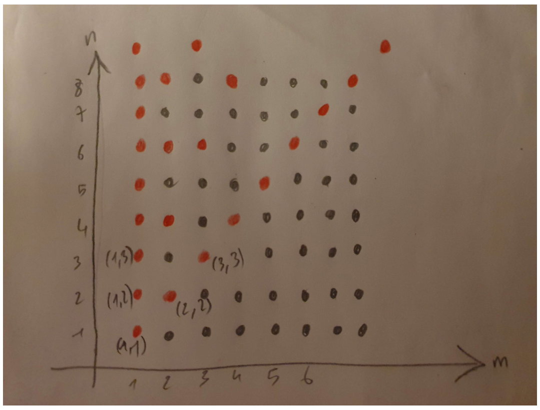
Parler de relation binaire ou de sous-ensemble de $A \times A$ est donc équivalente et on peut facilement voir que :

$$G_{\mathcal{R}_G} = G \quad \text{et} \quad \mathcal{R}_{G_{\mathcal{R}}} = \mathcal{R}$$

Exemple 1.8. Le signe égal sur $\mathbb{R}$ est une relation binaire car on écrit $x = y$ si et seulement si x et y sont identiques. Le sous-ensemble $G_= \subset \mathbb{R} \times \mathbb{R}$ associé à la relation binaire $=$ (c'est à dire ici que $\mathcal{R}$ est le signe $=$) est l'ensemble des paires $(x, y) \in \mathbb{R} \times \mathbb{R}$ telles que $x = y$. C'est à dire c'est la droite d'équation $y = x$ dans $\mathbb{R}^2$, c'est la diagonale $\Delta_{\mathbb{R}}$! On observe donc – mais c'est un cas particulier – que la diagonale de $\mathbb{R}^2$ est à la fois : le graphe $G_{\mathrm{id}_{\mathbb{R}}}$ de la fonction identité, et l'ensemble $G_=$ associé à la relation binaire symbolisant le signe $=$.

Exemple 1.9. Le sous ensemble $G_\leq \subset \mathbb{R}^2$ associé à la relation binaire $\leq$ est la partie du plan au dessus de la diagonale d'équation $y = x$ (incluse). En effet, $(x, y) \in G_\leq$ si $x \leq y$, donc lorsque l'ordonnée y est plus grande ou égale à l'abscisse x donc on est nécessairement au dessus de la diagonale ou dessus.

Exemple 1.10. Soit $G \subset \mathbb{N}^* \times \mathbb{N}^*$ le sous-ensemble de $\mathbb{N}^* \times \mathbb{N}^*$ formé des points rouges. Il s'écrit donc :



$$G = \{(1,1), (1,2), (1,3), \ldots, (2,2), (2,4), (2,6), \ldots,\\ (3,3), (3,6), (3,9), \ldots, (4,4), (4,8), \ldots,\\ (5,5), (5,10), \ldots, \text{etc.}\}$$

Ce sous-ensemble induit une relation binaire sur $\mathbb{N}^*$:

$$m\mathcal{R}_G n \iff (m, n) \in G$$

On voit que $\mathcal{R}_G$ est la relation binaire "m divise n".

Les relations binaires peut parfois avoir des propriétés intéressantes. Soit $\mathcal{R}$ une relation binaire sur A et soit $G_\mathcal{R}$ le sous-ensemble de $A \times A$ associé à cette relation binaire, c'est à dire tous les couples (x, y) d'éléments de A tels que $x \,\mathcal{R}\, y$. On dit que $\mathcal{R}$ est :

- *réflexive* si $\forall\, x \in A,\ x \,\mathcal{R}\, x$, c'est à dire si $\Delta_A \subset G_\mathcal{R}$;
- *symétrique* si $\forall\, x, y \in A,\ x \,\mathcal{R}\, y \implies y \,\mathcal{R}\, x$, c'est à dire que $G_\mathcal{R}$ est symétrique par rapport à la diagonale Δ_A (dans le sens défini plus haut) ;
- *antisymétrique* si $\forall\, x, y \in A,\ (x \,\mathcal{R}\, y \text{ et } y \,\mathcal{R}\, x) \implies x = y$, c'est à dire que la seule partie symétrique de $G_\mathcal{R}$ est sur la diagonale Δ_A ;
- *transitive* si $\forall\, x, y, z \in A,\ (x \,\mathcal{R}\, y \text{ et } y \,\mathcal{R}\, z) \implies x \,\mathcal{R}\, z$.
- *totale* si $\forall\, x, y \in A,\ x \,\mathcal{R}\, y$ ou $y \,\mathcal{R}\, x$

Exemple 1.11. La relation binaire $=$ sur $\mathbb{R}$ a toutes les propriétés sauf la dernière, en particulier elle est symétrique et antisymétrique. La relation $\leq$ sur $\mathbb{R}$ est réflexive, antisymétrique, transitive et totale tandis que la relation $<$ est seulement transitive. La division euclidienne sur $\mathbb{N}^*$ est réflexive, antisymmétrique et transitive. Remarquons que sur $\mathbb{Z}$, la division enclidienne n'est plus antisymétrique car $1|-1$ et $-1|1$ mais $1 \neq -1$.

On vient donc de fournir un dictionnaire entre le domaine de l'algèbre et celui de la géométrie. Cela nous permet d'éclairer différemment certaines notions et de parfois mieux comprendre ou faciliter des preuves. Maintenant, dans notre cas, deux types de relations sont d'importance :

Définition 1.12. *On appelle* relation d'équivalence sur A *tout relation binaire réflexive, symétrique et transitive. On appelle* relation d'ordre sur A *toute relation binaire réflexive, antisymétrique et transitive. Un ensemble muni d'une relation d'ordre est appelé* ensemble ordonné. *Si la relation d'ordre est totale alors l'ensemble est* totalement *ordonné, sinon il est* partiellement *ordonné.*

Exemple 1.13. L'égalité $=$ est une relation d'équivalence sur $\mathbb{R}$. La relation binaire "inférieur ou égal à" $\leq$ est une relation d'ordre sur $\mathbb{R}$. Soit A un ensemble, l'inclusion $\subset$ est une relation d'ordre sur l'ensemble $\mathcal{P}(A)$ des sous-ensembles de A. Sur l'ensemble des nombre entiers naturels, la division est une relation d'ordre. La relation d'ordre $\leq$ est totale, et les deux dernières relations d'ordre sont partielles.

Soit A un ensemble et $\mathcal{R}$ une relation d'équivalence sur A. On appelle *classe d'équivalence* de l'élément $x \in A$ par rapport à la relation d'équivalence $\mathcal{R}$ le sous ensemble $\mathcal{C}_x \subset A$ défini par :

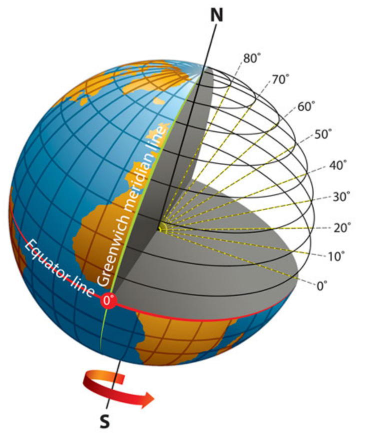
$$\mathcal{C}_x = \{y \in A \text{ tel que } x\mathcal{R}y\}$$

Nous prouverons en exercice que les classes d'équivalences de $\mathbb{R}$ sur A sont soit disjointes soit confondues, c'est à dire :

1. si x et y ne sont pas reliés par la relation d'équivalence $\mathcal{R}$, alors $\mathcal{C}_x \cap \mathcal{C}_y = \emptyset$
2. si $x\mathcal{R}y$, alors $\mathcal{C}_x = \mathcal{C}_y$

Si d'autre part la relation d'équivalence est totale alors les classes d'équivalence $(C_x)_{x\in A}$ forment une *partition* de A, c'est à dire une famille de sous ensemble qui sont deux à deux disjoints mais tels que tout point de A appartient à au moins un de ces sous-ensembles.

Exemple 1.14. Sur la Terre, on peut définir la relation d'équivalence suivante entre deux points de la Terre : un point A et un point B de la Terre sont équivalents si et seulement si ils possèdent la même latitude. Les classes d'équivalence de cette relation d'équivalence sont l'ensemble des points de même latitude, c'est à dire les lignes parallèles sur la Terre. A la latitude de $\pm\frac{\pi}{2}$, les classes d'équivalence sont des points – les pôles Nord et Sud – car il sont seuls dans leur classe d'équivalence : il n'y a aucun autre point équivalent à chacun d'entre eux.



Exemple 1.15. Soit $a \in \mathbb{R}$ et soit $\sim$ la relation d'équivalence définie sur $\mathbb{R}^2$ par la relation suivante :

$$(u,v) \sim (x,y) \quad \text{si et seulement si} \quad y - v = a(x-u)$$

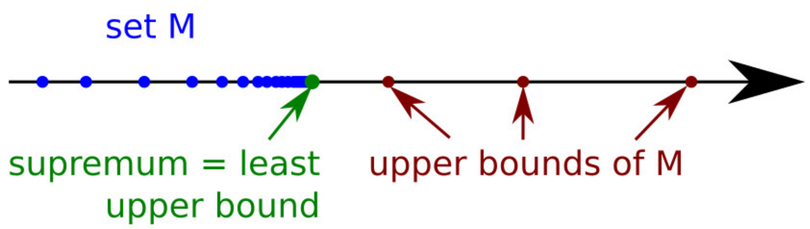
Prenons le point $(0,0)$ de $\mathbb{R}^2$ et décrivons sa classe d'équivalence. C'est l'ensemble des points (x,y) tels que $y-0 = a(x-0)$, c'est à dire la droite $y = ax$, de coefficient directeur a et passant par l'origine $(0,0)$, comme on pouvait s'y attendre. Maintenant, pour un point général $(u,v) \in \mathbb{R}^2$, décrivons sa classe d'équivalence. C'est l'ensemble des points (x,y) tels que $y-v = a(x-u)$, c'est à dire la droite $y = ax + (v - au)$, de coefficient directeur a et d'ordonnée à l'origine $b = v - au$. Cette droite est parallèle à la première, et passe par le point (u,v). Elle est donc confondue avec la première si $(u,v) \sim (0,0)$, ou bien elle est disjointe avec la première si $(u,v) \nsim (0,0)$. Les classes d'équivalence de la relation d'équivalence $\sim$ sont dont les droites parallèles de coefficient directeur a.

Soit A un ensemble ordonné, c'est à dire muni d'une relation d'ordre (notée $\mathcal{R}$), et soit $B \subset A$ un sous-ensemble. Un *majorant* de B (par rapport à la relation d'ordre $\mathcal{R}$) est un élément M de A tel que tout $x \in B$ satisfait $x\mathcal{R}M$. Un *minorant* de B (par rapport à la relation d'ordre $\mathcal{R}$) est un élément m de A tel que tout $x \in B$ satisfait $m\mathcal{R}x$. Un sous-ensemble majoré et minoré est dit *borné*. Supposons que B contient un de ses majorants – disons M – c'est à dire que $M \in B$ et que tout $x \in B$, on a $x\mathcal{R}M$. On appelle M le plus grand élément de B. Il est unique en effet car si jamais il existe $M' \in B$, lui aussi majorant de B, alors nous avons $M\mathcal{R}M'$ par définition car M' est un majorant de B, mais aussi $M'\mathcal{R}M$ car M est un majorant de B. Par antisymétrie de la relation d'ordre, nous avons alors $M = M'$. De façon similaire, on dit qu'un élément de B est le plus petit élément de B si c'est un minorant de B ; il est lui aussi unique. Le plus grand et le plus petit élément n'existent pas forcément !

Exemple 1.16. Prenons $A = \mathbb{N}^*$ muni de la relation d'ordre partiel $|$ (la division euclidienne). Prenons $B = \{1,2,3\}$. Un minorant $m \in \mathbb{N}^*$ de B (par rapport à la relation d'ordre "division euclidienne") est caractérisé par le fait que m divise tous les éléments de B, c'est à dire $m|1$, $m|2$ et $m|3$. Seulement $m = 1$ satisfait la condition $m|1$ (et automatiquement les autres), donc $m = 1$ est l'unique minorant de B. Comme c'est un élément de B, c'est le plus petit élément de B. Un majorant $M \in \mathbb{N}^*$ de B (par rapport à la relation d'ordre "division euclidienne") est caractérisé par la propriété suivante : $\forall n \in B,\ n|M$. Autrement dit, M est divisible par 1 (c'est toujours le cas), 2 et 3, c'est donc un multiple de 6. Les majorants de B sont donc 6, 12, 18, 24, etc. Aucun de ces multiples de 6 n'appartient à B donc B n'a pas de plus grand élément. Si on prend $C = B \cup \{6\}$ le sous-ensemble C a les mêmes majorants et minorants que B, mais comme $6 \in C$, alors C possède un plus grand élément.

Exemple 1.17. Prenons $A = \mathbb{R}$ muni de la relation d'ordre total $\leq$. Prenons $B = [0, 1[$. Un minorant de l'ensemble $[0, 1[$ est un nombre réel m inférieur ou égal à tout élément de $[0, 1[$, donc nécessairement $m \leq 0$. L'ensemble des minorants de $[0, 1[$ est donc $]-\infty, 0]$. Comme 0 est dans $[0, 1[$, on déduit que 0 est le plus petit élément de $[0, 1[$. Un majorant de l'ensemble $[0, 1[$ est un nombre réel M supérieur ou égal à tout élément de $[0, 1[$, donc nécessairement $1 \leq M$. L'ensemble des majorants de $[0, 1[$ est donc $[1, +\infty[$. Aucun de ces majorants n'appartient à $[0, 1[$ donc $[0, 1[$ n'a pas de plus grand élément.

Bien entendu, tout élément de A qui est plus grand – par rapport à la relation d'ordre $\mathcal{R}$ – qu'un majorant de B est lui aussi un majorant de B (par transitivité de la relation binaire). L'ensemble des majorants de B forme donc un sous-ensemble de A. Si l'ensemble des majorants de B admet un plus petit élément, il est appelé *borne supérieure* de B et noté $\sup(B)$. Similairement, tout élément plus petit qu'un minorant de B est lui aussi un minorant de B. L'ensemble des minorants de B forme donc un sous-ensemble de A. Si l'ensemble des minorants de B admet un plus grand élément, il est appelé *borne inférieure* de B et noté $\inf(B)$. Si A est totalement ordonné, on dit que A "a la propriété de la borne supérieure" si tout sous-ensemble non vide majoré de A a une borne supérieure.



Exemple 1.18. Reprenons le contexte de l'exemple 1.16. L'ensemble des minorants de $B = \{1, 2, 3\}$ est réduit à un élément $\{1\}$. Un majorant de 1 est un nombre entier naturel M tel que $1|M$. C'est à dire n'importe quel entier naturel, 1 compris. L'ensemble des majorants de $\{1\}$ est donc $\mathbb{N}^*$, qui contient 1. 1 est donc le plus grand élément de $\{1\}$ par rapport à la relation d'ordre partiel "division euclidienne" donc la borne inférieure de $B = \{1, 2, 3\}$. L'ensemble des majorants de $B = \{1, 2, 3\}$ est l'ensemble $E = \{6, 12, 18, 24, \ldots\}$ des multiples de 6. Les minorants de E (par rapport à la division euclidienne) sont les entiers naturels qui divisent tous les éléments de E, c'est à dire les nombres qui divisent 6 car tous les éléments de E sont des multiples de 6. Les diviseurs de 6 sont 1,2,3 et 6. Le nombre 6 – minorant de E – appartient à E donc c'est le plus petit élément de E. On en déduit que la borne supérieure de B est 6. On a donc $\inf(B) = 1 \in B$ et $\sup(B) = 6 \notin B$.

Exemple 1.19. Reprenons l'exemple 1.17. L'ensemble des minorants de $B = [0, 1[$ est $]-\infty, 0]$. Le plus grand élément de cet ensemble est 0, donc la borne inférieure de $[0, 1[$ est 0, et on note $\inf([0, 1[) = 0$. Ce nombre appartient à $[0, 1[$. L'ensemble des majorants de $[0, 1[$ est $[1, +\infty[$. Le plus petit élément de cet ensemble est 1, donc la borne supérieure de $[0, 1[$ est donc 1, et on note

$\sup([0,1[) = 1$. Ce nombre n'appartient pas à $[0,1[$. On remarque que la borne inférieure et la borne supérieure sont à la frontière de l'intervalle $[0,1[$.

Exemple 1.20. Prenons comme exemple les ensembles $A = \mathbb{R}$ avec la relation d'ordre usuelle $\leq$. Posons $B = \left\{\frac{1}{n},\text{ pour } n \in \mathbb{N}^*\right\} = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots\right\}$. Un majorant de B est n'importe quel nombre réel supérieur ou égal à 1. L'ensemble des majorants de B est donc $[1, +\infty[$. Cette ensemble des majorants a un plus petit élément – c'est 1 – qui est donc la borne supérieure de B, et on la note $\sup(B) = 1$. Nous voyons en outre qu'elle est un élément de B, c'est donc le plus grand élément de B. Un minorant de B est inférieur ou égal à 0. L'ensemble des minorants de B est donc $]-\infty, 0]$. Cet ensemble des minorants a un plus grand élément – c'est 0 – qui est donc la borne inférieure de B, et on écrit $\inf(B) = 0$. Par contre on voit que la borne inférieure n'est pas un élément de B ! L'ensemble B n'a pas de plus petit élément (ça se voit).

1.2 Nombres réels

Les nombres réels ne sont pas bien définis avant le XIXème siècle. En 1820 dans son cours pour l'Ecole Polytechnique, Cauchy se base sur "l'évidence géométrique" : l'ensemble des nombres réels consiste à une droite infinie d'un point de vue topologique, et quand on zoome dessus on ne voit jamais de trou. Ceci permet de faire de l'analyse de manière intuitive mais ce n'est pas rigoureux mathématiquement. Ce n'est qu'autour des années 1860 avec les travaux de Dedekind et Cantor que la définition topologique des nombres réels se fait rigoureusement. Dedekind construit $\mathbb{R}$ comme l'ensemble des *coupures de* $\mathbb{Q}$ (voir plus bas).

L'ensemble des nombre réels possède en outre des propriétés algébriques : c'est un corps totalement ordonné (et donc implicitement compatible avec les deux opérations). Pour comprendre algébriquement le crops des nombres réels on peut commencer par les entiers $\mathbb{N}$ qu'on muni de l'addition $+$. Si on admet les inverse par rapport à l'addition on obtient les entiers relatifs $\mathbb{Z}$ (c'est un groupe abélien/commutatif), avec 0 comme élément neutre de l'addition. On peut ajouter la multiplication $\times$ – avec 1 comme élément neutre de la multiplication – et $\mathbb{Z}$ devient alors un anneau. Si on ajoute ensuite tous les inverses $q \longrightarrow 1/q$ et leurs multiples, on obtient le corps des fractions – ou nombres rationnels, noté $\mathbb{Q}$). Un corps est un anneau c'est à dire un ensemble muni d'une opération $+$ et d'une opération $\times$ (les deux sont commutatives) dont tous les éléments sauf 0 sont inversibles par rapport à la multiplication. Il existe un corps de nombres – appelé nombre réels et noté $\mathbb{R}$ – qui est la plus petite extension topologique de $\mathbb{Q}$ dans le sens où $\mathbb{R}$ n'a pas de "trous". Les nombres de $\mathbb{R}\backslash\mathbb{Q}$ sont appelés nombres irrationnels La construction de $\mathbb{R}$ est difficile à axiomatiser donc on ne la donne pas.

Pour énoncer les propriétés du corps des réels, nous devons affiner ce que nous savons sur les relations d'ordre. La relation d'ordre "être inférieure ou égale à" sur dénotée $\leq$ est une relation d'ordre *total* car deux éléments $x, y \in \mathbb{R}$ sont toujours ordonnés : soit $x \leq y$ ou bien $y \leq x$. D'autre part, cette relation ensembliste a d'autre part les propriétés suivantes de compatibilité avec les opérations algébriques addition et multiplication :

— pour tout $x \leq y$ et $w \leq z$, on a $x + w \leq y + z$, et

— pour tout $x \leq y$ et $c > 0$, on a $cx \leq cy$.

Le corps $\mathbb{R}$ est totalement ordonné, et de même que le corps des nombres rationnels $\mathbb{Q}$. Les corps des nombres rationnels et des nombres réels sont aussi Archimédiens :

Lemme 1.21. ***Propriété d'Archimède*** *: pour tous,* $0 < x < y$*, il existe* $n \in \mathbb{N}$ *tel que* $y < nx$*.*

Cela a comme conséquence d'une part que le corps des réels (et celui des rationnels) possède la division euclidienne avec quotient entier. Soit $x, y \in \mathbb{R}$ avec $y > 0$, il existe $q \in \mathbb{Z}$ unique tel que $qy \leq x < (q + 1)y$. L'autre conséquence est que la fonction partie entière d'un nombre réel est bien définie.

Définition 1.22. *Soit $x \in \mathbb{R}$, on appelle* **partie entière de** *x et on dénote $E(x)$ le plus grand entier relatif inférieur ou égal à x :*

$$E(x) \leq x < E(x) + 1$$

La partie entière définie une fonction par escalier $E : \mathbb{R} \to \mathbb{N}, x \mapsto E(x)$, qu'on appelle fonction partie entière*.*

La valeur absolue d'un réel x est notée $|x|$, si x est positif, $|x| = x$ et si x est négatif, $|x| = -x \geq 0$. On appelle la distance entre x et y quelconques, la valeur absolue $|y - x|$. Pour tout réel strictement positif r on a :

$$|x| \leq r \quad \Longleftrightarrow \quad -r \leq x \leq r$$

Voici deux propriétés fondamentales de la valeur absolue, pour tout $x, y \in \mathbb{R}$:

$$|xy| = |x||y|$$
$$|x + y| \leq |x| + |y| \qquad \text{inégalité triangulaire}$$

Le module des nombres complexes a les mêmes propriétés. L'inégalité triangulaire prend sens dans $\mathbb{C}$ (dessiner ce que cela veut dire). On a d'autre part l'inégalité suivante, parfois utile :

$$\forall\, x, y \in \mathbb{R} \quad \big||x| - |y|\big| \leq |x - y|$$

Soit $a < b \in \mathbb{R}$, le *segment* – ou *intervalle fermé* – $[a, b]$ est l'ensemble des réels x tels que $a \leq x \leq b$:

$$[a, b] = \{x \in \mathbb{R} \text{ tel que } a \leq x \leq b\}$$

L'*intervalle ouvert* $]a, b[$ est l'ensemble des réels x tels que $a < x < b$

$$[a, b] = \{x \in \mathbb{R} \text{ tel que } a < x < b\}$$

Si c'est fermé à l'une des extremités on parle d'intervalle semi-ouvert (ou semi-fermé). Soit $a \in \mathbb{R}$, on peut aussi définir les deux *demi-droites ouvertes* suivantes :

$$] - \infty, a[= \{x \in \mathbb{R} \text{ tel que } x < a\} \quad \text{et} \quad]a, +\infty[= \{x \in \mathbb{R} \text{ tel que } x > a\}$$

Les *demi-droites fermées* sont similairement données par :

$$] - \infty, a] = \{x \in \mathbb{R} \text{ tel que } x \leq a\} \quad \text{et} \quad [a, +\infty[= \{x \in \mathbb{R} \text{ tel que } x \geq a\}$$

Attention les symboles $+\infty$ et $-\infty$ ne sont pas des éléments de $\mathbb{R}$. On considère que les demi-droites ouvertes sont des intervalles ouverts, et les demi-droites fermées sont des intervalles semi)ouverts. Dans la suite un intervalle I pourra désigner un segment, un intervalle ouvert ou un intervalle semi-ouvert. Nous allons maintenant donner une caractérisation topologique des intervalles de $\mathbb{R}$, qui veut dire qu'ils ne sont pas *discontinus*.

Définition 1.23. *Soit $A \subset \mathbb{R}$ un sous-ensemble de $\mathbb{R}$. On dit que A est* convexe *si, pour tout $x < y$ deux éléments de A le segment $[x, y]$ est inclus dans A, i.e. si $\forall\, x < y \in A,\ [x, y] \subset A$.*

Exemple 1.24. Le sous-ensemble $A = [0, 1[\cup\{2\}$ n'est pas convexe car si on prend $x = 0$ et $y = 2$, deux points de A tels que $x < y$, on observe que le segment $[x, y]$ n'est pas inclus dans A. On voit que le problème vient de la discontinuité de A. De même $\mathbb{Q}$ n'est pas convexe.

Proposition 1.25. *Les intervalles sont les convexes de $\mathbb{R}$, c'est à dire que tout intervalle est convexe, et que si un sous-ensemble de $\mathbb{R}$ est convexe alors c'est un intervalle.*

Nous avons besoin des notions topologiques suivantes, qui seront très utiles par la suite. On appelle *voisinage du point* $x \in \mathbb{R}$ tout sous-ensemble V de $\mathbb{R}$ qui contient un intervalle ouvert contenant x, autrement dit :

$$V \text{ est un voisinage de } x \quad \Longleftrightarrow \quad \exists\, \delta > 0 \text{ tel que }]x-\delta, x+\delta[\subset V$$

On appelle *voisinage de* $+\infty$ (resp $-\infty$) tout sous-ensemble V de $\mathbb{R}$ qui contient un intervalle ouvert avec borne supérieure $+\infty$ (resp. $-\infty$), autrement dit :

$$\begin{array}{lcl} V \text{ est un voisinage de } +\infty & \Longleftrightarrow & \exists\, M \in \mathbb{R} \text{ tel que }]M, +\infty[\subset V \\ V \text{ est un voisinage de } -\infty & \Longleftrightarrow & \exists\, m \in \mathbb{R} \text{ tel que }]-\infty, m[\subset V \end{array}$$

Notons que la notion de voisinage de $+\infty$ est valide aussi sur $\mathbb{N}$ si on considère les ensembles du type $\mathbb{N} \cap]E(M), +\infty[= \{E(M)+1, E(M)+2, E(M)+3, \ldots\}$.

Exemple 1.26. Par exemple, $A = [0, 1[\cup\{2\}$ n'est pas un voisinage de 0, de 1, ni de 2, mais c'st un voisinage de tous les points du segment ouvert $]0, 1[$. Tout d'abord $1 \notin A$ donc d'après la définition, A ne peut pas être un voisinage de 1. Pour 0 et 2, quelque soit $\epsilon > 0$, les intervalles $]-\epsilon, +\epsilon[$ et $]2-\epsilon, 2+\epsilon[$ ne sont pas inclus dans A donc d'après la définition A n'est pas un voisinage de 0 et 2. Maintenant soit $x \in]0, 1[\subset A$, alors posons $\epsilon = \min\left(\frac{x}{2}, \frac{1-x}{2}\right)$. Dans ce cas, l'intervalle $]x-\epsilon, x+\epsilon[$ est strictement inclus dans $]0, 1[$ donc dans A, donc A est un voisinage de x. L'ensemble A est donc un voisinage de chacun des points de $]0, 1[$, mais pas de chacun de ses points car n'est pas un voisinage de 0 et 2.

Définition 1.27. *Un sous-ensemble A de $\mathbb{R}$ est dit* ouvert *si il est le voisinage de chacun de ses points, i.e. si pour tout $x \in A$, il existe $\epsilon > 0$ tel que $]x-\epsilon, x+\epsilon[\subset A$. Un sous-ensemble B de $\mathbb{R}$ est dit* fermé *si son complémentaire B^c est ouvert.*

Exemple 1.28. L'ensemble $A = [0, 1[\cup\{2\}$ n'est pas ouvert car A n'est pas un voisinage de 0 et 2, par contre nous avons démontré que $]0, 1[$ est un voisinage de chacun de ses points dans l'exemple 1.26. L'intervalle ouvert $]0, 1[$ est donc un ouvert. Plus généralement, un intervalle ouvert est ouvert au sens topologique (ci dessus), un segment est fermé au sens topologique (ci-dessus).

Remarque 1.29. L'ensemble des sous-ensembles ouverts de $\mathbb{R}$ est appelé la *topologie* de $\mathbb{R}$ et est notée $\mathcal{T}(\mathbb{R})$. Par convention on pose que l'ensemble vide $\emptyset$ est ouvert. En particulier, nous avons que :

$$\mathcal{T}(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$$

C'est à dire que la topologie est un sous-ensemble de l'ensemble des parties de $\mathbb{R}$.

Le corps des nombres réels $\mathbb{R}$ possède la propriété suivante, que le corps des rationnels $\mathbb{Q}$ ne possède pas (on ne la prouvera pas) :

Propriété des segments emboîtés. *S'il existe une suite de segments $I_1, I_2, I_3, \ldots, I_n, \ldots$ de $\mathbb{R}$ telle que pour tout $n \in \mathbb{N}$, on a $I_{n+1} \subset I_n$, c'est à dire telle que :*

$$\ldots \subset I_{n+1} \subset I_n \subset I_{n-1} \subset \ldots \subset I_2 \subset I_1 \subset I_0,$$

alors il existe $x \in \mathbb{R}$ tel que $x \in I_n$, pour tout $n \in \mathbb{N}$ (tous les segments I_n ont au moins un points en commun).

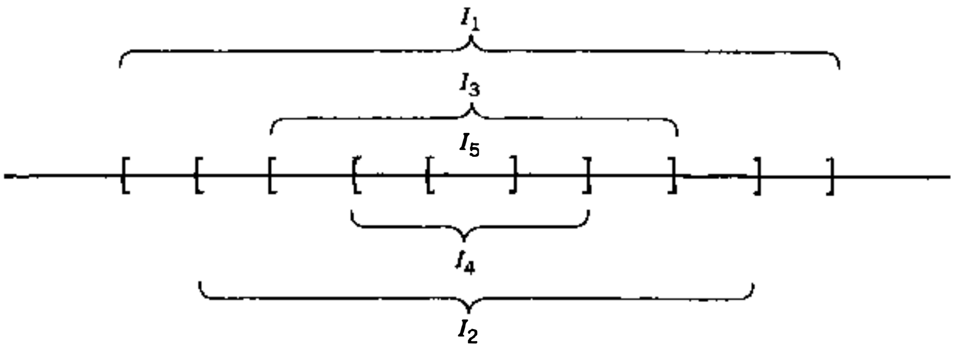


Figure 2.5.1 Nested intervals

Attention la propriété ne compte pas si les intervalles I_n sont semi-ouverts. Par exemple si on prend $I_n =]0, \frac{1}{n}]$ pour tout $n \geq 1$, alors ces intervalles semi-ouverts sont emboîtés, c'est à dire que $I_{n+1} \subset I_n$. Cependant leur intersection est l'ensemble vide, et on note :

$$\bigcap_{n=0}^{+\infty} I_n = \emptyset$$

En effet, pour tout $x \in]0, 1]$, il existe $N \in \mathbb{N}$ tel que $\frac{1}{N} < x$. Dans ce cas, pour tout $n \geq N$, l'intervalle I_n ne contient pas x. La propriété des segments emboités nous permet de montrer le résultat important suivant :

Théorème 1.30. *Les nombres réels sont indénombrables.*

Démonstration. Les personnes intéressées peuvent aller voir la preuve (très belle) page 49 du livre de Bartle et Sherbert, *Introduction to Real Analysis* (4th Edition). □

De cela nous déduisons que les nombres irrationnels $\mathbb{R} \backslash \mathbb{Q}$ sont indénombrables, car s'ils étaient dénombrables, alors $\mathbb{Q} \cup \mathbb{R} \backslash \mathbb{Q} = \mathbb{R}$ le serait aussi comme union d'ensembles dénombrables, ce qu'il n'est pas. Nous avons en outre le joli résultat suivant :

Proposition 1.31.
— *$\mathbb{Q}$ est **dense** dans $\mathbb{R}$: dans tout intervalle ouvert non vide de $\mathbb{R}$, il y a une infinité (dénombrable) de nombres rationnels.*
— *$\mathbb{R} \backslash \mathbb{Q}$ est **dense** dans $\mathbb{R}$: dans tout intervalle ouvert non vide de $\mathbb{R}$, il y a une infinité (indénombrable) de nombres irrationnels.*

Démonstration. Il suffit de faire la preuve sur un intervalle ouvert $]a, b[$, où $a < b$. Par la propriété d'Archimède, il existe $q \in \mathbb{N}$ tel que $0 < 1 < q(b - a)$. Notons $p = E(qa) + 1$. On a donc $qa < p \leq qa + 1 < qa + q(b - a) = qb$ donc il vient que $qa < p < qb$ donc $a < p/q < b$. Il existe un rationnel dans $]a, b[$.

Du fait que tout intervalle ouvert non vide contient un rationnel, appliquons ce résultat sur l'intervalle ouvert $]a - \sqrt{2}, b - \sqrt{2}[$. Il existe un rationnel $r \in \mathbb{Q}$ tel que $r \in]a - \sqrt{2}, b - \sqrt{2}[$. On a donc $a < r + \sqrt{2} < b$, mais comme le nombre $\sqrt{2}$ est irrationnel, et que la somme d'un rationnel et d'un irrationnel est irrationnel, on déduit qu'un irrationnel est dans l'intervalle ouvert $]a, b[$.

Maintenant montrons qu'il y en a une infinité de chaque dans l'intervalle. Soit $N \in \mathbb{N}$. Posons $l = b - a > 0$ et partageons l'intervalle $]a, b[$ en N intervalles de longueur l/N (avec coupures) :

$$\left]a, a + \frac{l}{N}\right[\ , \ \left]a + \frac{l}{N}, a + \frac{2l}{N}\right[\ , \ \left]a + \frac{2l}{N}, a + \frac{3l}{N}\right[\ , \ \dots \ , \ \left]a + \frac{(N-1)l}{N}, b\right[$$

Dans chacun de ces intervalles, il y a au moins un rationnel et un irrationnel. et ces intervalles sont disjoints deux à deux. Il y a donc au moins N rationnels et N irrationnels dans l'intervalle $]a, b[$. Ceci étant vrai quel que soit l'entier N, on en déduit qu'il y a une infinité (dénombrable) de nombre rationnels et de nombre irrationnels dans l'intervalle. $\square$

D'autre part, la propriété des segments emboités est équivalente à la propriété très importante suivante (que nous ne démontrerons pas), possédée par $\mathbb{R}$ et pas par $\mathbb{Q}$:

Propriété de la borne supérieure (ou inférieure). *Toute partie non vide majorée admet une borne supérieure. De façon équivalente, toute partie non vide minorée admet une borne inférieure.*

Soit A un sous ensemble non-vide de $\mathbb{R}$. Un majorant (resp. minorant) de A est un réel M (resp. m) tel que $\forall\ a \in A, a \leq M$ (resp. $m \leq a$). Le plus petit (resp. grand) des majorants (resp. minorants) de A – s'il existe – est appelé la **borne supérieure** (resp. **inférieure**) de A et noté $\sup(A)$ (resp. $\inf(A)$). Si A est majoré, l'ensemble des majorants de A forme la demi-droite fermée à gauche $[\sup(A), +\infty[$. Si A est minoré, l'ensemble des minorants de A forme la demi-droite fermée à droite $]-\infty, \inf(A)]$.

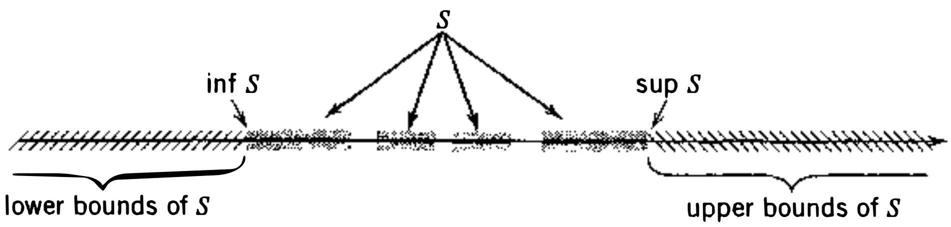


Figure 2.3.1 inf S and sup S

Proposition 1.32. *Soit A un sous-ensemble non-vide majoré de $\mathbb{R}$. La borne supérieure* $\sup(A)$ *est l'unique nombre réel caractérisé par les deux propriétés suivantes :*

— $\forall\, x \in A,\ x \leq \sup(A)$ *(la borne supérieure est un majorant de A),*

— $\forall\, \epsilon > 0,\ \exists\, x \in A,\ \sup(A) - \epsilon \leq x$ *(c'est le plus petit majorant de A)*

Exemple 1.33. La partie $A = \left\{1 - \frac{1}{n},\ n \in \mathbb{N}^*\right\}$ de borne supérieure $\sup(A) = 1$.

La propriété de la borne supérieure, et de façon équivalente celle des segments emboîtés, est aussi équivalente au deux propriété suivantes, qui illustrent la *complétude topologique* de $\mathbb{R}$: c'est une droite qui n'a pas de trous, contrairement à $\mathbb{Q}$ qui est *lacunaire/discret.*

— **Propriété des coupures de Dedekind** Si (A, B) forme une partition de $\mathbb{R}$, de façon que $\forall\, a \in A$ et $\forall\, b \in B$, on a $a < b$ alors il existe un élément $x_0 \in \mathbb{R}$ tel que :

  - ou bien $A = \{a \in \mathbb{R} \mid a \leq x_0\}$ et $B = \{b \in \mathbb{R} \mid x_0 < b\}$

  - ou bien $A = \{a \in \mathbb{R} \mid a < x_0\}$ et $B = \{b \in \mathbb{R} \mid x_0 \leq b\}$

— **Propriété de convergence des suites de Cauchy** Les suites de Cauchy convergent, i.e. admettent une limite dans $\mathbb{R}$.

Nous n'utiliserons pas les coupures de Dedekind, mais nous expliquerons plus tard ce que sont les suites de Cauchy et leur lien à la convergence.

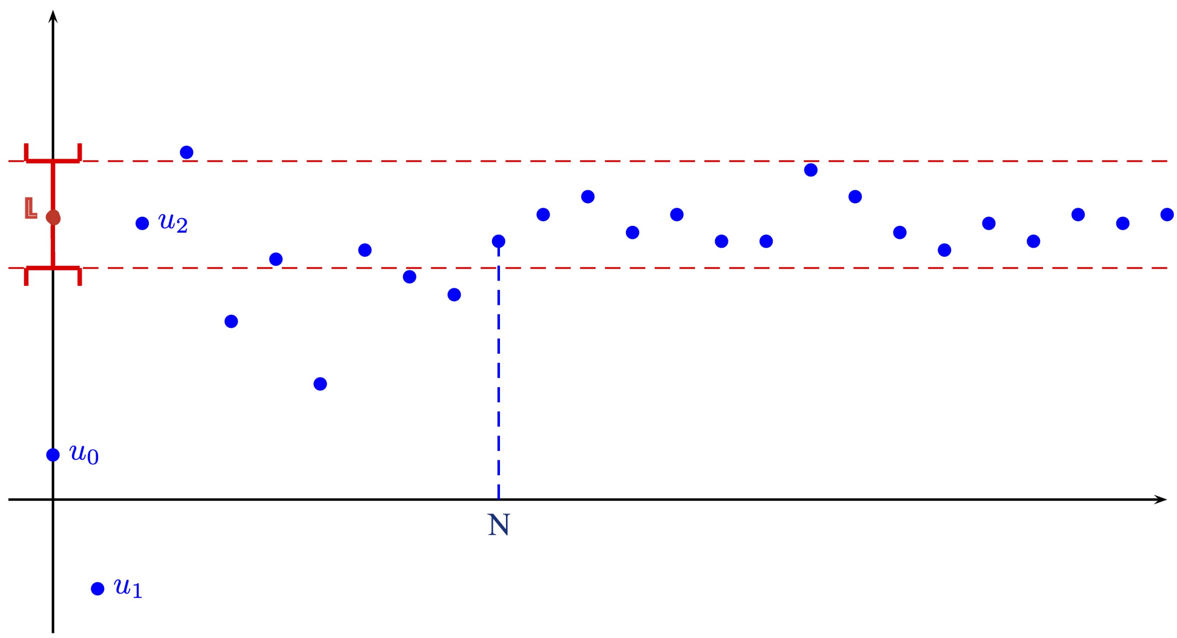


FIGURE 1 – A partir d'un rang N tous les termes de la suite sont dans l'intervalle $]\ell-\epsilon, \ell+\epsilon[$.

1.3 Suites, limites et comparaisons

Dans la suite, si on note $\mathbb{K}$ c'est pour signifier $\mathbb{R}$ ou $\mathbb{C}$. L'étude rigoureuse des suites date du début XIXème siècle avec le cours de Cauchy à l'Ecole Polytechnique. La définition de la convergence à l'aide de ϵ date des années 1860 avec les travaux de Weierstrass.

Définition 1.34. *Une **suite** est une application $u : \mathbb{N} \to \mathbb{K}, n \mapsto u_n$ qui à chaque entier associe un nombre réel ou un nombre complexe. On la note habituellement $(u_n)_{n\in\mathbb{N}}$ ou juste $(u_n)_n$, et on appelle u_n le terme général.*

Remarque 1.35. Attention, ne confondez pas $(u_n)_n$ la suite et u_n le terme général de rang n. La suite est une application de $\mathbb{N}$ dans $\mathbb{K}$, tandis que le terme général est un nombre de $\mathbb{K}$.

On dit qu'une suite réelle (ou complexe) converge vers un nombre ℓ réel (ou complexe) si, intuititvement, la suite se rapproche de ce nombre en terme de distance lorsque n devient de plus en plus grand (tend vers l'infini).

Définition 1.36. *Soit $u : \mathbb{N} \to \mathbb{K}$ une suite réelle ou complexe et soit $\ell \in \mathbb{K}$. On dit que $u = (u_n)_n$* converge vers *ℓ ou* tend vers *ℓ* (quand *n* tend vers l'infini) *si :*

$$\forall\, \epsilon > 0,\ \exists\, N \in \mathbb{N} \textit{ tel que } \forall\, n \geq N,\ |u_n - \ell| < \epsilon$$

Dans ce cas on appelle ℓ la limite *de la suite $(u_n)_n$ et on note :*

$$u_n \xrightarrow[n\to+\infty]{} \ell \quad \textit{ou encore} \quad \lim_{n\to+\infty} u_n = \ell$$

On dit qu'une suite $(u_n)_n$ est convergente *si elle converge vers une limite $\ell \in \mathbb{K}$; on dit qu'elle est divergente sinon.*

Exemple 1.37. Prenons la suite $u : \mathbb{N} \to \mathbb{R}, n \mapsto \frac{1}{n}$. Montrons qu'elle converge vers 0 avec la définition. Soit $\epsilon > 0$, posons $N = E\left(\frac{1}{\epsilon}\right) + 1$. Alors les propriétés de la fonction partie entière sont telles que :

$$\frac{1}{\epsilon} < E\left(\frac{1}{\epsilon}\right) + 1 \qquad \text{c'est à dire} \qquad \frac{1}{\epsilon} < N$$

Ce qui donne donc, pour tout $n \geq N$:

$$-\epsilon < 0 < \frac{1}{n} < \frac{1}{N} < \epsilon$$

On a donc, pour tout $n \geq N$, $|u_n| < \epsilon$, ce qui montre que la suite $(u_n)_n$ converge vers 0.

Remarque 1.38. Pour une suite complexe $(u_n)_n$ convergente vers une limite $\ell \in \mathbb{C}$, la partie réelle (resp. complexe) de u_n tend vers la partie réelle (resp. complexe) de ℓ, c'est à dire :

$$u_n \xrightarrow[n \to +\infty]{} \ell \quad \iff \quad \begin{cases} \mathrm{Re}(u_n) & \xrightarrow[n \to +\infty]{} \mathrm{Re}(\ell) \\ \mathrm{Im}(u_n) & \xrightarrow[n \to +\infty]{} \mathrm{Im}(\ell) \end{cases}$$

La convergence d'une suite complexe se ramène donc à l'étude de la convergence des suites réelles $(\mathrm{Re}(u_n))_n$ et $(\mathrm{Im}(u_n))_n$. Lorsqu'il est question de convergence, nous pouvons donc nous concentrer sur l'étude des suites réelles, c'est à dire $\mathbb{K} = \mathbb{R}$ sauf exception.

Remarque 1.39. Pour les suites réelles, la phrase mathématique dans la définition de la convergence peut se lire comme suit : la suite réelle $(u_n)_n$ converge vers $\ell \in \mathbb{R}$ si $u : \mathbb{N} \to \mathbb{R}$ envoie tout voisinage de $+\infty$ (dans $\mathbb{N}$) dans un voisinage de ℓ (dans $\mathbb{R}$). La notion de convergence est une notion *asymptotique*, cela veut dire qu'on ne peut pas savoir si une suite converge en ne regardant qu'un nombre fini de termes. C'est en regardant la partie de la suite dont le rang est proche de l'infini qu'on peut seulement savoir si la suite converge. Couper les N premiers termes d'une suite ne change ainsi pas son caractère convergent ou divergent.

Proposition 1.40. *La limite d'une suite convergente est unique.*

Démonstration. Comme dit plus haut, on peut se ramener au cas des suites réelles. Soit $u : \mathbb{N} \to \mathbb{R}$ une suite et soit $\ell, \ell' \in \mathbb{K}$. On veut montrer que si j'amais $u_n \xrightarrow[n \to +\infty]{} \ell$ et $u_n \xrightarrow[n \to +\infty]{} \ell'$, alors $\ell = \ell'$. Par l'absurde : on suppose que $(u_n)_n$ converge vers les deux limites, mais que $\ell \neq \ell'$, de façon à ce que $|\ell - \ell'| > 0$. Soit $\epsilon = \frac{|\ell - \ell'|}{2}$. Il suffit ensuite d'écrire la définition de la convergence. Comme u_n tend vers ℓ, il existe N tel que $|u_n - \ell| < \epsilon$ pour tout $n \geq N$ et comme u_n tend vers ℓ', il existe N' tel que $|u_n - \ell'| < \epsilon$ pour tout $n \geq N'$. Notons $\mathcal{N} = \max(N, N')$. Alors, pour tout $n \geq \mathcal{N}$, on a les deux inégalités : $|u_n - \ell| < \epsilon$ et $|u_n - \ell'| < \epsilon$. On a alors, par l'inégalité triangulaire :

$$|\ell - \ell'| \leq |u_n - \ell| + |u_n - \ell'| < \epsilon + \epsilon = 2\epsilon = |\ell - \ell'|$$

donc $|\ell - \ell'| < |\ell - \ell'|$, ce qui est absurde. $\square$

Exemple 1.41. Voici les termes généraux de plusieurs exemple de suites convergentes et divergentes :

$$u_n = \cos(n) \qquad u_n = \frac{1}{n} \qquad u_n = \sin\left(\frac{1}{n}\right)$$

$$u_n = e^{i\frac{n2\pi}{3}} \qquad u_n = \left(1 + \frac{1}{n}\right)^n \qquad u_n = \sum_{k=1}^{n} \frac{1}{k}$$

La première est divergente (pas de limite) car cosinus est périodique et $\cos(n)$ se comporte comme $\cos(x)$. La deuxième et la troisième convergent vers 0 car sin est continue. La quatrième est périodique donc divergente (dessiner). La cinquième tend vers le nombre d'Euler (constante de Napier) e et la cinquième diverge vers $+\infty$. Notons que la suite de terme général $v_n = \sum_{k=1}^{n} \frac{1}{k} - \ln(n)$ converge vers un nombre réel, noté γ et appelé constante d'Euler-Mascheroni. Elle vaut environ $0,5772156649$ et on ne sait toujours pas si elle est rationnelle ou irrationnelle ! !

Nous avons les résultats (évidents ?) sur les suites convergentes et leurs limites : si $(u_n)_n$ converge vers $\ell \in \mathbb{K}$ et $(v_n)_n$ converge vers $\ell' \in \mathbb{K}$ alors :

— la suite de terme général $a_n = u_n + v_n$ converge vers $\ell + \ell'$ et celle de terme général $b_n = u_n v_n$ converge vers $\ell\ell'$;

— la suite $(|u_n|)_n$ converge vers $|\ell|$ et la suite $(\overline{u_n})_n$ (complexe conjugué dans $\mathbb{C}$) converge vers $\overline{\ell}$;

Exemple 1.42. Soit $(u_n)_n$ la suite de terme général $u_n = \frac{2n+1}{3n+5}$. Ceci s'écrit aussi $u_n = \frac{2+\frac{1}{n}}{3+\frac{5}{n}}$, et si on écrit $v_n = 2 + \frac{1}{n}$ et $w_n = \frac{1}{3+\frac{5}{n}}$, alors la suite $(v_n)_n$ converge vers 2 et la suite $(w_n)_n$ converge vers $\frac{1}{3}$. Dans ce cas la suite $(u_n)_n$ converge vers $\frac{2}{3}$.

Maintenant étudions le cas où la limite est en l'infini (réel ou complexe).

Définition 1.43. *Dans $\mathbb{K} = \mathbb{R}$, on dit que la suite réelle $(u_n)_n$* tend vers $+\infty$ *si on a :*

$$\forall M \in \mathbb{R},\ \exists N \in \mathbb{N} \text{ tel que } \forall n \geq N,\ M \leq u_n$$

Dans $\mathbb{K} = \mathbb{R}$, on dit que la suite $(u_n)_n$ tend vers $-\infty$ *si on a :*

$$\forall m \in \mathbb{R},\ \exists N \in \mathbb{N} \text{ tel que } \forall n \geq N,\ u_n \leq m$$

Dans $\mathbb{K} = \mathbb{C}$, on dit que la suite complexe $(u_n)_n$ tend vers l'infini *lorsqu'on est arbitrairement grand en module :*

$$\forall M > 0,\ \exists N \in \mathbb{N} \text{ tel que } \forall n \geq N,\ |u_n| > M$$

Remarque 1.44. Dans $\mathbb{C}$ il n'y a pas de relation d'ordre donc cela n'a pas de sens d'écrire $+\infty$ et $-\infty$, car on peut tendre vers l'infini dans toutes les directions, voire même en spirale, par exemple prendre la suite complexe de terme général $u_n = ne^{in}$.

Remarque 1.45. Autrement dit, la suite réelle $(u_n)_n$ tend vers $+\infty$ (resp. $-\infty$) si $u : \mathbb{N} \to \mathbb{R}$ envoie tout voisinage de $+\infty$ (dans $\mathbb{N}$) dans un voisinage de $+\infty$ (resp. $-\infty$).

On a donc les propriétés suivantes vis à vis de la convergence en zéro et la divergence en l'infini : Soit $(u_n)_n$ une suite à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, dont on suppose qu'elle ne s'annule jamais, c'est à dire que $u_n \neq 0$ pour tout $n \in \mathbb{N}$, ou du moins qu'elle ne s'annule jamais à partir d'un certain rang. Nous avons les trois observations suivantes :

— Si la suite $(u_n)_n$ converge vers $\ell \neq 0$, la suite $\left(\frac{1}{u_n}\right)_n$ converge vers $1/\ell$;

— Si la suite $(u_n)_n$ converge vers $0 \in \mathbb{K}$, alors on a que la suite $\left(\frac{1}{u_n}\right)_n$ diverge, et plus précisément la suite $\left(\frac{1}{|u_n|}\right)_n$ tend vers $+\infty$;

— Inversement, si la suite $(u_n)_n$ tend vers l'infini quand n tend vers l'infini, on a que la suite $\left(\frac{1}{u_n}\right)_n$ converge vers $0 \in \mathbb{K}$.

On a aussi les résultats suivants, à connaitre par coeur :

Proposition 1.46. *Soit $q \in \mathbb{R}$, alors :*

si $|q| > 1$, $(q^n)_n$ diverge, si $q = 1$, $(q^n)_n$ est constante $= 1$, et si $q < 1$, $(q^n)_n$ converge vers 0.

Soit $\alpha \in \mathbb{R}$, alors :

si $\alpha > 0$, $n^\alpha \xrightarrow[n\to+\infty]{} +\infty$, si $\alpha = 0$, $(n^\alpha)_n$ est constante $= 1$, et si $\alpha < 0$, $(n^\alpha)_n$ converge vers 0.

Dans $\mathbb{R}$ comme il y a une relation d'ordre, nous avons les notions suivantes (noter bien la différence des phrases mathématiques avec la notion de tendre vers $\pm\infty$ de la définition 1.43) :

- la suite $(u_n)_n$ est majorée si $\exists\, M \in \mathbb{R}$ tel que $\forall\, n \in \mathbb{N},\ u_n \leq M$;
- la suite $(u_n)_n$ est minorée si $\exists\, m \in \mathbb{R}$ tel que $\forall\, n \in \mathbb{N},\ m \leq u_n$;
- la suite $(u_n)_n$ est bornée si elle est majorée et minorée. Alternativement on peut écrire de façon équivalente que $(u_n)_n$ est bornée si $\exists\, M \in \mathbb{R}$ tel que $\forall\, n \in \mathbb{N},\ |u_n| \leq M$ (cette définition est aussi valide pour les suites complexes).

Nous avons alors le résultat suivant qui caractérise les suites convergentes :

Proposition 1.47. *Toute suite convergente est bornée.*

Démonstration. On comence pour une suite $(u_n)_n$ qui tend vers 0. Soit $\epsilon = 1$, alors il existe N tel que pour tout $n \geq N$, on a $|u_n - 0| < \epsilon$ c'est à dire $|u_n| < 1$. Posons $M = \max(|u_0|, |u_1|, \ldots, |u_{N-1}|, 1)$. Alors, pour tout $0 \leq k \leq N-1$ nous avons que $|u_k| \leq M$. D'autre part, pour tout $n \geq N$, $|u_n| \leq M$ d'après notre choix de ϵ et de N. Ceci prouve la proposition pour $\ell = 0$.

Pour une limite finie $\ell \in \mathbb{K}$ quelconque, on pose $v_n = u_n - \ell$ pour tout $n \in \mathbb{N}$. Dans ce cas, comme $(u_n)_n$ tend vers ℓ, la limite de v_n est zéro. Donc d'après notre premier point, la suite $(v_n)_n$ est bornée, par un certain $M > 0$ disons. C'est à dire : pour tout $n \in \mathbb{N}$ on a $|v_n| \leq M$, et donc $|u_n - \ell| \leq M$. Cela veut dire que $-M + \ell \leq u_n \leq M + \ell$ pour tout $n \in \mathbb{N}$. La suite $(u_n)_n$ étant majorée et minorée, elle est bornée. Ceci conclut la preuve. □

Remarque 1.48. Ceci est une condition nécessaire, elle n'est absolument pas suffisante puisqu'on a des contres-exemples faciles : par exemple la suite de terme général $u_n = (-1)^n$ est bornée mais n'est pas convergente. D'autre part, avoir une inégalité stricte $|u_n| < M$ n'implique PAS qu'en passant à la limite on ait $|\ell| < M$, mais plutôt $|\ell| \leq M$. Par exemple, cela se voit avec la suite de terme général $u_n = 1 - \frac{1}{n}$ qui tend vers 1 donc $|u_n| < 1$ mais $|\ell| = 1 \leq 1$.

On a donc que l'ensemble des suites convergentes est strictement inclus dans l'ensemble des suites bornées :

$$\{u : \mathbb{N} \to \mathbb{K} \text{ convergente}\} \ \subset\ \{u : \mathbb{N} \to \mathbb{K} \text{ bornée}\}$$

Maintenant on aimerait voir comment avoir l'inclusion inverse, et savoir quelle hypothèse ajouter pour déduire qu'une suite bornée est convergente. Nous avons une telle réciproque sur $\mathbb{K} = \mathbb{R}$, comme il y a une relation d'ordre, si on rajoute une hypothèse de monotonie. Définissons les notions suivantes :

- la suite $(u_n)_n$ est *croissante* si $\forall\, n \in \mathbb{N},\ u_n \leq u_{n+1}$, et strictement croissante si l'inégalité est stricte ;
- la suite $(u_n)_n$ est *décroissante* si pour tout $\forall\, n \in \mathbb{N},\ u_{n+1} \leq u_n$, et strictement décroissante si l'inégalité est stricte ;
- la suite $(u_n)_n$ est *monotone* si elle est croissante ou décroissante ; strictement monotone si elle est strictement croissante ou décroissante ;
- la suite $(u_n)_n$ est *périodique* s'il existe $p \in \mathbb{N}$ (la période) tel que pour tout $n \in \mathbb{N}$, $u_{n+p} = u_n$ (cette définition est aussi valide pour les suites complexes) ;
- la suite $(u_n)_n$ est *stationnaire* si elle est 1-périodique, c'est à dire si elle est constante : $u_{n+1} = u_n$ pour $n \in \mathbb{N}$.

Exemple 1.49. Soit $q \in \mathbb{R}$, on pose $u_n = q^n$, pour tout $n \in \mathbb{N}$; c'est la suite géométrique de paramètre q. Soit $a, b \in \mathbb{R}$ on pose $u_n = a + n \times b$, c'est la suite arithmétique de raison b. On a aussi par exemple $v : \mathbb{N}^* \to \mathbb{R}, n \mapsto \frac{1}{n}$. Les deux premières sont croissantes ou décroissantes, selon les valeurs de q, a, b. La dernière est décroissante.

Une suite qui tend vers $+\infty$ n'est pas forcément croissante comme le montre par exemple la suite suivante :

$$u : \mathbb{N} \longrightarrow \mathbb{R}$$
$$n \longmapsto u_n = \begin{cases} n \text{ si } n \text{ est pair} \\ n - 2n \text{ si } n \text{ est impair} \end{cases}$$

Et d'autre part, une suite croissante ne tend pas forcément vers l'infini, comme le montre la proposition suivante, qui répond à la Proposition 1.47 :

Proposition 1.50. *Toute suite monotone et bornée est convergente. Plus précisément :*

- *Soit $(u_n)_n$ une suite réelle croissante : si elle est majorée, elle converge ; si elle n'est pas majorée, elle diverge vers $+\infty$.*
- *Soit $(v_n)_n$ une suite réelle décroissante : si elle est minorée, elle converge ; si elle n'est pas minorée, elle diverge vers $-\infty$.*

Démonstration. On commence par écrire ce qu'on a si la suite $(u_n)_n$ est majorée. L'ensemble

$$U = \{u_n, n \in \mathbb{N}\}$$

est l'ensemble des valeurs de la suite $(u_n)_n$. Dire que la suite est majorée revient à dire que l'ensemble U est majoré. Par la propriété de la borne supérieure sur $\mathbb{R}$, il existe une borne supérieure à l'ensemble U, qu'on note ℓ.

Montrons que $(u_n)_n$ converge vers ℓ. On a trois faits (croissance de la suite $(u_n)_n$ et ℓ borne supérieure) :

- $\forall n \in \mathbb{N},\ u_n \leq u_{n+1}$
- $\forall n \in \mathbb{N},\ u_n \leq \ell$
- $\forall \epsilon > 0,\ \exists N \in \mathbb{N}$ tel que $\ell - \epsilon < u_N$

On a que $u_N \leq u_n$ pour tout $n \geq N$, et donc on a :

$$\ell - \epsilon < u_N \leq u_n \leq \ell < \ell + \epsilon$$

that is to say :

$$-\epsilon < u_n - \ell < \epsilon$$

donc $\forall n \geq N, |u_n - \ell| < \epsilon$. Cela veut dire que la suite $(u_n)_n$ converge vers ℓ.

Deuxième cas, quand u_n n'est pas majorée, écrivons ce que cela veut dire :

- $\forall n \in \mathbb{N},\ u_n \leq u_{n+1}$
- $\forall M \in \mathbb{R},\ \exists N \in \mathbb{N}$ tel que $M \leq u_N$

Fixons $M \in \mathbb{R}$. On a que $u_N \leq u_n$ pour tout $n \geq N$. Mais alors pour tout $n \geq N$, on a $M \leq u_n$. C'est la définition de tendre vers $+\infty$. Pour les suites décroissantes, les preuves sont similaires. □

Remarque 1.51. On voit que si la suite $(u_n)_n$ est croissante et majorée, alors sa limite est $\sup\{u_n,\ n \in \mathbb{N}\}$. Si la suite est décroissante minorée, alors sa limite est $\inf\{u_n,\ n \in \mathbb{N}\}$.

Pour les suites réelles non-monotones, on peut se débrouiller pour montrer la divergence d'une suite à partir de la connaissance d'une autre suite :

Proposition 1.52. *Soit deux suites réelles $(u_n)_n$ et $(v_n)_n$ telles que $u_n \leq v_n$ pour tout n assez grand. Si $(u_n)_n$ tend vers $+\infty$ alors $(v_n)_n$ tend vers $+\infty$. Si $(v_n)_n$ tend vers $-\infty$ alors $(u_n)_n$ tend vers $-\infty$.*

Démonstration. Il suffit d'appliquer la définition de limite infinie, de la Définition 1.43. □

Pour des limites finies, il faut un encadrement plus fin (à connaitre par coeur) :

Théorème des gendarmes (ou des encadrements). *Soit* $(u_n)_n, (v_n)_n, (w_n)_n$ *trois suites telles que pour tout* n *assez grand on a l'encadrement suivant :*

$$u_n \leq v_n \leq w_n$$

Si $(u_n)_n$ *et* $(w_n)_n$ *convergent toutes deux vers une limite* $\ell \in \mathbb{R}$*, alors* $(v_n)_n$ *converge vers* ℓ*.*

Démonstration. Soit $\epsilon > 0$, comme $(u_n)_n$ (resp. $(w_n)_n$) converge vers $\ell \in \mathbb{R}$, il existe $N \in \mathbb{N}$ (resp. $N' \in \mathbb{N}$) tel que pour tout $n \geq N$ (resp. $n \geq N'$), on a $|u_n - \ell| < \epsilon$ (resp. $|w_n - \ell| < \epsilon$). Posons $\mathcal{N} = \max(N, N')$, alors pour tout $n \geq \mathcal{N}$ on a les deux inégalités à la fois. Donc pour tout $n \geq \mathcal{N}$ on a :

$$\ell - \epsilon < u_n \leq v_n \leq w_n < \ell + \epsilon$$

Cela veut dire que pour tout $n \geq \mathcal{N}$ on a $|v_n - \ell| < \epsilon$. Donc la suite $(v_n)_n$ converge vers ℓ. □

Exemple 1.53. Soit $a > 0$ un nombre réel et $(u_n)_n$ la suite de terme général $u_n = \frac{1}{1+na}$. Alors on a l'encadrement suivant pour tout $n \geq 1$:

$$0 \leq \frac{1}{1+na} \leq \frac{1}{na}$$

Le terme de droite tend vers 0 quand n tend vers l'infini donc la suite $(u_n)_n$ converge vers 0.

Exemple 1.54. On définit la suite $(v_n)_n$ de terme général suivant :

$$v_n = \frac{1}{n} + \frac{1}{n+\sqrt{1}} + \frac{1}{n+\sqrt{2}} + \ldots + \frac{1}{n+\sqrt{n}}$$

Pour tout $0 \leq k \leq n$, on a :

$$\frac{1}{n+\sqrt{n}} \leq \frac{1}{n+\sqrt{k}} \leq \frac{1}{n}$$

Sommons sur tous les k entre 0 et n, et on obtient :

$$\sum_{k=0}^{n} \frac{1}{n+\sqrt{n}} \leq \sum_{k=0}^{n} \frac{1}{n+\sqrt{k}} \leq \sum_{k=0}^{n} \frac{1}{n}$$

En notant que $\sum_{k=0}^{n} \lambda = \lambda \sum_{k=0}^{n} 1 = \lambda(n+1)$, on obtient :

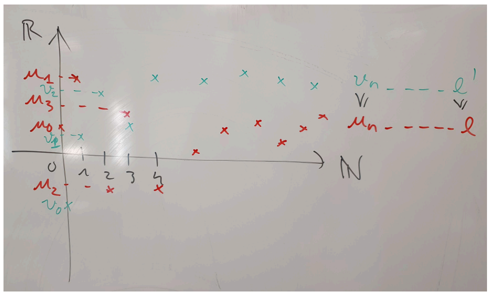
$$\frac{n+1}{n+\sqrt{n}} \leq \sum_{k=0}^{n} \frac{1}{n+\sqrt{k}} \leq \frac{n+1}{n}$$

On note que le terme du milieu est v_n. On pose $u_n = \frac{n+1}{n+\sqrt{n}}$ et $w_n = \frac{n+1}{n}$. Alors on a $u_n \leq v_n \leq w_n$. Comme $\lim\limits_{n \to +\infty} u_n = \lim\limits_{n \to +\infty} w_n = 1$, on en déduit que $\lim\limits_{n \to +\infty} v_n = 1$.

Exemple 1.55. Soit u_n une suite telle que $\left|\frac{u_{n+1}}{u_n}\right| \leq L < 1$ pour tout $n \in \mathbb{N}$, alors $0 \leq |u_n| \leq L^n$ (on peut le montrer par récurrence). Comme $0 < L < 1$, la suite $(L^n)_n$ tend vers 0, donc par encadrement la suite $(u_n)_n$ aussi.

1.4 Suites adjacentes et conséquences

Proposition 1.56. Passage à la limite. *Soit $(a_n)_n$ et $(b_n)_n$ deux suites réelles. On suppose qu'elles sont convergentes vers ℓ et ℓ' respectivement. Si à partir d'un certain rang $N \in \mathbb{N}$, on a $a_n \leq b_n$ pour tout $n \geq N$, alors $\ell \leq \ell'$.*



Démonstration. Par l'absurde. On suppose que $\ell > \ell'$. Prenons $\epsilon = \frac{\ell-\ell'}{2}$. Il existe N tel que $|a_n - \ell| < \epsilon$ et il existe N' tel que $|b_n - \ell'| < \epsilon$. Il existe un entier N'' tel que $\forall\, n \geq N''$, $a_n \leq b_n$. Posons $\mathcal{N} = \max(N, N, N'')$ de façon à ce qu'on a les trois inégalités valides en même temps :

$$\forall\, n \geq \mathcal{N} \qquad |a_n - \ell| < \epsilon, |b_n - \ell'| < \epsilon \text{ et } a_n \leq b_n$$

D'après les deux premières inégalités, nous avons $\ell - \epsilon < a_n$ et $b_n < \ell' + \epsilon$ comme application de la valeur absolue, et donc on a par choix de ϵ :

$$b_n < \ell' + \epsilon = \ell - \epsilon < a_n$$

c'est à dire $b_n < a_n$, pour tout $n \geq \mathcal{N}$, ce qui est en contradiction avec l'inégalité $a_n \leq b_n$. □

Remarque 1.57. Attention, si $a_n < b_n$ (inférieur stricte) pour tout n assez grand, on ne peut PAS déduire que $\ell < \ell'$ et on a même des contre-exemples, comme par exemple si on pose $a_n = -\frac{1}{n}$ et $b_n = \frac{1}{n}$.

Cette proposition permet de déduire une caractéristique de $\mathbb{R}$, la **Propriété des segments emboités**. Soit $(a_n)_n$ une suite croissante et $(b_n)_n$ une suite décroissante, telles que $a_n \leq b_n$ pour tout $n \in \mathbb{N}$. Dans ce cas le segment $I_{n+1} = [a_{n+1}, b_{n+1}]$ est inclus dans $I_n = [a_n, b_n]$, et nous avons donc une suite d'inclusions :

$$\ldots \subset I_{n+1} \subset I_n \subset I_{n-1} \subset \ldots \subset I_2 \subset I_1 \subset I_0,$$

Alors, la propriété des segments emboîtés nous dit qu'il existe au moins un élément $x \in \mathbb{R}$ tel que $a_n \leq x \leq b_n$ pour tout $n \in \mathbb{N}$, c'est à dire que x appartient donc à tous les segments I_n. Nous ne démontrons pas cette propriété des segments emboités. Nous dirons juste que pour la démontrer, nous utilisons le fait qu'une suite croissante majorée (resp. décroissante minorée) est convergente. Or dans la preuve de la proposition 1.50 nous avons utilisée la **Propriété de la borne supérieure**, caractéristique de $\mathbb{R}$. Il est donc équivalent de demander comme axiome de $\mathbb{R}$ l'une ou l'autre des deux propriétés, l'autre s'en déduit.

Théorème 1.58. *Soit $(a_n)_n$ et $(b_n)_n$ deux suites telles que*

— $(a_n)_n$ *une suite croissante,*

— $(b_n)_n$ *une suite décroissante, et*

— $\lim\limits_{n\to+\infty} b_n - a_n = 0.$

Alors les suites $(a_n)_n$ et $(b_n)_n$ sont convergentes et tendent vers la même limite finie ℓ. De plus, pour tout $n \in \mathbb{N}$, $a_n \leq \ell \leq b_n$.

Démonstration. On va montrer que la suite $(a_n)_n$ est majorée, et la suite $(b_n)_n$ est minorée, puis utiliser la propriété 1.50 pour déduire qu'elles convergent. La troisième hypothèse nous dira que leur limite respective coincide. Montrons d'abord que la suite $(b_n - a_n)_n$ est une suite décroissante. Soit $n \in \mathbb{N}$, comme $(b_n)_n$ est décroissante, on a $b_{n+1} \leq b_n$, et comme $(a_n)_n$ est croissante, on a $-a_{n+1} \leq -a_n$. On a donc $b_{n+1} - a_{n+1} \leq b_n - a_n$. La suite $(b_n - a_n)_n$ est donc bien décroissante. Comme elle tend vers 0 (troisième hypothèse), elle est nécessairement positive, et donc on conclut que $a_n \leq b_n$ pour tout $n \in \mathbb{N}$.

Par décroissance de la suite $(b_n)_n$, on a que $a_n \leq b_n \leq b_{n-1} \leq \ldots \leq b_1 \leq b_0$, c'est à dire que $a_n \leq b_0$ pour tout $n \in \mathbb{N}$. De façon identique, par croissance de la suite $(a_n)_n$, on a que $a_0 \leq a_1 \leq \ldots \leq a_{n-1} \leq a_n \leq b_n$, c'est à dire que $a_0 \leq b_n$ pour tout $n \in \mathbb{N}$. Ainsi, la suite décroissante $(b_n)_n$ est minorée et la suite croissante $(a_n)_n$ est majorée. La Proposition 1.50 nous dit que la suite $(a_n)_n$ converge vers une limite ℓ_a et la suite $(b_n)_n$ converge vers une limite ℓ_b. Avec la troisième hypothèse, comme $b_n - a_n \underset{n\to+\infty}{\longrightarrow} 0$, on en déduit en passant à la limite que $\ell_a - \ell_b = 0$, c'est à dire $\ell_a = \ell_b$. Les deux suite convergent vers une limite commune, qu'on note ℓ. Par construction on a que $a_n \leq \ell \leq b_n$ pour tout $n \in \mathbb{N}$. □

Le théorème 1.58 ci dessus est tellement important est utile qu'on donne un nom aux suites ayant cette propriété :

Définition 1.59. *Deux suites $(a_n)_n$ et $(b_n)_n$ satisfaisant les hypothèses du théorème 1.58 sont dites* ***adjacentes****.*

Exemple 1.60. Pour tout $n \in \mathbb{N}^*$, posons :

$$u_n = \sum_{k=0}^{n} \frac{1}{k!} \quad \text{et} \quad v_n = u_n + \frac{1}{n \cdot n!}$$

Alors les suites $(u_n)_n$ et $(v_n)_n$ sont adjacentes et convergent vers le nombre transcendant e. Le nombre d'Euler est irrationnel, car si jamais il était rationnel et s'écrivait $e = \frac{p}{q}$ alors en choisissant $n = q$, on a que $n \cdot n!e$ est un nombre entier, or c'est absurde car on a $n \cdot n!u_n < n \cdot n!e < n \cdot n!u_n + 1$, qui vient de l'encadrement $u_n < e < v_n$.

Exemple 1.61. Approximation décimale d'un réel : soit $x \in \mathbb{R}$ et les deux suites suivantes :

$$a_n = \frac{E(10^n x)}{10^n} \text{ et } b_n = \frac{E(10^n x) + 1}{10^n}$$

Alors on a les inégalités suivantes (en utilisant les propriétés de la partie entière) :

$$0 \leq a_{n+1} - a_n < \frac{1}{10^n}, \qquad 0 \leq b_n - b_{n+1} < \frac{1}{10^n}, \qquad b_n - a_n = \frac{1}{10^n}$$

La suite $(a_n)_n$ est croissante, la suite $(b_n)_n$ est décroissante, et $\lim\limits_{n\to+\infty} b_n - a_n = 0$, ce sont donc des suites adjacentes. La limite des deux suites adjacentes est le nombre réel x. Par exemple, pour $x = \pi = 3,14159\ldots$, on a :

$$a_0 = 3, \; a_1 = 3,1, \; a_2 = 3,14, \; a_3 = 3,141, \; \text{etc. et } b_0 = 4, \; b_1 = 3,2, \; b_2 = 3,15, \; b_3 = 3,142, \; \text{etc.}$$

Exemple 1.62. Soit les deux suites $(a_n)_n$ et $(b_n)_n$ définies par $a_0 = 1$, $b_0 = 2$, et les relations de récurrence suivantes :

$$a_{n+1} = \frac{4}{a_n + b_n} = \frac{2}{b_{n+1}} \qquad \text{et} \qquad b_{n+1} = \frac{a_n + b_n}{2}$$

Montrer qu'elles forment des suites adjacentes, c'est à dire que a_n est croissante et b_n et décroissante, et que $\lim_{n\to+\infty}(b_n - a_n) = 0$. Tout d'abord rang 1 on a $a_1 = 4/3$ et $b_1 = 3/2$ donc on a bien $0 < a_0 < a_1 < b_1 < b_0$. Supposons qu'au rang n on a $0 < a_{n-1} < a_n < b_n < b_{n-1}$. On a d'abord que $a_n + b_n < 2b_n$ donc $b_{n+1} < b_n$, mais donc aussi (comme tout le monde est positif) $\frac{2}{b_n} < \frac{2}{b_{n+1}}$ c'est à dire $a_n < a_{n+1}$. Donc au rang $n+1$ on a bien que $(a_n)_n$ est croissante et $(b_n)_n$ est décroissante. Mais d'autre part, on note que pour tout n, on a $a_nb_n = 2$, donc $b_{n+1} - a_{n+1} = \frac{a_n+b_n}{2} - \frac{4}{a_n+b_n} = \frac{(a_n-b_n)^2}{2(a_n+b_n)} > 0$ donc on a $0 < a_n < a_{n+1} < b_{n+1} < b_n$.

De cette formule, comme $a_0 = 1$ et que tout le monde est plus grand, on a que $a_n + b_n > 2$ donc que

$$b_{n+1}-a_{n+1} = \frac{(a_n - b_n)^2}{2(a_n + b_n)} < \frac{(a_n - b_n)^2}{4} < \frac{(a_{n-1} - b_{n-1})^4}{4 \times 4^2} < \ldots < \frac{(a_0 - b_0)^{2(n+1)}}{4 \times 4^2 \times \ldots \times 4^{2n}} = \frac{1}{4^{\sum_{k=0}^{n} 2^k}}$$

ce qui tend vers zéro donc on a bien la limite $\lim_{n\to+\infty}(b_n - a_n) = 0$. Les suites sont adjacentes. Comme elles tendent vers la même limite ℓ, et que pour tout n on a $a_nb_n = 2$, on en déduit que $\ell^2 = 2$ donc $\ell = \sqrt{2}$.

On va maintenant s'intéresser à une propriété importante des suites réelles (et complexes).

Définition 1.63. *Soit $\varphi : \mathbb{N} \to \mathbb{N}$ une application strictement croissante, c'est à dire que $\varphi(n) < \varphi(n+1)$. On appelle **sous-suite** ou **suite extraite** de la suite réelle (ou complexe) $(u_n)_n$ par rapport à φ, la suite $(v_n)_n$ définie par $v_n = u_{\varphi(n)}$. On note habituellement cette suite $(u_{\varphi(n)})_n$.*

Exemple 1.64. Définissons deux fonctions extractrices strictement croissantes :

$$\begin{aligned} \varphi : \mathbb{N} &\longrightarrow \mathbb{N} \\ n &\longmapsto 2n \end{aligned} \qquad\qquad \begin{aligned} \psi : \mathbb{N} &\longrightarrow \mathbb{N} \\ n &\longmapsto 2n+1 \end{aligned}$$

La suite de terme général $u_n = (-1)^n$ admet deux suites extraites particulières par rapport à φ and ψ. Pour tout $n \in \mathbb{N}$, nous avons $u_{\varphi(n)} = (-1)^{\varphi(n)} = (-1)^{2n} = 1$ et $u_{\psi(n)} = (-1)^{\psi(n)} = (-1)^{2n+1} = -1$. Ces deux suites extraites $(u_{\varphi(n)})_n$ et $(u_{\psi(n)})_n$ sont stationnaires/constantes (et donc convergent) mais la suite originale $(u_n)_n$ ne converge pas.

Exemple 1.65. Prenons la suite périodique de terme général $u_n = \sin\left(\frac{2\pi n}{17}\right)$, de période 17, c'est à dire que $u_{n+17} = u_n$. Posons $\varphi : \mathbb{N} \to \mathbb{N}, n \mapsto 17n$. Alors pour tout $n \in \mathbb{N}$ nous avons $u_{\varphi(n)} = \sin\left(\frac{2\pi 17n}{17}\right) = \sin(2\pi) = 0$. La sous-suite extraite $(u_{\varphi(n)})_n$ est une suite stationnaire (constante) nulle. Par contre, si on prend $\psi : \mathbb{N} \to \mathbb{N}, n \mapsto 17n+1$, on a $u_{\psi(n)} = \sin\left(\frac{2\pi(17n+1)}{17}\right) = \sin\left(\frac{2\pi}{17}\right) \neq 0$. La sous-suite extraite $(u_{\psi(n)})_n$ est aussi une suite stationnaire (constante), mais non nulle.

Proposition 1.66. *Soit $(u_n)_n$ une suite qui tend vers une limite L, soit finie soit infinie. Alors toute suite extraite de $(u_n)_n$ tend vers L.*

Démonstration. Prenons une application strictement croissante $\varphi : \mathbb{N} \to \mathbb{N}$ et on notera $v_n = u_{\varphi(n)}$ ce qui définit une nouvelle notation plus simple pour suite extraite : $(v_n)_n = (u_{\varphi(n)})_n$. Montrons maitenant par récurrence que la stricte croissance de φ implique que $\varphi(n) \geq n$. Pour $n = 0$ on a $\varphi(0) \geq 0$ car $\varphi(0) \in \mathbb{N}$. Ensuite, supposons que $\varphi(n) \geq n$. Comme $\varphi(n+1) > \varphi(n)$, alors $\varphi(n+1) > n$, mais donc $\varphi(n+1) \geq n+1$ car $\varphi(n+1)$ est un entier. Ceci prouve la propriété pour tout $n \in \mathbb{N}$.

On écrit ce que veut dire tendre vers une limite finie ℓ pour la suite $(u_n)_n$:

$$\forall \epsilon > 0, \ \exists N \in \mathbb{N} \text{ tel que } \forall n \geq N, \ |u_n - \ell| < \epsilon$$

Finalement, si pour tout $n \geq N$ on a $|u_n - \ell| < \epsilon$, alors nécessairement on a aussi $|u_{\varphi(n)} - \ell| < \epsilon$ comme on a $\varphi(n) \geq n \geq N$. Ce qui veut dire $|v_n - \ell| < \epsilon$ ce qui est la convergence de la suite $(v_n)_n$ de terme général $v_n = u_{\varphi(n)}$.

Si maintenant, la suite $(u_n)_n$ tend vers $+\infty$, on écrit :

$$\forall M \in \mathbb{R}, \ \exists N \in \mathbb{N} \text{ tel que } \forall n \geq N, \ u_n > M$$

Pour la même raison que précédemment, pour tout $n \geq N$, on a aussi $u_{\varphi(n)} > M$ soit $v_n > M$ donc on a bien que la suite $(v_n)_n$ tend vers $+\infty$. Même argument si $(u_n)_n$ tend vers $-\infty$. $\square$

Remarque 1.67*.* Cette proposition permet de savoir si $(u_n)_n$ n'a pas de limite. Si on peut extraire de $(u_n)_n$ deux sous-suites qui ont deux limites différentes, on sait que la suite $(u_n)_n$ ne tend ni vers une limite finie, ni vers une limite infinie. Par exemple, la suite de terme général $u_n = (-1)^n$ n'admet pas de limites car on peut définir au moins deux sous-suites extraites convergentes vers deux limites différentes. Un autre exemple moins facile est la suite de terme général $u_n = (-1)^n n^2 \sin\left(\frac{1}{n}\right)$.

Une autre propriété des suites extraites est la suivante qui est surprenante mais profonde :

Proposition 1.68. *De toute suite réelle on peut extraire une sous-suite monotone.*

Démonstration. Voir la preuve du Théorème 3.4.7 dans la quatrième édition du livre de Bartle et Sherbert (3.4.6 dans d'autres éditions). $\square$

Théorème de Bolzano-Weierstrass. *De toute suite réelle (ou complexe) bornée, on peut extraire une sous-suite convergente.*

Démonstration. On donne la démonstration dans le cas des suites réelles. Pour le cas des suites complexes, il suffira ensuite d'extraire une première sous-suite telle que les parties réelles convergent, puis de cette première sous-suite, on extrait une deuxième sous-suite telle que les parties imaginaires convergent également.

Preuve comme corollaire de la Proposition 1.68. Toute suite réelle possède une sous-suite extraite monotone. Si la suite d'origine est bornée alors la sous-suite extraite est bornée. Etant monotone et bornée, elle est convergente (c'est la Proposition 1.50).

Preuve avec les segments emboités. Supposons d'abord que la suite $(u_n)_n$ n'est pas stationnaire car dans ce cas il n'y a rien à prouver (la suite est convergente vers la valeur de la suite donc toute suite extraite converge vers cette valeur). Comme la suite $(u_n)_n$ est bornée, l'ensemble

$$U = \{u_n, n \in \mathbb{N}\}$$

est un sous-ensemble borné de $\mathbb{R}$. Comme la suite $(u_n)_n$ prends donc au moins deux valeurs différentes, il existe $a < b \in \mathbb{R}$ tel que $U \subset [a, b]$. On procède par dichotomie, c'est-à-dire que l'on va couper l'intervalle $[a, b]$ en deux en gardant une moitié qui contient une infinité de valeurs de $(u_n)_n$. Puis on va recouper ce nouvel intervalle en deux en gardant une moitié qui contient une infinité de termes et ainsi de suite, pour construire une suite de segments emboités dont la longueur tend vers 0. La propriété des segments emboités de $\mathbb{R}$ permettra de conclure. La preuve se fait par récurrence.

On pose $I_0 = [a, b]$. On divise le segment en deux moitiés égales : $\left[a, \frac{a+b}{2}\right]$ et $\left[\frac{a+b}{2}, b\right]$. La suite $(u_n)_n$ contient une infinité de termes, et comme $U \subset I_0$, nous comprenons que I_0 contient tous les termes de la suite $(u_n)_n$, c'est à dire une infinité. Les deux moitiés gauche et droite ne peuvent *pas* contenir un nombre *fini* de termes de la suite. Au moins une des deux moitiés contient un nombre infini de termes. Nous procédons alors de façon algorithmique :

— si la moitié de gauche possède un nombre infini de termes de la suite $(u_n)_n$, alors on pose $I_1 = \left[a, \frac{a+b}{2}\right]$, $a_1 = a$ et $b_1 = \frac{a+b}{2}$;

— sinon, la moitié de droite contient un nombre infini de termes de la suite $(u_n)_n$, alors on pose $I_1 = \left[\frac{a+b}{2}, b\right]$, $a_1 = \frac{a+b}{2}$ et $b_1 = b$.

Soit $m \in \mathbb{N}$. On suppose tous les intervalles $I_k = [a_k, b_k]$ construits jusqu'au rang $m \geq 1$, avec les propriétés suivantes, pour tout $1 \leq k \leq m$:

— $I_k \subset I_{k-1}$, c'est à dire $a_{k-1} \leq a_k$ et $b_k \leq b_{k-1}$;

— l'intervalle I_k possède une infinité de termes de la suite $(u_n)_n$;

— la longueur de I_k est $b_k - a_k = \frac{b-a}{2^k}$.

Donc en particulier $I_m = [a_m, b_m]$ possède une infinité de termes de la suite $(u_n)_n$. Donc il existe nécessairement une infinité de termes dans au moins l'un des deux intervalles $\left[a_m, \frac{a_m+b_m}{2}\right]$ ou $\left[\frac{a_m+b_m}{2}, b_m\right]$. On procède ainsi :

— si la moitié de gauche possède un nombre infini de termes de la suite $(u_n)_n$, alors on pose $I_{m+1} = \left[a_m, \frac{a_m+b_m}{2}\right]$, $a_{m+1} = a_m$ et $b_{m+1} = \frac{a_m+b_m}{2}$;

— sinon, la moitié de droite contient un nombre infini de termes de la suite $(u_n)_n$, alors on pose $I_{m+1} = \left[\frac{a_m+b_m}{2}, b_m\right]$, $a_{m+1} = \frac{a_m+b_m}{2}$ et $b_{m+1} = b_m$.

Quel que soit le cas, nous avons par construction :

— $I_{m+1} \subset I_m$, c'est à dire $a_m \leq a_{m+1}$ et $b_{m+1} \leq b_m$;

— l'intervalle I_{m+1} possède une infinité de termes de la suite $(u_n)_n$;

— la longueur de I_{m+1} est $b_{m+1} - a_{m+1} = \frac{b-a}{2^{m+1}}$ (car nous avons divisé le segment I_m en deux moitiés égales).

La propriété est donc vraie au rang $m + 1$. Par récurrence, on obtient une suite d'intervalles $(I_m)_m$ qui possèdent toutes les bonnes propriétés énoncées ci dessus. Il résulte de ces propriétés que la suite $(a_m)_m$ est croissante, $(b_m)_m$ est décroissante, et que la suite $(b_m - a_m)_m$ converge vers 0. Ce sont donc deux suites adjacentes, qui convergent nécessairement vers une même limite, notée ℓ, telle que $\ell \in I_m$ pour tout $m \in \mathbb{N}$.

Définissons maintenant la sous suite extraite de $(u_n)_n$ qui va tendre vers ℓ. On choisit $\varphi(0) = 0$, puis chaque $\varphi(n)$ est choisi tel que

1. $\varphi(n + 1) > \varphi(n)$
2. $u_{\varphi(n)} \in I_n$

Notons que le deuxième item est toujours possible car l'intervalle $I_n = [a_n, b_n]$ contient une infinité de termes de la suite. Il existe donc toujours au moins un élément de la suite dont l'indice est plus grand que n. On en conclut que pour tout $n \in \mathbb{N}$, on a $a_n \leq u_{\varphi(n)} \leq b_n$. Par le théorème des gendarmes, on en déduit que $(u_{\varphi(n)})_n$ converge vers ℓ. □

Corollaire 1.69. *Réciproquement, de toute suite non-bornée on peut extraire une sous-suite qui tend vers l'infini.*

Démonstration. D'après la Proposition 1.68, tout suite réelle possède une sous-suite monotone. Si la suite d'origine n'est pas bornée alors la sous-suite monotone n'est pas bornée. Elle tend vers $+\infty$ si elle est croissante, et vers $-\infty$ si elle est décroissante. □

Définition 1.70. *On dit que $\ell \in \mathbb{K}$ est une* valeur d'adhérence *de la suite $(u_n)_n$ s'il existe une sous-suite extraite convergeant vers ℓ.*

Exemple 1.71. On a déjà vu des sous-suite convergentes dans le cas $u_n = (-1)^n$ et $u_n = \sin\left(\frac{2\pi n}{17}\right)$.

Exemple 1.72. Si on prend la suite $(u_n)_n$ de terme général $u_n = sin(n)$, il y a beaucoup de sous-suites convergentes : on peut montrer que pour tout $\ell \in [-1,1]$ il existe une fonctions extractrice $\varphi : \mathbb{N} \to \mathbb{N}$ telle que la sous-suite extraite $(u_{\varphi(n)})_n$ converge vers ℓ. Cela veut dire que l'ensemble des valeurs de la suite – dénoté $U = \{\sin(n) \mid n \in \mathbb{N}\}$ – est dense dans $[-1,1]$.

Cette définition nous permet de reformuler le théorème de Bolzano-Weierstrass de manière plus courte :

Théorème de Bolzano-Weierstrass. *Toute suite réelle (ou complexe) bornée admet une valeur d'adhérence.*

C'est un phénomène courant en mathématiques : on donne un nom aux notions qui sont importantes, à l'aide des définitions. Cela permet ensuite d'écrire les théorèmes et les propositions de manière plus courte et plus précise. Il existe une version encore plus courte du théorème de Bolzano-Weierstrass, c'est sa version topologique :

Théorème de Bolzano-Weierstrass topologique. *Une boule fermée est séquentiellement compacte.*

Dans $\mathbb{R}$ une boule fermée est un segment $[a,b]$ et dans $\mathbb{C}$ une boule fermée est un disque (frontière incluse). La notion de *compacité* est une notion de topologie que nous verrons plus tard, et qui correspond au fait que toute suite d'élément de la boule fermée admette une sous-suite extraite convergente (dans la boule fermée).

Comme vu dans l'exemple 1.72, une suite peut avoir plusieurs valeurs d'adhérence. D'autre part, on déduit de la définition de la valeur d'adhérence que si une suite est convergente, il ne peut y avoir qu'une valeur d'adhérence : sa limite. C'est une reformulation de la Proposition 1.66. Par contre, si une suite $(u_n)_n$ possède une unique valeur d'adhérence, ça n'implique pas que la suite est convergente. En effet, si on prend comme suite celle définie par le terme général suivant :

$$u_n = \begin{cases} 0 \text{ si } n \text{ pair} \\ n \text{ si } n \text{ impair} \end{cases}$$

On a que la suite est non bornée donc diverge (ne converge pas), et que d'autre part si on définit l'application extractrice $\varphi : \mathbb{N} \to \mathbb{N}, n \mapsto 2n$, on a que $u_{\varphi(n)} = u_{2n} = 0$ pour tout $n \in \mathbb{N}$. Et donc 0 est une valeur d'adhérence de la suite, et il n'en existe pas d'autre (cela se voit). Par contre la réciproque correcte est la suivante :

Corollaire 1.73. *Toute suite bornée admettant une unique valeur d'adhérence est convergente.*

Démonstration. Soit ℓ la valeur d'adhérence de $u : \mathbb{N} \to \mathbb{K}$. Montrons que u est convergente. On raisonne par l'absurde et on suppose que u n'est pas convergente. On peut alors fixer $\epsilon > 0$ tel que pour tout $N \in \mathbb{N}$, il existe $n \geq N$ tel que $|u_n - \ell| > \epsilon$. On construit alors, par récurrence une extractrice $\varphi : \mathbb{N} \to \mathbb{N}$ telle que $|u_{\varphi(n)} - \ell| > \epsilon$, et on pose $v_n = u_{\varphi(n)}$ pour tout $n \in \mathbb{N}$.

D'après Bolzano-Weierstrass, il existe $\ell' \in \mathbb{K}$ et $\psi : \mathbb{N} \to \mathbb{N}$ extractrice telle que la suite $(v_{\psi(n)})_n$ converge vers ℓ'. Pour tout $n \in \mathbb{N}$, on a

$$\left|v_{\psi(n)} - \ell\right| = \left|u_{\varphi(\psi(n))} - \ell\right| > \epsilon$$

En faisant tendre dans le membre de gauche n vers l'infini, on obtient $|\ell' - \ell| \geq \epsilon$, donc $\ell \neq \ell'$, ce qui est absurde. $\square$

Remarque 1.74. Ce résultat est à mettre en parallèle avec la Proposition 1.50, réciproque de la Proposition 1.47. Nous voyons que la Proposition 1.73 est une autre réciproque de la Proposition 1.47, mais qui cette fois s'applique aussi aux suites complexes !

Une caractérisation de la notion de valeur d'adhérence à l'aide des quantificateurs universels et existentiels se fait ainsi :

Proposition 1.75. *Soit $(u_n)_n$ une suite de $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) et soit $\ell \in \mathbb{K}$. ℓ est une valeur d'adhérence de u si et seulement si :*

$$\forall \epsilon > 0,\ \forall N \in \mathbb{N},\ \exists n \geq N \text{ tel que } |u_n - \ell| < \epsilon$$

Remarque 1.76. Cela s'interprète topologiquement par le slogan suivant : "Pour tout voisinage V de ℓ dans $\mathbb{K}$, il existe une infinité d'indices n tels que $u_n \in V$". Comparer avec la phrase définissant la convergence d'une suite.

Définition 1.77. *Soit A une partie de $\mathbb{K}$. On appelle* adhérence *de A et on note $\overline{A}$ toutes les valeurs d'adhérence des suites d'éléments de A :*

$$\overline{A} = \{\ell \in \mathbb{R} \text{ tel que il existe une suite } (u_n)_n \text{ d'éléments de } A \text{ convergeant vers } \ell\}$$

Remarque 1.78. Le slogan de cette définition c'est que l'adhérence d'une partie A de $\mathbb{K}$ est l'ensemble de tous les points de $\mathbb{K}$ qu'on peut atteindre comme limite d'une suite de points de A.

Dans $\mathbb{R}$, pour tout intervalle I (ouvert, semi-ouvert ou fermé), l'adhérence de I est l'intervalle fermé formé de l'union de I avec ses points frontières. Par exemple, si $a < b \in \mathbb{R}$ alors voici une liste des adhérences de différents intervalles :

$$\begin{aligned} \overline{]a,b[} = \overline{[a,b[} = \overline{]a,b]} = \overline{[a,b]} &= [a,b], \\ \overline{]a,+\infty[} = \overline{[a,+\infty[} &= [a,+\infty[, \\ \overline{]-\infty,b[} = \overline{]-\infty,b]} &=]-\infty,b] \end{aligned}$$

Exemple 1.79. L'adhérence de l'ensemble $A = \left\{\frac{1}{n} \mid n \in \mathbb{N}^*\right\}$ est $\overline{A} = A \cup \{0\}$. On a ajouté à A l'origine, car c'est la limite (donc une valeur d'adhérence) de la suite de terme général $\frac{1}{n}$.

Exemple 1.80. Pour un domaine $D \subset \mathbb{R}$ formé d'une union d'intervalle $D = \bigcup_{i\in\mathbb{N}} I_i$, l'adhérence de D n'est *pas* l'union de toutes les adhérences de ses parties : $\bigcup_{i\in\mathbb{N}} \overline{I_i}$. En effet prenons par exemple $I_i = \left]\frac{1}{i+1}, \frac{1}{i}\right[$ pour tout $i \geq 1$. Alors $\overline{I_i} = \left[\frac{1}{i+1}, \frac{1}{i}\right]$ et l'union de tous ces intervalles est $\bigcup_{i\in\mathbb{N}} \overline{I_i} =]0,1]$. Ceci n'est pas l'adhérence de $D = \bigcup_{i\in\mathbb{N}} I_i$, qui vaut $\overline{D} = [0,1]$, car on peut trouver une suite d'éléments de D qui converge vers 0.

L'ensemble des termes d'une suite définit un sous-sensemble de $\mathbb{R}$ ou $\mathbb{C}$:

$$U = \{u_n \mid n \in \mathbb{N}\}$$

C'est un ensemble de points, de cardinal (nombre d'éléments) fini ou dénombrable (autant que dans $\mathbb{N}$). Par exemple : $u_n = (-1)^n$ donne $U = \{-1, 1\}$ (cardinal fini) et $u_n = \frac{1}{n}$ donne $U = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right\}$ (cardinal dénombrable). Etudier les suites c'est étudier les ensembles fini ou dénombrables de $\mathbb{R}$ (ou $\mathbb{C}$). Etudier les sous-suites extraites convergentes, c'est étudier les valeurs d'adhérence de ces ensembles, c'est à dire leur adhérence dans le sens donné ci-dessus. Les valeurs d'adhérences nous donnent donc une caractérisation séquentielle de la topologie de $\mathbb{K}$.

Proposition 1.81. Caractérisation séquentielle des fermés. *Soit A un sous-ensemble de $\mathbb{K}$, alors les conditions suivantes sont équivalentes :*

— *A est fermé ;*
— *$A = \overline{A}$;*
— *si une suite d'éléments de A converge, alors sa limite est dans A.*

Définition 1.82. *Soit A et B deux parties de $\mathbb{K}$. On dit que A est* dense *dans B si $\overline{A} = B$.*

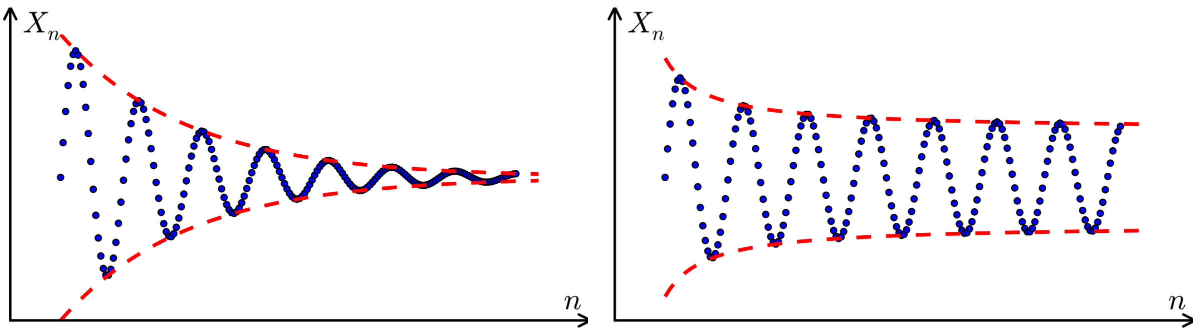
Exemple 1.83. Dans la Proposition 1.31, nous avons vu que les nombres rationnels $\mathbb{Q}$ sont denses dans $\mathbb{R}$. Cela correspond exactement à la nouvelle Définition 1.82.

Exemple 1.84. L'ensemble $U = \{\sin(n) \mid n \in \mathbb{N}\}$ est dense dans $[-1, 1]$: tout point de $[-1, 1]$ peut être atteint comme la limite d'une suite d'éléments de U, c'est à dire comme la limite d'une sous-suite extraite de la suite $(u_n)_n$ de terme général $u_n = \sin(n)$.

Nous allons désormais nous tourner vers une notion importante cette année comme l'année prochaine, qui s'appuie sur la notion de suite extraite :

Définition 1.85. *On dit qu'une suite réelle ou complexe $(u_n)_n$ est une* ***suite de Cauchy*** *si elle possède la propriété suivante :*

$$\forall\, \epsilon > 0,\ \exists\, N \in \mathbb{N} \text{ tel que } \forall\, p, q \geq N,\ |u_p - u_q| < \epsilon$$



Exemple 1.86. Tout simplement la suite de terme général $u_n = \frac{1}{n}$.

Remarque 1.87. Attention, avoir la propriété :

$$\forall\, \epsilon > 0,\ \exists\, N \in \mathbb{N} \text{ tel que } \forall\, n \geq N,\ |u_{n+1} - u_n| < \epsilon$$

n'est pas équivalente à la propriété satisfaite par les suites de Cauchy. En effet la propriété ci-dessus est beaucoup plus faible. Par exemple prenons la suite de terme général $u_n = \ln(n)$, alors

$$u_{n+1} - u_n = \ln(n+1) - \ln(n) = \ln\left(\frac{n+1}{n}\right) = \ln\left(1 + \frac{1}{n}\right)$$

qui tend vers $\ln(1) = 0$ lorsque n vers l'infini. Donc elle satisfait la condition 1.87, mais elle n'est pas une suite de Cauchy car $(u_n)_n$ tend vers $+\infty$, donc pour tout $N \in \mathbb{N}$ on peut toujours trouver $p, q \geq N$ tels que $|u_p - u_q| > 1$.

Lemme 1.88. *Toute suite de Cauchy est bornée.*

Démonstration. Etudions le cas des suites réelles, car pour les suites complexes, il suffit d'appliquer la preuve aux suites des parties réelles et imaginaires. Soit $(u_n)_n$ une suite de Cauchy réelle. Montrons que la suite est bornée. Soit $\epsilon = 1$, alors le critère de Cauchy nous dit qu'il existe $N \in \mathbb{N}$ tel que $\forall\, p, q \geq N$ tels que $|u_p - u_q| < 1$. En particulier, pour $q = N$ on a $|u_p - u_N| < 1$

pour tout $p \geq N$. On a donc, pour tout $p \geq N$, $-1 < u_p - u_N < 1$, donc $-1+u_N < u_p < 1+u_N$. Et comme $-|u_N| \leq u_N \leq |u_N|$, on a par la suite :

$$-1-|u_N| \leq -1+u_N < u_p < 1+u_N \leq 1+|u_N|$$

et donc pour tout $p \geq N$, on a $|u_p| < 1+|u_N|$. On note $M = \max(|u_0|, |u_1|, \ldots, |u_{N-1}|, 1+|u_N|)$, et donc on obtient :

$$\begin{cases} |u_p| \leq M & \text{si } 0 \leq p \leq N-1 \\ |u_p| \leq 1+|u_N| \leq M & \text{si } p \geq N \end{cases}$$

Ce qui montre que $\forall\, p \in \mathbb{N}, |u_p| \leq M$, c'est à dire que la suite est bornée. □

Théorème 1.89. *Dans $\mathbb{R}$ et $\mathbb{C}$, une suite est convergente si et seulement si elle est de Cauchy.*

Démonstration. La preuve que toute suite convergente est une suite de Cauchy sera faite en exercice. Inversement, montrons que toute suite réelle de Cauchy est convergente, car pour les suites complexes, il suffit d'appliquer la preuve aux suites des parties réelles et imaginaires. Comme la suite $(u_n)_n$ est de Cauchy, si on fixe $\epsilon > 0$, alors il existe $N \in \mathbb{N}$ tel que pour tous $p, q \geq N$, on a $|u_p - u_q| < \frac{\epsilon}{2}$. Nous savons que toute suite de Cauchy est bornée. Par le théorème de Bolzano-Weierstrass, la suite $(u_n)_n$ admet donc une valeur d'adhérence, qu'on note ℓ. En appliquant la Proposition 1.75, on voit qu'il existe $n_0 \geq N$ tel que $|u_{n_0} - \ell| < \frac{\epsilon}{2}$. Soit $n \geq N$, alors on a que $|u_n - u_{n_0}| < \frac{\epsilon}{2}$ car $n, n_0 \geq N$. Alors on a, par l'inégalité triangulaire :

$$|u_n - \ell| = |u_n - u_{n_0} + u_{n_0} - \ell| \leq \underbrace{|u_n - u_{n_0}|}_{<\frac{\epsilon}{2} \text{ par suite de Cauchy}} + \underbrace{|u_{n_0} - \ell|}_{<\frac{\epsilon}{2} \text{ par valeur d'adhérence}} < \epsilon$$

C'est la définition que la suite $(u_n)_n$ converge vers ℓ. □

Exemple 1.90. Comme application très importante de ce résultat, nous avons la notion de convergence absolue des séries numériques. Soit $u : \mathbb{N} \to \mathbb{K}$ une série complexe ou réelle, et posons, pour tout $n \in \mathbb{N}$:

$$S_n = \sum_{k=0}^{n} u_k \quad \text{et} \quad T_n = \sum_{k=0}^{n} |u_k|$$

Montrons que si la suite réelle positive $(T_n)_n$ converge, alors la suite $(S_n)_n$ converge. Soit donc $0 \leq q < p \in \mathbb{N}$, nous avons que :

$$|S_p - S_q| = \left|\sum_{k=0}^{p} u_k - \sum_{k=0}^{q} u_k\right| = \left|\sum_{k=q+1}^{p} u_k\right| \leq \sum_{k=q+1}^{p} |u_k| = \sum_{k=0}^{p} |u_k| - \sum_{k=0}^{q} |u_k| = T_p - T_q$$

Autrement dit, pour tout $p \neq q \in \mathbb{N}$, on a que :

$$|S_p - S_q| \leq |T_p - T_q| \tag{1.1}$$

Si la suite $(T_n)_n$ converge, alors elle est de Cauchy donc on a la propriété de Cauchy suivante :

$$\forall\, \epsilon > 0,\ \exists\, N \in \mathbb{N} \text{ tel que } \forall\, p, q \geq N,\ |T_p - T_q| < \epsilon$$

Mais dans ce cas, on a que la suite $(S_n)_n$ est aussi une suite de Cauchy du fait de la majoration (1.1). Donc la suite $(S_n)_n$ converge.

Définition 1.91. *Une partie A de $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) est dite complète si, pour toute suite de Cauchy d'éléments de A (qui converge nécessairement donc), la limite est dans A.*

Exemple 1.92. L'ensemble $\mathbb{Q}$ des nombres rationnels n'est pas complet, car une suite de nombres rationnels qui tend vers $\sqrt{2}$ est de Cauchy, mais la limite est un irrationnel (donc pas dans $\mathbb{Q}$). Le corps des réels $\mathbb{R}$ est complet par le Théorème 1.89. C'est d'ailleurs le complété de $\mathbb{Q}$: c'est le plus petit ensemble complet contenant $\mathbb{Q}$. Autrement dit pour obtenir $\mathbb{R}$, on a ajouté à $\mathbb{Q}$ toutes les limites des suites de Cauchy rationnelles.

Exemple 1.93. Un exemple de suite de Cauchy rationnelle qui converge en dehors de $\mathbb{Q}$ est la suite de terme général suivant :

$$u_n = \sum_{k=0}^{n} \frac{1}{k!}$$

Alors on peut montrer que pour tout $p > q \geq 2$, on a :

$$0 \leq u_p - u_q \leq \frac{1}{(q+1)!}\left(1 + \frac{1}{q+2} + \ldots + \frac{1}{(q+2)^{p-q-1}}\right)$$

De cela, en notant que la parenthèse de droite est la somme partielle de rang $p-q-1$ de la suite géométrique de raison $\frac{1}{q+2}$, et que $\frac{q+2}{(q+1)^2} \leq \frac{1}{2}$ (car $q \geq 2$), nous avons que :

$$0 \leq u_p - u_q \leq \frac{1}{(q+1)!}\left(\frac{1-\left(\frac{1}{q+2}\right)^{p-q}}{q+1}\right)(q+2) \leq \frac{1}{(q+1)!} \times \frac{q+2}{q+1} \leq \frac{1}{2 \times (q!)} \leq \frac{1}{q}$$

Maintenant, pour tout $1 \leq \epsilon > 0$, si on prend $N = E\left(\frac{1}{\epsilon}\right) + 1 \geq 2$ alors $\frac{1}{\epsilon} < N$, et pour tout $p, q \geq N$, on a $|u_p - u_q| \leq \frac{1}{N} < \epsilon$. Ainsi, la suite $(u_n)_n$ est une suite de Cauchy (rationnelle). Elle converge dans $\mathbb{R}$, vers le nombre irrationnel transcendant e. La suite ne converge donc pas dans $\mathbb{Q}$.

Exemple 1.94. Dans $\mathbb{R}$, l'intervalle semi-ouvert $]0,1]$ n'est pas complet car la suite de terme général $u_n = 1/n$ est de Cauchy, mais elle converge vers 0 qui n'est pas dans l'intervalle. Par contre le segment $[0,1]$ est complet. On voit que le fait d'être fermé et le fait d'être complet est relié. Nous donnons la caractérisation suivante que nous ne démontrerons pas :

Proposition 1.95. *Dans $\mathbb{R}$ ou $\mathbb{C}$, un sous-ensemble est complet si et seulement si il est fermé.*

Intuitivement, un espace est complet s'il n'a pas de trou ou s'il n'a pas de point manquant. On peut "compléter" un espace métrique en remplissant les trous, en rajoutant toutes les limites des suites de Cauchy. Ainsi on obtient $\mathbb{R}$ en "complétant" $\mathbb{Q}$ en rajoutant toutes ces limites : c'est le plus petit espace métrique complet dans lequel $\mathbb{Q}$ est dense. C'est une construction alternative des réels. Pour passer de $\mathbb{Q}$ à $\mathbb{R}$ on requiert donc des propriétés ensemblistes (coupures de Dedekind) ou topologiques (convergence de Cauchy), mais pour passer de $\mathbb{R}$ à $\mathbb{C}$ on requiert une propriété algébrique : $\mathbb{R}$ est le complété topologique de $\mathbb{Q}$, et $\mathbb{C}$ est la clôture algébrique de $\mathbb{R}$.

2 Fonctions réelles

Rappel : soit E et F deux ensembles. On appelle "application de E dans F" et on note $f : E \to F$ lorsqu'on assigne à certains points de E un point de F. Lorsque $E = \mathbb{R}$ et $F = \mathbb{R}$, ou tout sous-ensemble de $\mathbb{R}$, on parle de *fonction*. Ce chapitre va être dédié à étudier la notion de continuitié, de dérivabilité, de convexité et les propriétés des fonctions sur $\mathbb{R}$.

2.1 Limites et continuité

Dans ce chapitre, notons D une partie de $\mathbb{R}$, et soit $f : D \to \mathbb{R}$ une fonction. La notation de la définition d'une fonction se fait comme suit :

$$\begin{aligned} f :\ D &\longrightarrow \mathbb{R} \\ x &\longmapsto \text{expression de } f(x) \end{aligned}$$

Bien noter la différence des flèches (avec et sans pied). On appelle *domaine de définition* de la fonction f le sous-ensemble D de $\mathbb{R}$ sur lequel f est définie. Par exemple, si $D = \mathbb{N}$ alors une fonction $f : \mathbb{N} \to \mathbb{R}$ est une suite. Mais en général D est un intervalle ou une union disjointe d'intervalles. On dit que f *est définie en* $x \in \mathbb{R}$ si $x \in D$, c'est à dire si $f(x)$ existe. La notion de limite s'applique à tout point de l'adhérence de D – notée $\overline{D}$ – dont on rappelle la Définition 1.77.

Définition 2.1. *Soit $f : D \to \mathbb{R}$ une fonction. Soit $a \in \overline{D}$ et soit $\ell \in \mathbb{R}$. On dit que f* a pour limite ℓ en a *– ou que f* tend vers ℓ quand x tend vers a *– et on note $\lim\limits_{\substack{x \to a \\ x \neq a}} f(x) = \ell$ si :*

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 \textit{ tel que } \forall\, x \in D,\ \textit{si } \begin{cases} |x-a| < \delta \\ x \neq a \end{cases} \quad \textit{alors } |f(x) - \ell| < \epsilon$$

La définition signifie que "$f(x)$ est aussi près que l'on veut de ℓ à condition de choisir x suffisamment près de a, mais différent de a". Cette notion, malgré l'intuition qu'on peut en avoir, a été extrêmement longue à être clarifiée par les mathématiciens et n'a reçu ses fondements définitifs que vers 1850. On peut réécrire la condition comme suit :

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 \ \text{ tel que } \ \forall\, x \in \left(D \backslash \{a\} \cap]a-\delta, a+\delta[\right),\ \text{on a } f(x) \in]\ell - \epsilon, \ell + \epsilon[$$

On peut récrire la définition de limite de f au point a à l'aide des voisinages :

$$\forall\, V_\ell \text{ voisinage de } \ell,\ \exists\, V_a \text{ voisinage de } a \text{ tel que, si } x \in D\backslash\{a\} \cap V_a \text{ alors } f(x) \in V_\ell$$

Exemple 2.2. Soit $I = [0, +\infty[$, et soit $a = 0$, $f(x) = \sqrt{(x)}$. On sait que la fonction est continue (la définition rigoureuse viendra plus tard). Mais montrons que la fonction tend vers 0 en 0 avec la définition ci dessus. Soit $\epsilon > 0$. On pose $\delta = \epsilon^2 > 0$. Soit $0 < x < \delta$ alors comme la fonction $x \mapsto \sqrt{x}$ est croissante, on a $0 \leq \sqrt{x} < \sqrt{\delta} = \epsilon$. On a donc bien que $|f(x) - 0| < \epsilon$ pour tout x tel que $|x - 0| < \delta$.

Montrons maintenant que pour tout $a > 0$, la fonction f tend vers $\sqrt{a}$ quand x tend vers a (c'est la continuité de la racine carrée). Soit $a > 0$. Comme $u^2 - v^2 = (u-v)(u+v)$ on l'applique à $u = \sqrt{x}$ et $v = \sqrt{y}$, de sorte qu'on a $\sqrt{x} - \sqrt{y} = \frac{x-y}{\sqrt{x}+\sqrt{y}}$ pour tout $x, y > 0$. On a donc, en appliquant la valeur absolue :

$$|\sqrt{x} - \sqrt{y}| = \frac{|x-y|}{\sqrt{x}+\sqrt{y}} \leq \frac{|x-y|}{\sqrt{y}}$$

Soit donc maintenant $\epsilon > 0$. On pose $\delta = \epsilon\sqrt{a}$. Prenons $x \in I$ tel que $x \neq a$ et $|x-a| \leq \delta = \epsilon\sqrt{a}$ cela veut dire que $\frac{|x-a|}{\sqrt{a}} \leq \epsilon$. D'après l'équation du dessus, on en déduit que $|\sqrt{x} - \sqrt{a}| \leq \epsilon$, ce qui correspond bien à la définition.

Proposition 2.3. *La limite d'une fonction en un point donné est unique.*

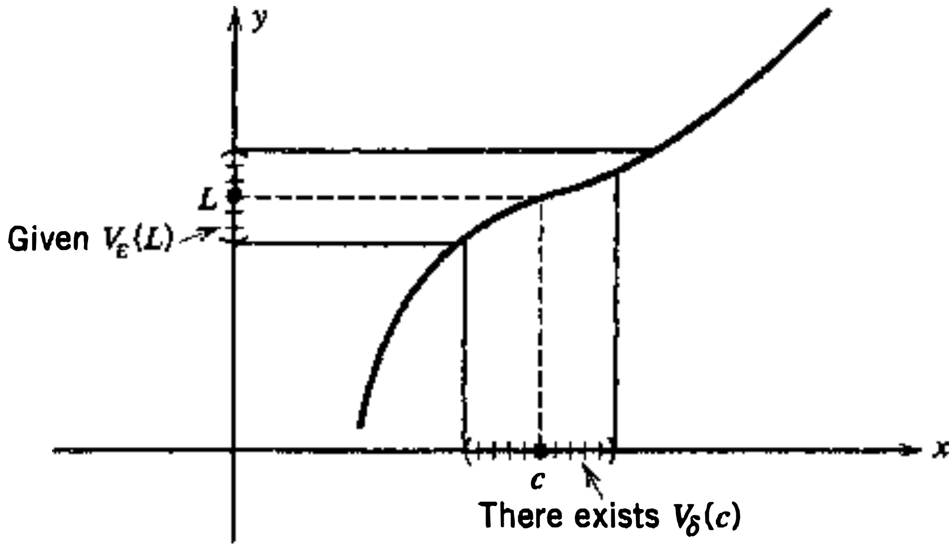


Figure 4.1.1 The limit of f at c is L

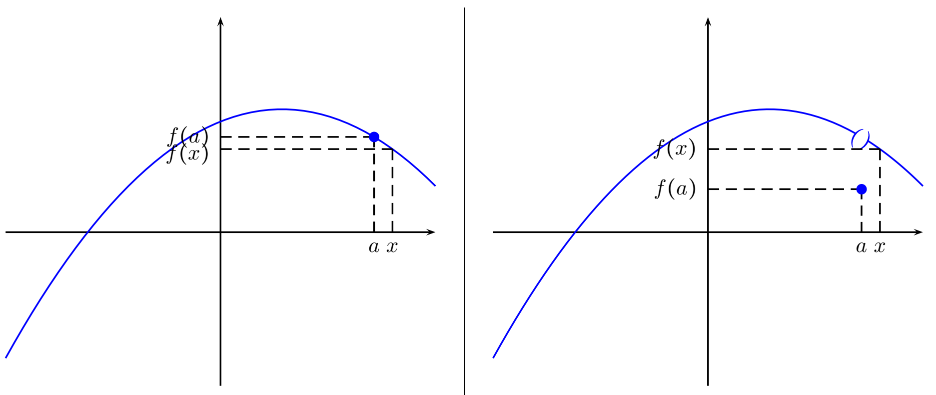
FIGURE 2 – Dans cette image tiré du livre de Bartle et Sherbert, le voisinage V_ϵ est le segment ouvert $]L-\epsilon, L+\epsilon[$ et $V_\delta =]c-\delta, c+\delta[$. C'est une interprétation géométrique de la définition d'une limite (où f est définie en c).

Démonstration. C'est le même type de preuve que pour la Proposition 1.40, où on remplace $\exists\, N \in \mathbb{N}$ tq $\forall n \geq N$ par $\exists\, \delta > 0$ tel que $\forall\, x \in (D\backslash\{a\} \cap]a-\delta, a+\delta[)$. □

Définition 2.4. *Soit $f : D \to \mathbb{R}$ une fonction et $a \in D$ (c'est à dire que f est définie en a). Supposons que f admette une limite ℓ en a. Si la limite ℓ de f en a coincide avec la valeur de f en a, c'est à dire si $\ell = f(a)$, ce qui s'écrit mathématiquement :*

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 \text{ tel que } \forall\, x \in D \cap]a-\delta, a+\delta[\,,\ \text{ on a } |f(x) - f(a)| < \epsilon$$

alors on dit que f est continue au point a*, et on peut écrire $\lim\limits_{x \to a} f(x) = f(a)$. On dit que f est continue sur D si f est continue en tout point de D. Par contre, si $\ell \neq f(a)$ on dit que f est discontinue en a.*



Quand x tend vers a, $f(x)$ tend vers $f(a)$. f est continue en a.

Quand x tend vers a, $f(x)$ ne tend pas vers $f(a)$. f n'est pas continue en a.

Exemple 2.5. Toutes les fonctions dont on peut dessiner le graphe au tableau sans lever la main sont continues. Attention il existe plein de fonctions continues qu'on ne peut pas dessiner au tableau. Il en existe d'ailleurs infiniment plus que celles qu'on peut dessiner.

Remarque 2.6. On peut récrire la définition de continuité de f au point a à l'aide des voisinages :

$$\forall\, V_{f(a)} \text{ voisinage de } f(a),\ \exists\, V_a \text{ voisinage de } a \text{ tel que, si } x \in D \cap V_a \text{ alors } f(x) \in V_{f(a)}$$

Il existe même une formulation encore plus abstraite (topologique) que nous donnons sans prouver : $f : D \to \mathbb{R}$ est continue au point $a \in D$ si et seulement si, pour tout voisinage $V_{f(a)}$ de $f(a)$, l'ensemble $f^{-1}(V_{f(a)})$ est un voisinage de a.

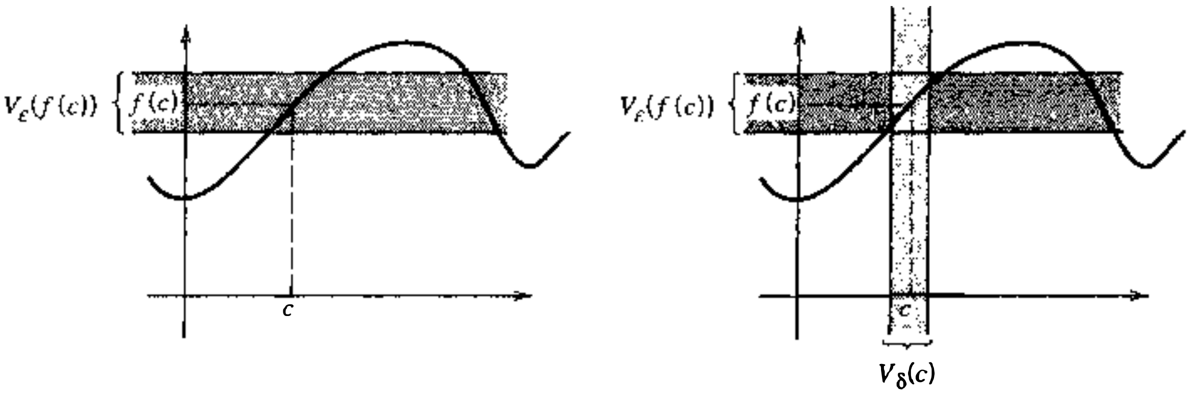


Figure 5.1.1 Given $V_\varepsilon(f(c))$, a neighborhood $V_\delta(c)$ is to be determined

Proposition 2.7. Caractérisation séquentielle de la limite. *Soit $f : D \to \mathbb{R}$ une fonction, $\ell \in \mathbb{R}$ et $a \in \overline{D}$. Les propriétés suivantes sont équivalentes :*

1. *f tend vers ℓ quand x tend vers a ;*
2. *pour toute suite $(x_n)_n$ à valeurs dans $D\backslash\{a\}$ convergeant vers a, la suite $(f(x_n))_n$ converge vers ℓ.*

Démonstration. Montrons 1. $\Longrightarrow$ 2. et supposons f admet pour limite ℓ en a. Cela veut dire que la condition de la Définition 2.1 est satisfaite. Fixons $\epsilon > 0$, alors il existe $\delta > 0$ satisfaisant les conditions de la définition. Soit $(x_n)_n$ une suite d'éléments de D convergeant vers a tout en étant différents de a. Cela veut dire qu'à partir d'un certain N assez grand, on a $|x_n - a| < \delta$ pour tout $n \geq N$. Dans ce cas, comme la condition de la Définition 2.1 est satisfaite pour tous ces x_n, on a $|f(x_n) - \ell| < \epsilon$ pour tout $n \geq N$. Cela veut dire que la suite $(f(x_n))_n$ converge donc vers ℓ.

Montrons 2. $\Longrightarrow$ 1. en montrant la contraposée, c'est à dire que la négation de 1. implique la négation de 2. Supposons que f n'admet pas de limite en a. Cela veut dire qu'il existe $\epsilon > 0$ tel que pour tout $\delta > 0$, il existe $x \in (D\backslash\{a\} \cap]a-\delta, a+\delta[)$ tel que $|f(x) - \ell| \geq \epsilon$. Construisons une suite $(x_n)_n$ convergeant vers a mais telle que la suite $(f(x_n))_n$ ne converge pas vers ℓ (c'est la négation de 2.). Soit un ϵ satisfaisant la condition écrite ci dessus. Alors, pour tout $n \in \mathbb{N}$, il existe $x_n \in (D\backslash\{a\} \cap]a - \frac{1}{n}, a + \frac{1}{n}[)$ tel que $|f(x_n) - \ell| \geq \epsilon$. La suite $(x_n)_n$ ainsi construite converge vers a mais la suite $(f(x_n))_n$ reste à une distance au moins ϵ de ℓ, elle ne converge donc pas vers ℓ. On a montré la négation de 2. $\square$

Remarque 2.8. Le mot important de la Proposition 2.7 est *pour toute suite*. Car en particulier, si on trouve deux suites $(x_n)_n$ et $(y_n)_n$ d'éléments de $D\backslash\{a\}$ qui tendent vers a et telles que $\lim\limits_{n\to+\infty} f(x_n) \neq \lim\limits_{n\to+\infty} f(y_n)$ alors la fonction f n'admet pas de limite en a. En particulier, f ne peut pas être continue en a.

Proposition 2.9. Critère de discontinuité. *Soit $f : D \to \mathbb{R}$ une fonction et $a \in D$. f est discontinue au point a si et seulement si il existe une suite $(x_n)_n$ de point de D qui converge vers a, mais telle que la suite $(f(x)_n)_n$ ne converge pas vers $f(a)$.*

Exemple 2.10. La fonction de Dirichlet (1829) est une fonction $f : \mathbb{R} \to \mathbb{R}$ définie comme suit :

$$f(x) = \begin{cases} 0 \text{ si } x \in \mathbb{Q} \\ 1 \text{ si } \ \in \mathbb{R} \backslash \mathbb{Q} \end{cases}$$

La fonction n'est pas dessinable et n'admet de limites en aucun point. En effet, $\mathbb{Q}$ et $\mathbb{R} \backslash \mathbb{Q}$ sont denses dans $\mathbb{R}$. Donc, pour tout $x \in \mathbb{R}$ il existe une suite de rationnels $(a_n)_n$ et une suite d'irrationnels $(b_n)_n$ qui convergent vers x, mais pour tout n $f(a_n) = 0$ et $f(b_n) = 1$. Donc f ne peut admettre de limite en x. Par exemple pour $x = 0$ on peut prendre $a_n = \frac{1}{n}$ et $b_n = \frac{\pi}{n}$.

Remarque 2.11. On peut faire la différence entre deux types de discontinuité : si la fonction f tend vers une limite finie à gauche et à droite du point a, mais que ces limites sont différentes, ou bien que ces limites sont égales mais la valeur de $f(a)$ est différentes de cette valeur commune, alors on parle de *discontinuité de première espèce*. Si la fonction f n'admet pas de limite finie soit à gauche soit à droite alors on parle de *discontinuité de deuxième espèce* ou *discontinuité essentielle*. Le théorème de Froda, redécouvert en 1929 par le mathématicien roumain Alexandru Froda mais dont des versions plus générales avaient été trouvées de 1907 à 1910 par Grace Chisholm Young et William Henry Young, assure que l'ensemble des points de discontinuité de première espèce d'une fonction réelle d'une variable réelle (définie sur un intervalle) est au plus dénombrable.

D'autre part, on peut préciser si on tend vers ℓ par la gauche ou par la droite en écrivant $x < a$ ou $x > a$ plutôt que $x \neq a$ sous le symbole de la limite, et on parle alors de *limite à gauche* ou *limite à droite*, et on note :

$$\lim_{\substack{x \to a \\ x < a}} f(x) = \ell \qquad \text{et} \qquad \lim_{\substack{x \to a \\ x > a}} f(x) = \ell$$

Nous verrons plus loin qu'une fonction est continue en a si les limites à gauche, à droite, et la valeur de f en a coincident toutes. Dans la suite si on parle de limite de f en a sans préciser laquelle, on suppose implicitement que la limite à gauche de f en a et celle à droite coincident. Tous les résultats sur les limites se généralisent aux limites à gauche et aux limites à droite.

Exemple 2.12. Soit E la fonction partie entière. Soit $a = 1$. Calculons la limite à gauche de E en 1 et sa limite à droite. Soit $\epsilon > 0$. Prenons $\delta = 1/2$. Alors pour tout $1/2 < x < 1$, $E(x) = 0$ donc $|E(x) - 0| = 0 < \epsilon$. Cela étant vrai pour tout $\epsilon > 0$, la fonction partie entière tend vers 0 quand x tend vers 1 par la gauche. Maintenant prenons l'autre côté : pour tout $1 < x < 3/2$ on a $E(x) = 1$ et $|E(x) - 1| = 0 < \epsilon$. Cela étant vrai pour tout $\epsilon > 0$, la fonction partie entière tend vers 1 quand x tend vers 1 par la droite. La fonction partie entière a donc deux limites différentes en 1 ! ! En particulier elle n'est pas continue en 1.

Définition 2.13. *Soit $f : D \to \mathbb{R}$ une fonction et soit $a \in D$.*
- *On dit que f est continue à gauche au point a si $\lim_{\substack{x \to a \\ x < a}} f(x) = f(a)$.*
- *On dit que f est continue à droite au point a si $\lim_{\substack{x \to a \\ x > a}} f(x) = f(a)$.*

Exemple 2.14. La fonction partie entière n'est continue à gauche en aucun entier. Par contre, elle est continue à droite en tous les entiers.

Exemple 2.15. La fonction signe n'est pas continue en 0, ni à gauche ni à droite :

$$\begin{aligned} \operatorname{sgn} : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} 0 \text{ si } x = 0 \\ \frac{x}{|x|} \text{ si } x \neq 0 \end{cases} \end{aligned}$$

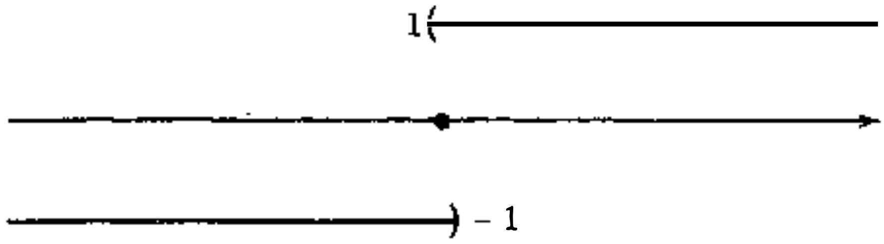


Figure 4.1.2 The signum function

Comme la convergence d'une fonction en un point a se ramène à l'étude de convergences de suites, toutes les opérations sur des limites (produits, sommes, inverses, comparaisons, théorème des gendarmes) sont transposables directement du monde des suites au monde des fonctions sur $\mathbb{R}$. En particulier pour les théorèmes de comparaisons nous définissons les notion suivantes qui sont très naturelles :

Définition 2.16. *Soit $f : D \to \mathbb{R}$ et $g : D \to \mathbb{R}$ deux fonctions, $a \in \overline{D}$ et $b \in D$. On dit que :*

— *f est* inférieure *à g et on note $f \leq g$ dans un voisinage V_a de a (resp. D) si $f(x) \leq g(x)$ pour tout $x \in D \cap V_a$ (resp. D) ;*

— *f est* bornée *dans un voisinage V_a de a (resp. D) si il existe $M > 0$ tel que $|f(x)| \leq M$ pour tout $x \in D \cap V_a$ (resp. D) ;*

— *admet un* maximum *(resp.* minimum*) global en $b \in D$ si $f(x) \leq f(b)$ (resp. $f(b) \leq f(x)$) pour tout $x \in D$.*

Remarque 2.17. Une fonction peut ne pas avoir de maximum ou de minimum global, par exemple la fonction $f : \mathbb{R}^* \to \mathbb{R}, x \mapsto \frac{1}{x}$. D'autre part, il peut y avoir des cas où le maximum (resp. minimum) global peut être atteint en plusieurs point de D, par exemple la fonction $f : [-1, 1] \to \mathbb{R}, x \mapsto x^2$, admet un maximum global en $x = -1$ et $x = 1$.

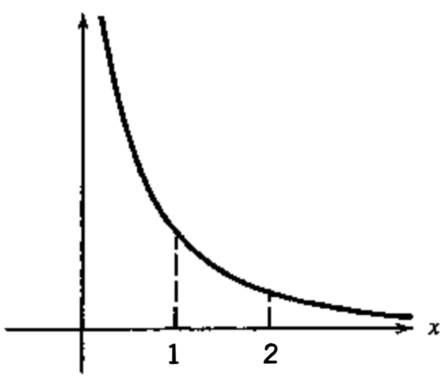


Figure 5.3.1 The function
$f(x) = 1/x \quad (x > 0)$

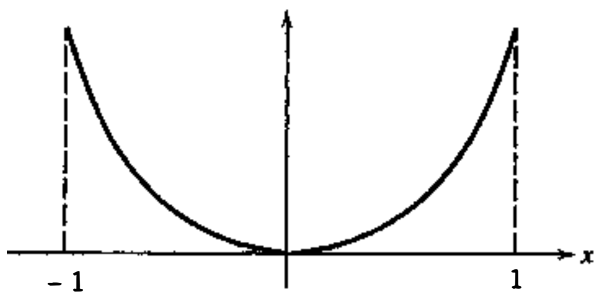


Figure 5.3.2 The function
$g(x) = x^2 \quad (|x| \leq 1)$

En particulier nous avons le résultat suivant qui est l'analogue de la Proposition 1.56 :

Proposition 2.18. Passage à la limite. *Soit* $f : D \to \mathbb{R}$ *et* $g : D \to \mathbb{R}$ *deux fonctions telles que* $f \leq g$ *dans un voisinage d'un point* $a \in \overline{D}$. *Si* $\lim\limits_{\substack{x\to a\\ x\neq a}} f(x) = \ell$ *et* $\lim\limits_{\substack{x\to a\\ x\neq a}} g(x) = \ell'$, *alors* $\ell \leq \ell'$.

Remarque 2.19. Ce résultat est aussi valide pour $f < g$ MAIS on a tjrs $\ell \leq \ell'$. Lorsqu'on passe à la limite, toutes les inégalités strictes deviennent des inégalités non-strictes. Du point de vue topologique : toutes les conditions ouvertes deviennent des conditions fermées. Le slogan c'est qu'"on peut fermer une ouverture, mais pas ouvrir une fermeture" .

La deuxième proposition qu'il faut noter est la version analogue de la Proposition 1.47 :

Proposition 2.20. *Soit* $f : D \to \mathbb{R}$ *et soit* $a \in \overline{D}$ *tel que* f *admet une limite en* a. *Alors* f *est bornée dans un certain voisinage de* a.

Démonstration. Si f admet une limite finie ℓ au point $a \in \overline{D}$, alors pour tout $\epsilon > 0$ il existe $\delta > 0$ tel que pour tout $x \in (D\backslash\{a\} \cap]a-\delta, a+\delta[)$, on a $|f(x) - \ell| < \epsilon$. Posons $I =]a-\delta, a+\delta[$, alors la condition de continuité en a se récrit :

$$\forall\, x \in D \cap I \qquad \Longrightarrow \qquad |f(x) - f(a)| < \epsilon$$

So we have that :

$$\forall\, x \in D \cap I \qquad \Longrightarrow \qquad |f(x)| < \epsilon + |f(a)|$$

C'est la condition de majoration dans un voisinage de a. □

Définition 2.21. *Soit* $f : D \to \mathbb{R}$ *et* $g : D' \to \mathbb{R}$ *deux fonctions, telles que* $f(D) \subset D'$. *On définit alors une troisième fonction de* D *dans* $\mathbb{R}$, *appelée la composée de* f *et* g *et notée* $g \circ f$ *("g rond f"), par :*

$$\forall\, x \in D \qquad g \circ f(x) = g(f(x)).$$

Proposition 2.22. *Supposons que* $\lim\limits_{\substack{x\to a\\ x\neq a}} f(x) = \ell$ *et que* $\lim\limits_{\substack{y\to \ell\\ y\neq \ell}} g(y) = \lambda$, *alors* $\lim\limits_{\substack{x\to a\\ x\neq a}} g \circ f(x) = \lambda$.

Démonstration. Se fait avec les suites. □

Proposition 2.23. *Soit* $f : D \to \mathbb{R}$ *et* $h : D \to \mathbb{R}$ *deux fonctions continues. Alors* $f+h$, fh *sont continues, et si jamais* h *ne s'annule pas, alors* $\frac{f}{h}$ *est continue. Soit* $D' \subset \mathbb{R}$ *un sous-ensemble tel que* $f(D) \subset D'$. *Si* $g : \mathbb{D}' \to \mathbb{R}$ *est continue alors la composition* $g \circ f : D \to \mathbb{R}$ *est continue.*

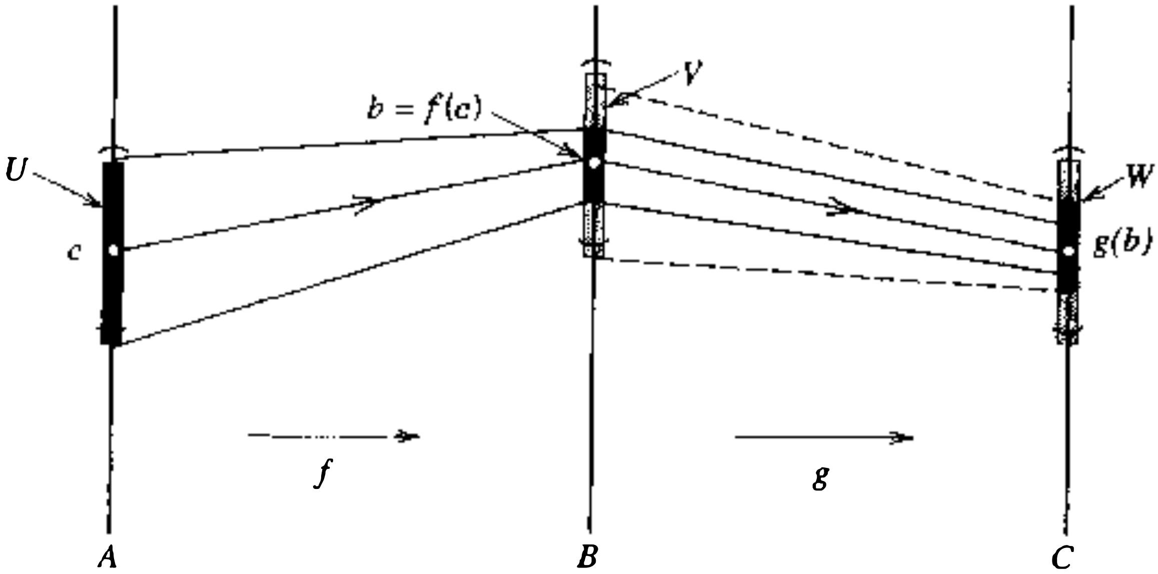


Figure 5.2.1 The composition of f and g

Définition 2.24. *Soient $f : D \to \mathbb{R}$ une fonction continue et $g : D' \to \mathbb{R}$, avec $D \subset D'$. On dit que g* est un prolongement par continuité de f *si*

1. *g est prolongement de f, i.e. $g(x) = f(x)$ pour tout $x \in D$;*
2. *g est continue en tout point de D'.*

Exemple 2.25. Prenons la fonction :

$$\begin{aligned} f :\ \mathbb{R}^* &\longrightarrow \mathbb{R} \\ x &\longmapsto x \sin\left(\frac{1}{x}\right) \end{aligned}$$

tend vers une limite commune nulle en 0. En effet, pour tout $x \in \mathbb{R}^*$, $\left|x \sin\left(\frac{1}{x}\right)\right| \leq |x|$ donc quand x tend vers zero, $f(x)$ tend vers 0 (à gauche comme à droite). On peut donc définir une fonction $g : \mathbb{R} \to \mathbb{R}$ telle que :

$$\begin{aligned} g :\ \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases} \end{aligned}$$

C'est un prolongement par continuité de f en 0.

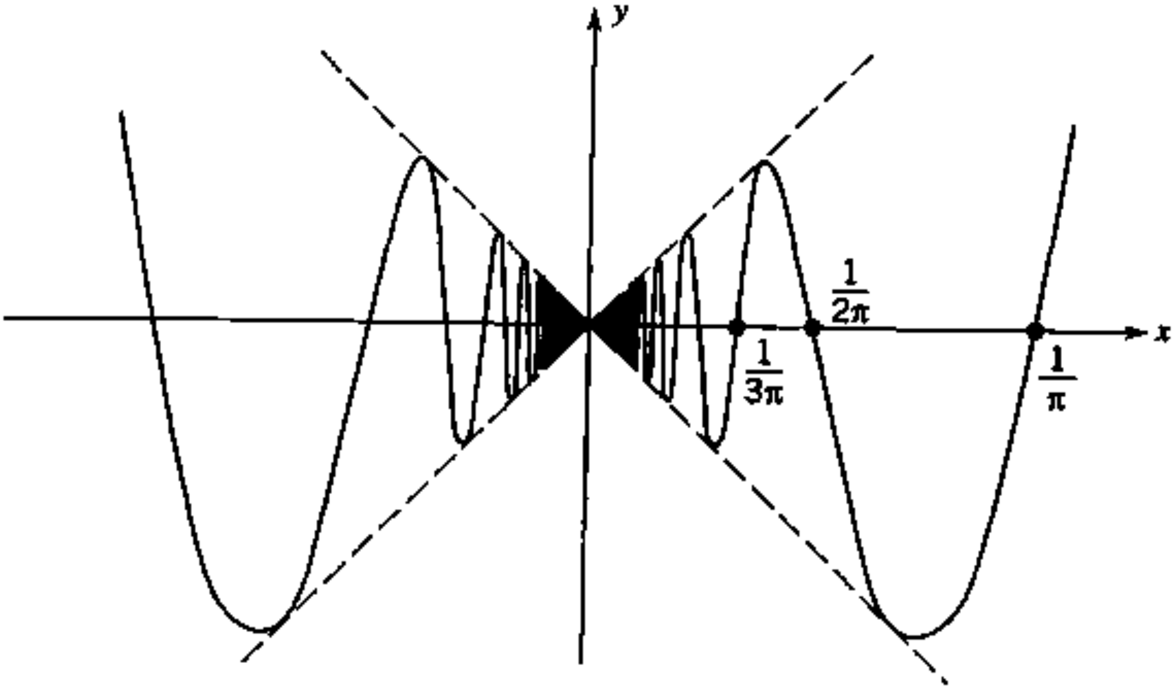


Figure 5.1.3 Graph of $f(x) = x\sin(1/x)$ $(x \neq 0)$

Exemple 2.26. Prenons comme fonction :

$$\begin{aligned} f : \mathbb{R}^* &\longrightarrow \mathbb{R} \\ x &\longmapsto \sin\left(\frac{1}{x}\right) \end{aligned}$$

Comme la fonction $x \mapsto 1/x$ est continue sur $]0, +\infty[$ et sur $]-\infty, 0[$, et que le sinus l'est partout, la composée est continue sur $\mathbb{R}^*$: c'est bien la fonction f. On se demande si f admet une limite à gauche en zéro qui soit égale à la limite à droite, auquel cas on peut définir un prolongement par continuité de f en 0.

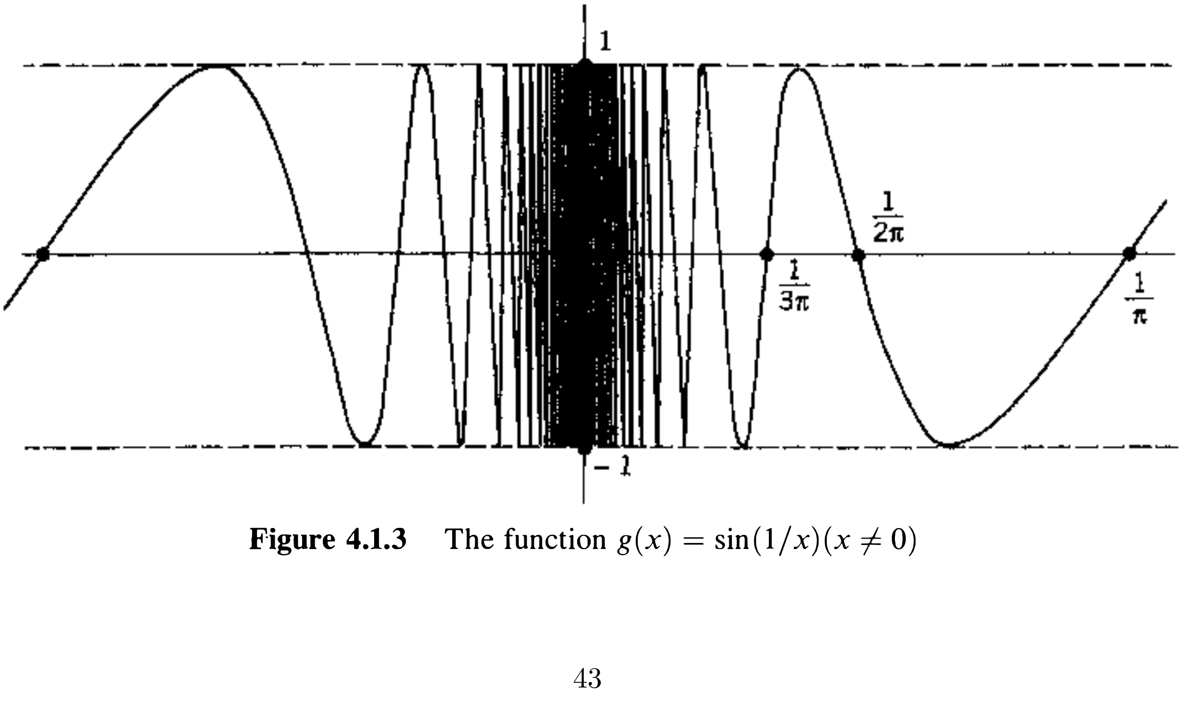


Figure 4.1.3 The function $g(x) = \sin(1/x)(x \neq 0)$

Malheureusement, pour tout $n \in \mathbb{N}^*$, posons $x_n = \frac{2}{\pi n}$. Alors, $f(x_n) = \sin\left(\frac{\pi n}{2}\right) = (-1)^n$. Mais alors, la suite $(x_n)_n$ converge vers 0, mais la suite $(f(x_n))_n$ ne converge pas (alterne), donc f ne tend vers aucune limite en 0 (par la droite, mais idem par la gauche). Donc la fonction n'est certainement pas prolongeable à une fonction continue sur $\mathbb{R}$.

Exemple 2.27. La fonction ζ de Riemann peut se prolonger sur le plan complexe.

Quelque chose de particulier pour les fonctions, c'est qu'on peut aussi tendre vers $\pm\infty$ au point $a \in \overline{D}$, et on peut tendre vers une limite finie ou infinie en $\pm\infty$. Commençons par le premier cas :

Définition 2.28. *Soit $f : D \to \mathbb{R}$ une fonction telle que D et soit $a \in \overline{D}$. On dit que f* tend vers $+\infty$ en a *si :*

$$\forall\, M \in \mathbb{R},\ \exists\, \delta > 0 \text{ tel que } \forall\, x \in D \text{ satisfaisant } \begin{cases} |x-a| < \delta \\ x \neq a \end{cases} \quad \text{on a } f(x) > M$$

et on écrit $\lim\limits_{\substack{x \to a \\ x \neq a}} f(x) = +\infty$. Pour $-\infty$, c'est la même condition mais avec $f(x) < M$.

Remarque 2.29. Comme dans le cas d'une limite finie, on peut définir la divergence vers $\pm\infty$ à gauche ou à droite en a. Cela est par exemple le cas pour la fonction $f : x \mapsto \frac{1}{x}$ définie sur $\mathbb{R}^*$ et qui tend vers $+\infty$ par la droite en 0, et vers $-\infty$ par la gauche en 0.

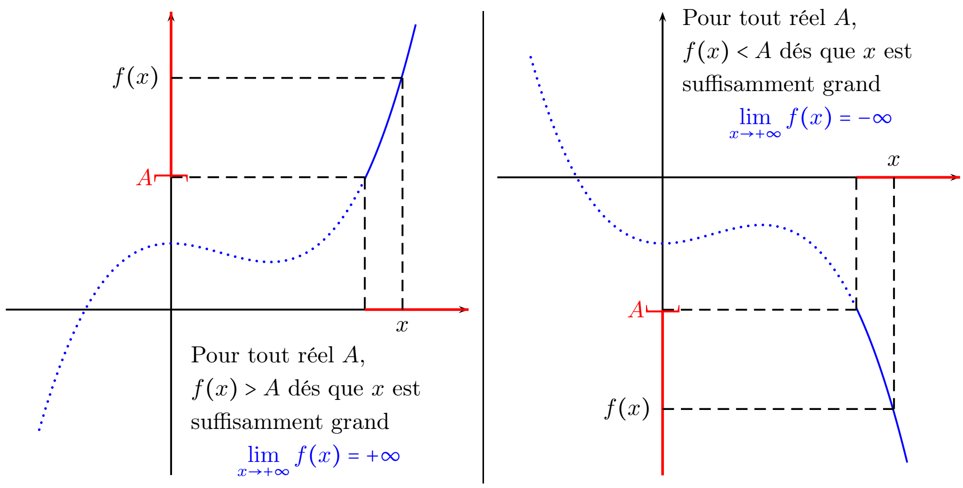
Définition 2.30. *Soit $f : D \to \mathbb{R}$ une fonction telle que D n'est pas majoré, et soit $\ell \in \mathbb{R}$. On dit que f* a pour limite ℓ en $+\infty$ *ou que f* tend vers ℓ quand x tend vers $+\infty$ *et on note $\lim\limits_{x \to +\infty} f(x) = \ell$ si :*

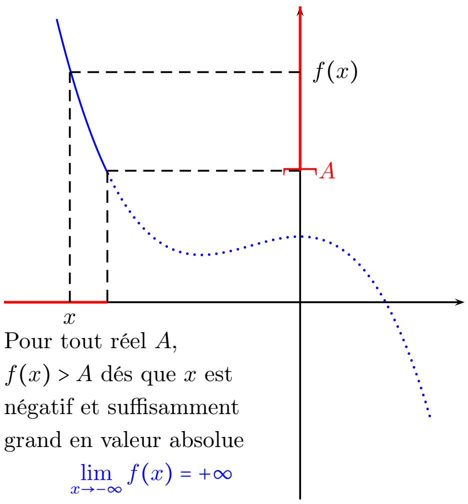
$$\forall\, \epsilon > 0,\ \exists\, M > 0 \text{ tel que } \forall\, x \in D \cap]M, +\infty[, \quad |f(x) - \ell| < \epsilon$$

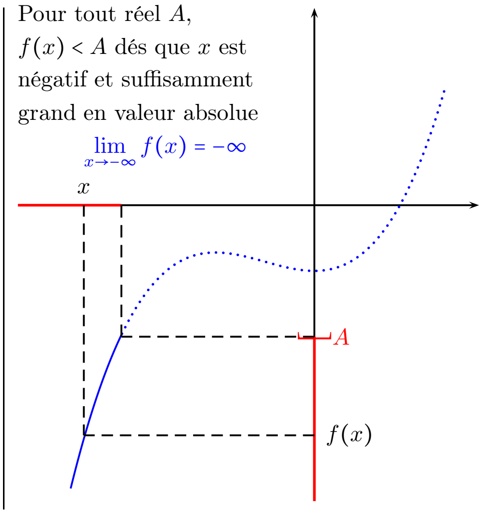
On dit que f tend vers $+\infty$ en $+\infty$ *si :*

$$\forall\, M \in \mathbb{R},\ \exists\, M' > 0 \text{ tel que } \forall\, x \in D \cap]M', +\infty[, \quad f(x) > M$$

et on écrit $\lim\limits_{x \neq +\infty} f(x) = +\infty$. Pour $-\infty$, c'est la même condition mais avec $f(x) < M$.







Remarque 2.31. On peut récrire la même définition avec D non minoré, et on définit le fait que *f tend vers ℓ (resp. $\pm\infty$) quand x tend vers $-\infty$* en écrivant les deux phrases mathématiques suivantes :

$$\forall\, \epsilon > 0,\ \exists\, M > 0 \text{ tel que } \forall\, x \in D \cap]M, +\infty[, \quad |f(x) - \ell| < \epsilon$$
$$\forall\, M \in \mathbb{R},\ \exists\, M' > 0 \text{ tel que } \forall\, x \in D \cap]M', +\infty[, \quad f(x) > M$$

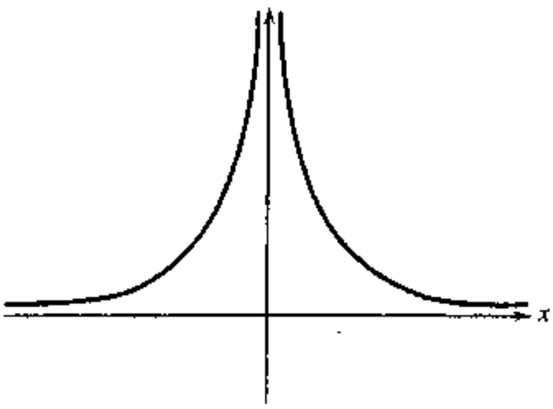


Figure 4.3.3 Graph of $f(x) = 1/x^2 \quad (x \neq 0)$

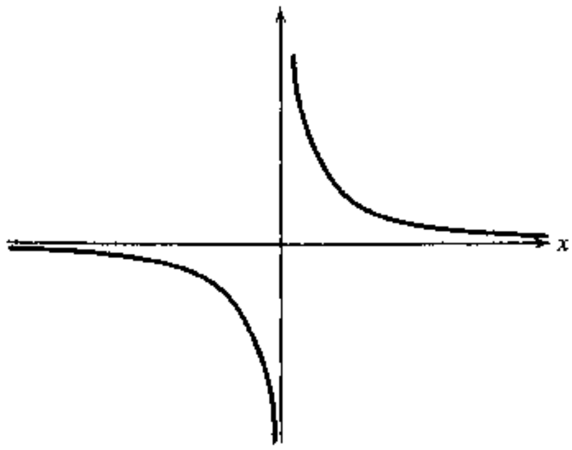


Figure 4.3.4 Graph of $g(x) = 1/x \quad (x \neq 0)$

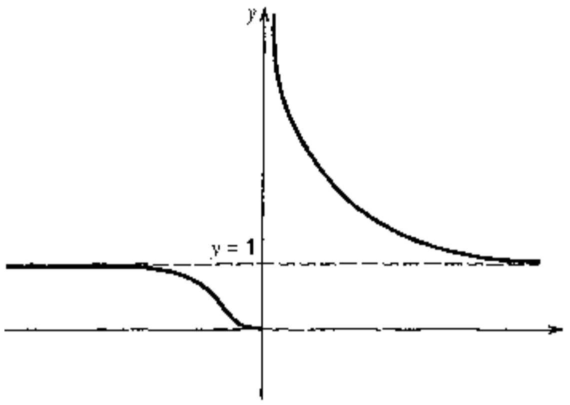


Figure 4.3.1 Graph of
$g(x) = e^{1/x} \quad (x \neq 0)$

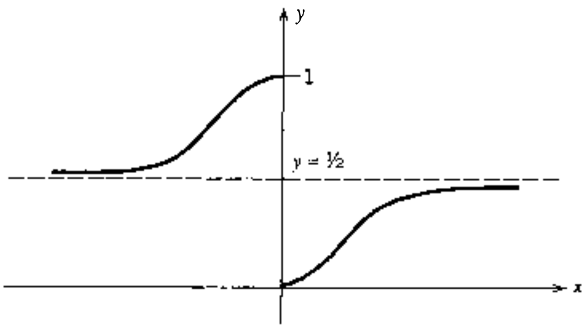


Figure 4.3.2 Graph of
$h(x) = 1/(e^{1/x} + 1) \quad (x \neq 0)$

Remarque 2.32. Lorsque $D = \mathbb{N}$, une fonction $u : \mathbb{N} \to \mathbb{R}$ est en réalité une suite réelle. Les définitions de limite finie (Définition 1.36) et infinie (Définition 1.43) d'une suite réelle sont en fait une application directement de la Definition 2.30 puisque la phrase

$$\forall \epsilon > 0,\ \exists M > 0 \text{ tel que } \forall x \in D \cap]M, +\infty[, \quad |u(x) - \ell| < \epsilon$$

est équivalent à écrire ;

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N} \text{ tel que } \forall x \in \mathbb{N} \cap]N, +\infty[, \quad |u(x) - \ell| < \epsilon$$

qui est à son tour équivalent à écrire la formule de la Définition 1.36 :

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N} \text{ tel que } \forall n \geq, \quad |u(n) - \ell| < \epsilon$$

De même pour les limites $\pm\infty$ d'une suite.

Toutes les définitions du type "f tend vers une limite L quand x tend vers A" (avec $L = \ell$ finie ou $L = \pm\infty$, et $A = a$ un point de $\overline{D}$ ou $A = \pm\infty$), se résume par la formule suivante :

$$\forall V_L \text{ voisinage de } L,\ \exists V_A \text{ voisinage de } A \text{ tel que si } x \in D \cap V_A, \text{ alors } f(x) \in V_L \tag{2.1}$$

On rappelle qu'un *voisinage* de L (c'est similaire pour A en remplaçant L par A) désigne une partie V_L de $\mathbb{R}$ contenant :

- pour $L = \ell$ finie, un intervalle du type $]\ell - \epsilon, \ell + \epsilon[$;
- pour $L = +\infty$, un intervalle du type $]M, +\infty[$;
- pour $L = -\infty$, un intervalle du type $]-\infty, M[$.

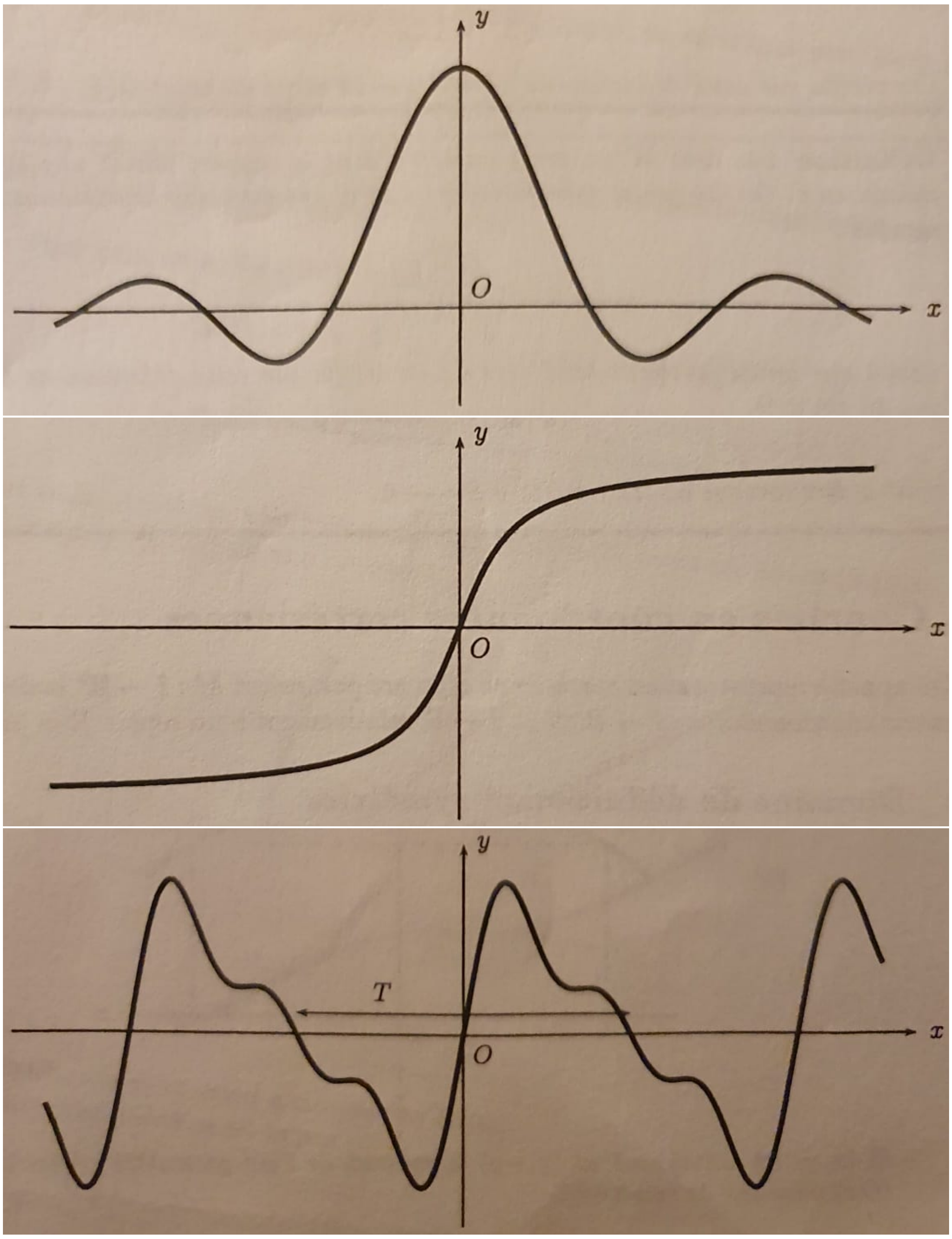
La phrase (2.1) peut être adaptée à chaque cas particulier, en utilisant des voisinages adéquats. Voici le tableau récapitulant toutes les limite possibles, subsumées par la phrase (2.1) :

Voisinage V_L / Voisinage V_A	$\ell \in \mathbb{R}$	$+\infty$	$-\infty$
$a \in \overline{D}$	$\forall]\ell - \epsilon, \ell + \epsilon[$ $\exists]a - \delta, a + \delta[$	$\forall]M, +\infty[$ $\exists]a - \delta, a + \delta[$	$\forall]-\infty, M[$ $\exists]a - \delta, a + \delta[$
$+\infty$	$\forall]\ell - \epsilon, \ell + \epsilon[$ $\exists]M, +\infty[$	$\forall]M, +\infty[$ $\exists]M', +\infty[$	$\forall]-\infty, M[$ $\exists]M', +\infty[$
$-\infty$	$\forall]\ell - \epsilon, \ell + \epsilon[$ $\exists]-\infty, M[$	$\forall]M, +\infty[$ $\exists]-\infty, M'[$	$\forall]-\infty, M[$ $\exists]-\infty, M'[$

Définition 2.33. *Soit $f : D \subset \mathbb{R} \to \mathbb{R}$ une fonction. On dit que :*

- f est paire *si* $\forall\, x \in \mathcal{D},\ -x \in \mathcal{D}$ *et* $f(-x) = f(x)$ *;*
- f est impaire *si* $\forall\, x \in \mathcal{D},\ -x \in \mathcal{D}$ *et* $f(-x) = -f(x)$ *;*
- f est T-périodique*, pour un certain* $T > 0$ *si* $\forall\, x \in \mathcal{D},\ f(x+T) = f(x)$*.*

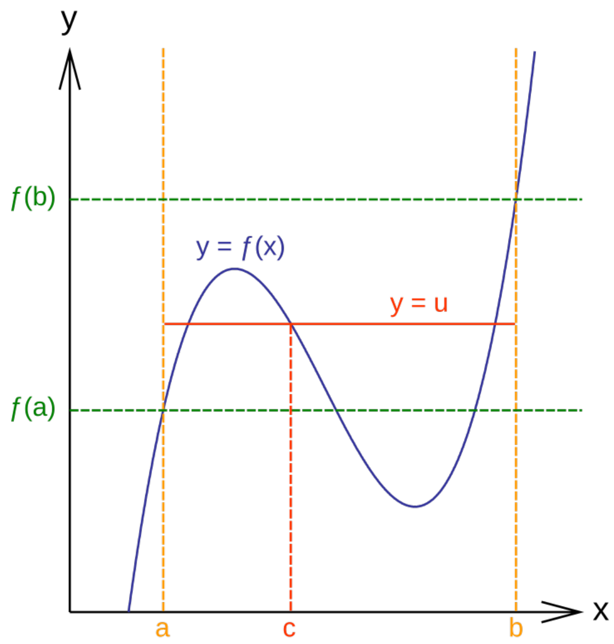
Remarque 2.34. La fonction f est paire si et seulement si son graphe est symétrique par rapport à l'axe vertical. La fonction f est impaire si et seulement si son graphe est symétrique par rapport à l'origine du plan. La fonction f est T-périodique si et seulement si son graphe se répète avec une période T.

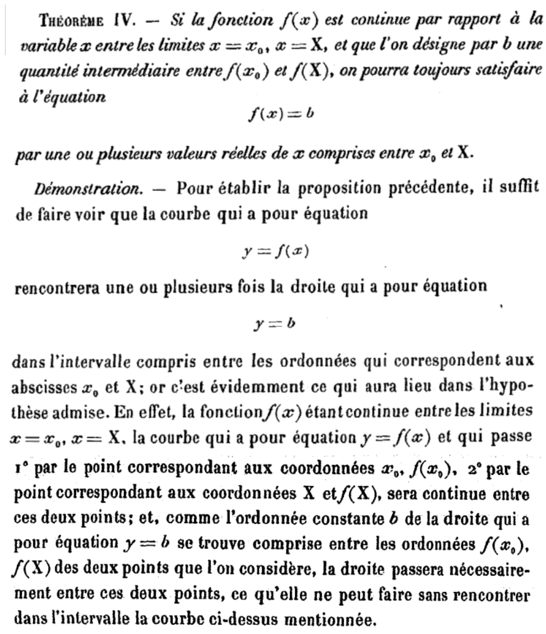


2.2 Propriétés des fonctions continues

Les mathématiques sont une éducation intellectuelle à ne pas prendre les choses qui se présentent à nous comme des évidences. C'est une éducation à embrasser la complexité et à l'accueillir comme une composante normale du monde. Le théorème suivant n'a eu de démonstration que fort tard. Il nécessite en effet une conception claire de la continuité, qui n'est apparue qu'au XIXème. En 1817, Bolzano (1781-1848) rejette les justifications usuelles basées sur des considérations liées à la géométrie, au mouvement, à l'espace, dans un domaine qu'il considère purement analytique. La première définition de la continuité, encore intuitive et sans donner de détails, a été publiée par Cauchy en 1821. On a longtemps pensé que la réciproque du théorème était vraie, et que la propriété des valeurs intermédiaires caractérisait les fonctions continues. Ce n'est qu'en 1875 que Darboux donna un contre-exemple d'une fonction discontinue ayant la propriété des valeurs intermédiaires dans son article Mémoire sur les fonctions discontinues, après avoir réalisé à la suite de Riemann qu'on pouvait intégrer des fonctions discontinues.

Théorème 2.35. Theoreme des valeurs intermédiaires. *Soit $f : [a, b] \to \mathbb{R}$ une fonction continue, alors pour tout réel y compris entre $f(a)$ et $f(b)$, il existe $x \in [a, b]$ tel que $y = f(x)$.*





THÉORÈME IV. — *Si la fonction $f(x)$ est continue par rapport à la variable x entre les limites $x = x_0$, $x = X$, et que l'on désigne par b une quantité intermédiaire entre $f(x_0)$ et $f(X)$, on pourra toujours satisfaire à l'équation*

$$f(x) = b$$

par une ou plusieurs valeurs réelles de x comprises entre x_0 et X.

Démonstration. — Pour établir la proposition précédente, il suffit de faire voir que la courbe qui a pour équation

$$y = f(x)$$

rencontrera une ou plusieurs fois la droite qui a pour équation

$$y = b$$

dans l'intervalle compris entre les ordonnées qui correspondent aux abscisses x_0 et X ; or c'est évidemment ce qui aura lieu dans l'hypothèse admise. En effet, la fonction $f(x)$ étant continue entre les limites $x = x_0$, $x = X$, la courbe qui a pour équation $y = f(x)$ et qui passe 1° par le point correspondant aux coordonnées x_0, $f(x_0)$, 2° par le point correspondant aux coordonnées X et $f(X)$, sera continue entre ces deux points ; et, comme l'ordonnée constante b de la droite qui a pour équation $y = b$ se trouve comprise entre les ordonnées $f(x_0)$, $f(X)$ des deux points que l'on considère, la droite passera nécessairement entre ces deux points, ce qu'elle ne peut faire sans rencontrer dans l'intervalle la courbe ci-dessus mentionnée.

Remarque 2.36. La formulation du théorème peut se faire de façon plus rigoureuse en écrivant que si $f(a) \leq f(b)$, alors pour tout $f(a) \leq y \leq f(b)$, il existe $x \in [a, b]$ tel que $y = f(x)$, et que si $f(a) \geq f(b)$, alors pour tout $f(b) \leq y \leq f(a)$, il existe $x \in [a, b]$ tel que $y = f(x)$. Les deux cas sont essentiellement équivalents et on peut passer de l'un à l'autre en changeant le signe de f, car si f satisfait $f(a) \leq f(b)$, alors $g = -f$ satisfait $g(a) \geq g(b)$.

Démonstration. On suppose pour cette preuve que $f(a) \leq f(b)$. La preuve est très élégante, car soit $y = f(a)$ ou $y = f(b)$ le théorème est valide, soit $f(a) < y < f(b)[$ et alors on va construire deux suites adjacentes $(a_n)_n$ et $(b_n)_n$ qui tendent vers une limite commune x telle que $f(x) = y$. On commence par poser $a_0 = a$ et $b_0 = b$. On va construire la suite par récurrence. Au rang 0 on a $f(a_0) < y < f(b_0)$. On suppose les suites construites jusqu'au rang n telles que :

$$\forall\, 0 \leq k \leq n \qquad \begin{cases} a_k \leq b_k \\ f(a_k) \leq y \leq f(b_k) \end{cases}$$

Pour construire les suites au rang $n+1$, on a alors deux cas :

— soit $f\left(\frac{a_n+b_n}{2}\right) < y$, on pose $a_{n+1} = \frac{a_n+b_n}{2}$ et $b_{n+1} = b_n$, et dans ce cas $a_{n+1} \leq b_n = b_{n+1}$ et $f(a_{n+1}) < y \leq f(b_n) = f(b_{n+1})$;

— soit $f\left(\frac{a_n+b_n}{2}\right) \geq y$, on pose $a_{n+1} = a_n$ et $b_{n+1} = \frac{a_n+b_n}{2}$, et dans ce cas $a_{n+1} = a_n \leq b_{n+1}$ et $f(a_{n+1}) = f(a_n) \leq y \leq f(b_{n+1})$.

Par construction la suite $(a_n)_n$ est croissante et la suite $(b_n)_n$ est décroissante. En outre, on a que :

$$b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2} = \frac{b-a}{2^{n+1}}$$

Donc la suite $(b_n - a_n)_n$ tend vers 0. Donc les deux suites sont adjacentes et convergent vers une limite commune, notée x. Cette limite est dans le segment $[a, b]$ car pour tout n on a $a \leq a_n \leq x \leq b_n \leq b$. Comme la fonction f est continue, on a que $(f(a_n))_n$ et $(f(b_n))_n$ tendent vers $f(x)$ quand n tend vers l'infini. Comme en outre on a l'encadrement $f(a_n) \leq y \leq f(b_n)$, en passant à la limite on a $f(x) \leq y \leq f(x)$ donc $f(x) = y$ (théorème des gendarmes). □

Remarque 2.37. Cette proposition ne dit absolument pas qu'il n'y a qu'une seule préimage à y, il peut y avoir plusieurs voire une infinité de préimages (prendre $f(x) = \sin(x)$ sur $[-\frac{3\pi}{2}, \frac{3\pi}{2}]$ par exemple et $y = 0$), mais juste qu'il en existe au moins une.

Exemple 2.38. Tout polynôme de degré maximal impair admet au moins une racine (réelle). C'est à dire si $P(x) = a_0 + a_1x + a_2x^2 + \ldots + a_{2n+1}x^{2n+1}$, il suffit de voir que le terme dominant de P – en l'occurrence a_{2n+1} en $\pm\infty$ tend vers $\pm\infty$ (en fonction du signe de a_{2n+1}), donc il existe $x \in \mathbb{R}$ tel que $P(x) = 0$.

La propriété des valeurs intermédiaires a été longtemps considéré comme une caractéristique des fonctions continues, c'est à dire qu'une fonction qui vérifie la propriété des valeurs intermédiaires est nécessairement continue. Cette croyance reposait sur la définition naïve et intuitive des fonctions continues qui présidait jusqu'au milieu du XIXème siècle. En 1875, Gaston Darboux mit un terme à cette conviction dans son *Mémoire sur les fonctions discontinues*, en prouvant d'une part que toute fonction dérivée vérifie la propriété des valeurs intermédiaires, et d'autre part qu'il existe des fonctions dérivables dont la fonction dérivée n'est continue sur aucun intervalle. Autrement dit, il existe des fonctions qui satisfont la propriété des valeurs intermédiaires sans être continues. On appelle *fonction de Darboux* toute fonction vérifiant la propriété des valeurs intermédiaires. De telles fonctions sont nombreuses et assez générales pour décrire l'ensemble des fonctions réelles, avec les premières remarques suivantes :

— toute fonction continue est une fonction de Darboux ;

— toute fonction réelle est somme de deux fonctions de Darboux ;

— une fonction de Darboux n'admet que des discontinuités essentielles.

Le théorème de Darboux (voir plus bas) énonce que la dérivée d'une fonction dérivable est une fonction de Darboux. Ce n'est pas une équivalence, donc une fonction de Darboux n'est pas forcément la dérivée d'une autre fonction. Autrement dit une fonction de Darboux n'admet pas forcément de primitive, mais il faut nécessairement être de Darboux pour admettre une primitive (c'est à dire pour être la dérivée d'une fonction). Ce résultat – le Corollaire 2.91 – peut servir à montrer qu'une fonction n'admet pas de primitive, en montrant qu'il existe un intervalle sur lequel cette fonction ne vérifie pas le théorème des valeurs intermédiaires. Un exemple trivial est donné par la fonction partie entière.

Corollaire 2.39. *Soit I un intervalle et $f : I \to \mathbb{R}$ une fonction continue. Alors $f(I)$ est un intervalle.*

Démonstration. Montrer que pour tous $y < w$ éléments de $f(I)$, on a que si $z \in \mathbb{R}$ tel que $y < z < w$, alors $z \in f(I)$. Soit $a \in I$ tel que $y = f(a)$, et $b \in I$ tel que $w = f(b)$. Alors $f(a) < z < f(b)$ donc par le théorème des valeurs intermédiaires, il existe $c \in [a, b]$ tel que $z = f(c)$ donc $z \in f(I)$. □

Remarque 2.40. Rappel qu'un intervalle satisfait la propriété de **convexité** : $\forall y < w \in I$, if $z \in \mathbb{R}$ est tel que $y < z < w$, alors $z \in I$. La partie $A = \{\frac{1}{n} n \in \mathbb{N}^*\}$ n'est pas un intervalle. Le corollaire nous dit qu'une fonction continue f envoie une partie convexe de $\mathbb{R}$ – l'intervalle de définition I – sur une partie convexe de $\mathbb{R}$ – son image $\mathrm{Im}(f)$.

Exemple 2.41. Par contre l'image de l'intervalle I n'est pas forcément de la même nature que I. L'image par $\sin : \mathbb{R} \to \mathbb{R}$ est le segment $[-1, 1]$. Donc un intervalle ouvert peut être envoyé sur un segment. L'image de $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2 + x + 1$ est l'intervalle $[3/4, +\infty[$. L'intervalle $\mathbb{R}$ est envoyé par l'exponentielle sur $]0, +\infty[$. La fonction $f : \mathbb{R} \to \mathbb{R}, x \mapsto \frac{1}{1+x^2}$ a pour image $]0, 1]$.

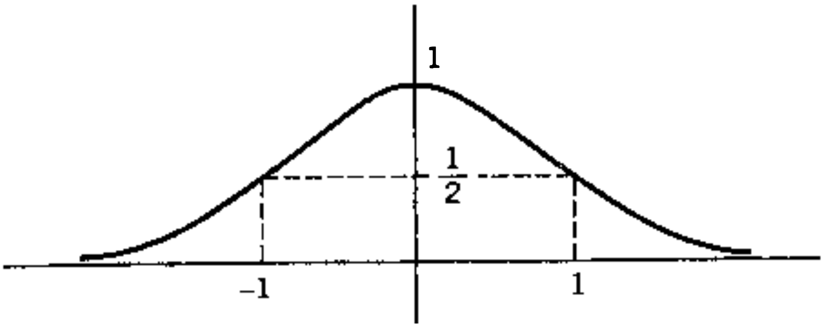


Figure 5.3.4 Graph of $f(x) = 1/(x^2 + 1)$ $(x \in \mathbb{R})$

Tout cela a à voir avec la continuité, qui est en réalité une limite. Or on avait vu dans la Remarque 2.19 qu'au passage à la limite, toute inégalité stricte $f(x) < g(x)$ (condition ouverte) pouvait devenir non-stricte $\ell \leq \ell'$ (condition fermée). En général lors des passages à la limite les conditions ouvertes (strictes) ne se préservent pas et on doit accepter de les affaiblir en conditions fermées (non-strictes). Le slogan on le rappelle c'est qu'"on peut fermer une ouverture, mais pas ouvrir une fermeture" Dans ce contexte, il y a une classe d'intervalles qui est préservée par les fonctions continues : les segments.

Proposition 2.42. *Soit $I = [a, b]$ un segment et $f : [a, b] \to \mathbb{R}$ une fonction continue. Alors $f([a, b])$ est un segment, i.e. f admet un maximum et un minimum sur $[a, b]$.*

Démonstration. Montrons que f est majorée sur $[a, b]$. Supposons qu'elle ne le soit pas. Cela veut dire que pour tout $n \in \mathbb{N}$, il existe $x_n \in [a, b]$ tel que $f(x_n) > n$. Cela nous donne une suite $(x_n)_n$ de points du segment $[a, b]$, telle que la suite $(f(x_n))_n$ tend vers $+\infty$ (ainsi que toute suite extraite donc). Par le théorème de Bolzano-Weierstrass, il existe une sous-suite extraite $(x_{\varphi(n)})_n$ qui converge, vers un point $x \in [a, b]$. Comme f est continue en x, on sait que dans un voisinage de x la fonction f est bornée donc la suite $(f(x_{\varphi(n)}))_n$ est bornée, donc contradiction.

Comme f est majorée sur $[a, b]$, l'ensemble $f([a, b])$ admet une borne supérieure M. Montrons qu'il existe $c \in [a, b]$ tel que $M = f(c)$. Comme M est la borne supérieure, pour tout $\epsilon = \frac{1}{2^n}$, il existe $c_n \in [a, b]$ tel que $M - \frac{1}{2^n} < f(c_n) \leq M$. On a donc une suite $(c_n)_n$ de points du segments. Par Bolzano-Weierstrass à nouveau, on a une sous-suite extraite convergent $c_{\varphi(n)}$ vers un point $c \in [a, b]$. La suite $(f(c_{\varphi(n)}))_n$ converge, par continuité, vers $f(c)$. Or, par définition de la sous-suite extraite, $\varphi : \mathbb{N} \to \mathbb{N}$ est une application strictement croissante, et qui satisfait donc :

$$M - \frac{1}{2^n} < M - \frac{1}{2^{\varphi(n)}} < f(c_{\varphi(n)}) \leq M$$

donc la limite de la suite $(f(c_{\varphi(n)}))_n$ est la borne supérieure M, c'est à dire que $M = f(c)$ donc M est le maximum de la fonction f sur $[a, b]$. La preuve pour le minimum m se fait en reproduisant cette preuve avec $-f$. □

Remarque 2.43. Cette proposition nous dit que l'image $f([a,b])$ est un segment $[m,M]$, cela ne veut pas dire que $[m,M] = [f(a), f(b)]$. La proposition veut dire que la fonction f est bornée et qu'elle atteint sa borne supérieure (resp. inférieure) sur le segment $[a,b]$, ce qui veut dire que c'est un maximum (resp. minimum) global. L'énoncé de la proposition est le cas particulier d'un énoncé en topologie, qui dit que l'image d'un ensemble compact par une fonction continue est un ensemble compact.

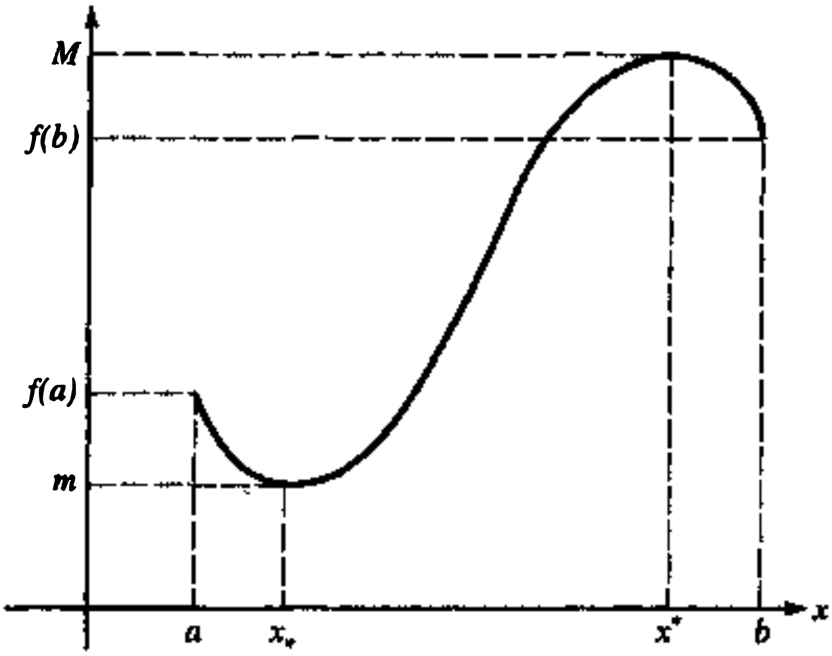


Figure 5.3.3 $f(I) = [m,M]$

Maintenant rappelons quelques notions de théorie des ensembles, appliquées aux fonctions réelles. Soit E et F deux sous-ensembles de $\mathbb{R}$ et soit $f : E \to F$ une fonction. On note D (ou D_f) l'ensemble de définition de l'application f, et $\mathrm{Im}(f) \subset F$ le sous-ensemble de F formé des éléments images de f. On dit que f est :

— injective si pour tous $x \neq y \in D$, on a $f(x) \neq f(y)$

— surjective si pour tout $z \in F$, il existe $x \in D \subset E$ tel que $z = f(x)$

— bijective si f est à la fois injective et surjective

En particulier, l'application $f : D \to \mathrm{Im}(f)$ est surjective. Nous allons maintenant caractériser les cas où la fonction f est injective. Pour cela introduisons les notions suivantes :

Définition 2.44. *Soit $f : D \to \mathbb{R}$ une fonction. On dit que f est :*

— croissante *(resp.* strictement croissante*) si pour tout $x < y \in D$, on a $f(x) \leq f(y)$ (resp. $f(x) < f(y)$) ;*

— décroissante *(resp.* strictement décroissante*) si pour tout $x < y \in D$, on a $f(x) \geq f(y)$ (resp. $f(x) > f(y)$) ;*

— (strictement) monotone *si f est (strictement) croissante ou décroissante.*

Remarque 2.45. Une fonction constante si et seulement si elle est à la fois croissante et décroissante.

Exemple 2.46. La fonction $x \mapsto x^2$ n'est ni croissante ni décroissante sur $\mathbb{R}$, par contre elle est croissante sur $[0, +\infty[$ et décroissante sur $]-\infty, 0]$. La fonction $x \mapsto x^3$ est croissante sur $\mathbb{R}$.

Remarque 2.47. Notons que la notion de croissance, décroissance et monotonie se restreint à celle des suites numériques lorsque $D = \mathbb{N}$.

Il est toujours vrai (voir preuve de la Proposition 2.48) qu'une fonction strictement monotone (continue ou non) est injective. Par contre voici un exemple d'une fonction injective qui n'est pas strictement monotone, car elle est non continue :

$$\begin{array}{rccl} f: & [0,1] & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \begin{cases} x & \text{si } x \in \left[0, \frac{1}{2}\right[\\ \frac{3}{2} - x & \text{si } x \in \left[\frac{1}{2}, 1\right] \end{cases} \end{array}$$

Il y a équivalence des deux notions si la fonction est continue sur un intervalle.

Proposition 2.48. *Soit I un intervalle et soit $f : I \to \mathbb{R}$ une fonction continue. La fonction f est injective si et seulement si elle est strictement monotone.*

Démonstration. Pour montrer l'implication "strictement monotone $\Longrightarrow$ injective", on peut par exemple supposer la fonction strictement croissante. Alors, si $x \neq y \in I$, soit $x < y$ et on a $f(x) < f(y)$, soit $y < x$ donc $f(y) < f(x)$. Dans tous les cas $f(x) \neq f(y)$ donc f est injective. Si la fonction est strictement décroissante, la même preuve fonctionne avec $-f$.

Pour montrer l'implication "injective $\Longrightarrow$ strictement monotone ", nous pouvons montrer la contraposée : "non-strictement monotone $\Longrightarrow$ non-injective",. Soit $a < b < c$ trois points de l'intervalle de définition I tels que $f(b) \leq f(a) \leq f(c)$ (la fonction n'est pas strictement monotone). Comme f est continue, par le théorème des valeurs intermédiaires, la fonction atteint toutes les valeurs entre $f(b)$ et $f(c)$ sur le segment $[b, c]$. En particulier, elle atteint la valeur $f(a)$. Cela veut dire qu'il existe $x \in [b, c]$ – donc $x \neq a$ – tel que $f(x) = f(a)$. Donc le nombre $f(a)$ a au moins deux antécédents différents, donc la fonction f n'est pas injective. □

Cela caractérise toute une famille de fonctions qui sont injectives sur leur ensemble de définition. Comme elle sont toujours surjectives sur leur image, nous avons que les fonctions continues strictement monotones sont des bijections entre leur domaine de définition et leur image. Nous allons maintenant approfondir l'étude de ces fonctions.

Définition 2.49. *Pour toute fonction bijective $f : D_f \to \mathrm{Im}(f)$ il existe une fonction f^{-1} : $\mathrm{Im}(f) \to D_f$ – dénommée "fonction inverse" ou "fonction réciproque" de f – qui est bijective et qui satisfait :*

$$\forall\, x \in D_f, \quad f^{-1}(f(x)) = x, \quad \textit{et} \quad \forall\, y \in \mathrm{Im}(f), \quad f(f^{-1}(y)) = y.$$

Remarque 2.50. Attention à ne pas confondre la notation f^{-1} qui est une fonction et la notation $f^{-1}(V)$ qui est le sous-ensemble de D_f formé de toutes les préimages/antécédents des éléments de V.

Exemple 2.51. La fonction exponentielle $e : \mathbb{R} \to \mathbb{R}$ est strictement croissante donc injective. Elle est aussi surjective sur son image $]0, +\infty[$, donc la corestriction à son image est une bijection de $\mathbb{R}$ dans $]0, +\infty[$.

La définition ne fait pas référence à la continuité de la fonction f, par contre nous avons le résultat suivant dans le cas où f est continue. Ce théorème synthétise à la fois ce que nous savons sur les fonctions continues strictement monotones, et spécifie à ce cas particulier le théorème des valeurs intermédiaires. Là où ce dernier ne nous donnait que l'existence d'une préimage, le théorème suivant nous dit que cette préimage est unique (et donc la fonction inversible) :

Théorème 2.52. Théorème des fonctions réciproques. *Soit I un intervalle et $f : I \to \mathbb{R}$ une fonction continue strictement monotone. Alors :*

1. *la fonction f est une bijection entre I et $\mathrm{Im}(f) = f(I)$;*
2. *la bijection réciproque de f strictement monotone et de même sens de variation que f ;*
3. *la fonction inverse f^{-1} est continue.*

Démonstration. Nous n'allons prouver que le premier point. Comme la fonction est strictement monotone elle est injective. Comme elle est surjective sur $\mathrm{Im}(f) = f(I)$, elle est bijective entre I et $f(I)$. □

Remarque 2.53. Le premier point peut se reformuler comme un cas particulier du Théorème des valeurs intermédiaires. Soit $]a, b[$ un intervalle ouvert (de bornes finies ou infinies) et $f :]a, b[\to \mathbb{R}$ une fonction continue strictement monotone. Alors, pour tout y compris entre la limite de f en a et la limite de f en b, *il existe un unique* $x \in]a, b[$ tel que $y = f(x)$.

Exemple 2.54. Ce théorème permet de définir la fonction racine carrée et plus généralement les fonctions racine n-ième. Quand n est pair, la fonction $x \mapsto x^n$ est strictement décroissante sur $]-\infty, 0]$ et strictement croissante sur $[0, +\infty[$. On choisit par exemple le deuxième intervalle. D'après le théorème, la fonction $x \mapsto x^n$ est une bijection continue strictement croissante de $[0, +\infty[$ sur $[0, +\infty[$. Elle admet donc une fonction inverse bijective continue et strictement croissante sur $[0, +\infty[$. On l'appelle *fonction racine n-ième* et on a note : $\sqrt[n]{} : x \to \sqrt[n]{x}$. Quand n est impair, la fonction $x \mapsto x^n$ est strictement croissant sur $\mathbb{R}$ entier donc la bijection se fait cette fois ci sur $\mathbb{R}$ et la fonction racine n-ième est définie sur $\mathbb{R}$ entier.

Remarque 2.55. Le graphe de la fonction réciproque f^{-1} est le symétrique du graphe de f par rapport à la droite diagonale $y = x$.

Exemple 2.56. On peut tracer la fonction exponentielle et logarithme et voir qu'elles sont effectivement symétriques par rapport à la diagonale d'équation $y = x$.

Exemple 2.57. La fonction sinus est définie sur $\mathbb{R}$, à valeurs dans $[-1, 1]$, et est 2π-périodique. La restriction du sinus au segment $[-\frac{\pi}{2}, \frac{\pi}{2}]$ est strictement croissante et atteint toutes les valeurs entre -1 et 1. Elle définit donc une bijection entre $[-\frac{\pi}{2}, \frac{\pi}{2}]$ et $[-1, 1]$, et admet donc, d'après le théorème, une fonction réciproque. Cette fonction, appelée *arcsinus*, et notée arcsin est continue et strictement croissante, et envoie $[-1, 1]$ sur $[-\frac{\pi}{2}, \frac{\pi}{2}]$, et satisfait :

$$\forall\, x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad \arcsin(\sin(x)) = x \quad \text{et} \quad \forall\, y \in [-1, 1], \quad \sin(\arcsin(y)) = y$$

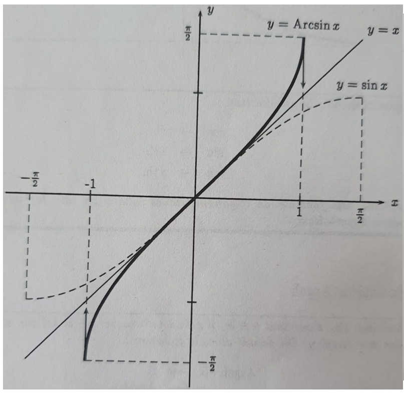
Exemple 2.58. La fonction cosinus est définie sur $\mathbb{R}$, à valeurs dans $[-1, 1]$, et est 2π-périodique. La restriction du cosinus au segment $[0, \pi]$ est strictement décroissante et atteint toutes les valeurs entre -1 et 1. Elle définit donc une bijection entre $[0, \pi]$ et $[-1, 1]$, et admet donc, d'après le théorème, une fonction réciproque. Cette fonction, appelée *arccosinus*, et notée arccos est continue et strictement décroissante, et envoie $[-1, 1]$ sur $[0, \pi]$, et satisfait :

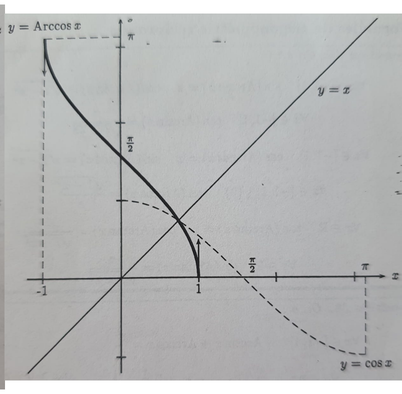
$$\forall\, x \in [0, \pi], \quad \arccos(\cos(x)) = x \quad \text{et} \quad \forall\, y \in [-1, 1], \quad \cos(\arccos(y)) = y$$

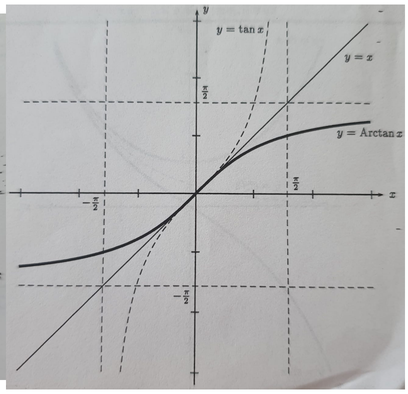
Exemple 2.59. La fonction tangente est définie sur $\mathbb{R} - \mathbb{Z}\frac{\pi}{2}$ et est π-périodique. La restriction du cosinus à l'intervalle ouvert $]-\frac{\pi}{2}, \frac{\pi}{2}[$ est strictement croissante et atteint toutes les valeurs réelles. Elle définit donc une bijection entre $]-\frac{\pi}{2}, \frac{\pi}{2}[$ et $\mathbb{R}$, et admet donc, d'après le théorème, une fonction réciproque. Cette fonction, appelée *arctangente*, et notée arctan est continue et strictement croissante, et envoie $\mathbb{R}$ sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$, et satisfait :

$$\forall\, x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[, \quad \arctan(\tan(x)) = x \quad \text{et} \quad \forall\, y \in \mathbb{R}, \quad \tan(\arctan(y)) = y$$

Remarque 2.60. Ce dernier exemple nous dit que $\mathbb{R}$ est en bijection avec $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, et donc en réalité avec tout intervalle ouvert $]a,b[$ où $a<b$, car on peut toujours envoyer $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ sur $]a,b[$ par translation et dilatation.







Définition 2.61. *Soit I un intervalle et soit $f : I \to \mathbb{R}$ une fonction continue. On dit que f est un* homéomorphisme *si la corestriction de f à son image $f|^{\mathrm{Im}(f)} : I \mapsto \mathrm{Im}(f)$ est une bijection et sa réciproque $f^{-1} : \mathrm{Im}(f) \to I$ est continue.*

Remarque 2.62. Attention, l'hypothèse que le domaine de définition de la fonction est un intervalle est importante. En effet, la fonction

$$\begin{aligned} g :\ [0,1[\,\cup\,[2,3] &\longrightarrow [0,2] \\ x &\longmapsto \begin{cases} x & \text{si } x \in [0,1[\\ x-1 & \text{si } x \in [2,3] \end{cases} \end{aligned}$$

est une bijection continue strictement croissante sur son ensemble de définition mais la réciproque :

$$\begin{aligned} g^{-1} :\ [0,2] &\longrightarrow [0,1[\,\cup\,[2,3] \\ x &\longmapsto \begin{cases} x & \text{si } x \in [0,1[\\ x+1 & \text{si } x \in [1,2] \end{cases} \end{aligned}$$

est une bijection strictement croissante mais discontinue en $g(1) = 1$.

2.3 Dérivabilité

Après la continuité, et avec les outils qu'on a à portée de main, on peut définir la dérivabilité d'une fonction réelle. C'est l'outil principal pour étudier une fonction. Ici nous procéderons comme avec les fonctions continues. D'abord nous définirons le nombre dérivé d'une fonction en un point a, ainsi que les nombres dérivés à gauche et à droite, puis nous en déduirons certaines conséquences importantes. Enfin nous établirons l'existence des fonctions dérivées si une fonction est dérivable en chacun des points de son intervalle de définition. A partir d'ici, tous les domaines de définition sont des intervalles ou des unions finies d'intervalles.

Définition 2.63. *Soit $f : D \to \mathbb{R}$ et $a \in D$. On dit que f* est dérivable en *a si la limite*

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

existe et est finie. On dénote cette limite $f'(a)$ (c'est un nombre réel) et on l'appelle nombre dérivé *ou* dérivée de *f* en *a.*

Exemple 2.64. Soit la fonction carré : $f : x \mapsto x^2$ et $a \in \mathbb{R}$. Alors $\frac{x^2-a^2}{x-a} = x + a$, qui tend vers $2a$ lorsque x tend vers a. Donc f est dérivable en a et sa dérivée est $2a$. C'est bien le résultat attendu.

Comme la continuité à gauche et à droite, on parle aussi de dérivabilité à gauche et à droite d'un point $a \in D$, si on se contraint à rester que d'un côté du point a.

Par exemple, le nombre dérivé à gauche (resp. à droite) en a est la limite suivante, si elle est finie :

$$\lim_{\substack{x\to a\\ x<a}} \frac{f(x)-f(a)}{x-a} \qquad \left(\text{resp. } \lim_{\substack{x\to a\\ x>a}} \frac{f(x)-f(a)}{x-a}\right)$$

Une fonction est dérivable en a si elle est dérivable à gauche et à droite, et que les deux valeurs coincident.

Définition 2.65. *Soit $f : D \to \mathbb{R}$ une fonction et soit $a \in D$. On dit que :*

— f est dérivable à gauche en a *si la limite* $f'_g(a) = \lim_{\substack{x\to a\\ x<a}} \frac{f(x)-f(a)}{x-a}$ *existe et est finie.*

— f est dérivable à droite en a *si la limte* $f'_d(a) = \lim_{\substack{x\to a\\ x>a}} \frac{f(x)-f(a)}{x-a}$ *existe et est finie.*

Proposition 2.66. *Soit $f : D \to \mathbb{R}$ une fonction et soit $a \in D$. f est dérivable au point a si et seulement si elle est dérivable à gauche et à droite en a et que $f'_g(a) = f'_d(a)$.*

Exemple 2.67. Prenons la fonction valeur absolue $|.| : x \mapsto |x|$ et $a = 0$. La dérivée à droite est :

$$f'_d(0) = \lim_{\substack{x\to 0\\ x>0}} \frac{|x|-|0|}{x-0} = \lim_{\substack{x\to 0\\ x>0}} \frac{|x|}{x} = 1$$

tandis que la dérivée à gauche est :

$$f'_g(0) = \lim_{\substack{x\to 0\\ x<0}} \frac{|x|-|0|}{x-0} = \lim_{\substack{x\to 0\\ x<0}} \frac{|x|}{x} = -1$$

Donc la dérivée à gauche et à droite de la fonction valeur absolue ne coincident pas en zéro, donc la fonction n'est pas dérivable en zéro.

Supposons $f : D \to \mathbb{R}$ dérivable en $a \in D$ et définissons une fonction $\epsilon : D \to \mathbb{R}$ en posant :

$$\forall x \in D, \qquad \epsilon(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} - f'(a) \text{ si } x \neq a \\ 0 \text{ si } x = a \end{cases}$$

Il vient que la fonction ϵ est continue au point a car

$$\lim_{\substack{x\to a\\ x\neq a}} \epsilon(x) = \lim_{\substack{x\to a\\ x\neq a}} \frac{f(x)-f(a)}{x-a} - f'(a) = f'(a) - f'(a) = 0 = \epsilon(a)$$

L'existence d'une telle fonction ϵ caractérise en fait les fonctions dérivables :

Proposition 2.68. *La fonction $f : D \to \mathbb{R}$ est dérivable en $a \in D$ si et seulement si il existe un nombre réel α et une fonction $\epsilon : D \to \mathbb{R}$ tels que :*

— *ϵ est continue en a, et $\epsilon(a) = 0$;*

— *$f(x) = f(a) + \alpha(x-a) + (x-a)\epsilon(x)$ pour tout $x \in D$.*

Dans ce cas, le nombre α est égal à $f'(a)$.

Démonstration. Une direction a déjà été montrée. Montrons la réciproque en supposant qu'une telle fonction existe et que f est dérivable au point a. Pour tout $x \neq a$, on a :

$$\frac{f(x) - f(a)}{x - a} = \alpha + \epsilon(x)$$

donc quand x tend vers a, $\epsilon(x)$ tend vers 0 par continuité, donc on déduit que le quotient de gauche tend vers la limite finie α. Cela veut dire que la fonction f est dérivable en 0 et le nombre α est par définition le nombre dérivé $f'(a)$. □

Corollaire 2.69. *Si la fonction $f : D \to \mathbb{R}$ est dérivable en $a \in D$, alors elle est continue en a.*

Démonstration. On a $f(x) = f(a) + \alpha(x - a) + (x - a)\epsilon(x)$ pour tout $x \in D$, or toutes les fonctions à droites sont continues en a, donc la fonction à gauche est continue en a. □

Démonstration. On a une preuve plus longue et plus formelle, qui ne s'appuie par sur la Proposition 2.68. Soit $\epsilon > 0$. Comme f est dérivable en a, le qutoient $\frac{f(x)-f(a)}{x-a}$ tend vers une limite finie en a, notée $f'(a)$. Alors on a qu'il existe $\delta > 0$ tel que pour tout $x \in D \cap]a - \delta, a + \delta[$ et différent de a, on a :

$$\left| \frac{f(x) - f(a)}{x - a} - f'(a) \right| < \epsilon$$

Quitte à prendre δ encore plus petit (c'est possible), on peut prendre δ tel que $\delta < \frac{\epsilon}{\epsilon + |f'(a)|}$. Or nous savons que $|u - v| < \epsilon$ implique $|u| < \epsilon + |v|$ donc on a :

$$\left| \frac{f(x) - f(a)}{x - a} \right| < \epsilon + |f'(a)| \text{ c'est à dire } |f(x) - f(a)| < (\epsilon + |f'(a)|) |x - a| < (\epsilon + |f'(a)|) \delta$$

Mais par notre choix de δ, on a précisément que $(\epsilon + |f'(a)|) \delta < \epsilon$, donc on a que :

$$|f(x) - f(a)| < \epsilon$$

c'est à dire que la $f(x)$ tend vers $f(a)$ quand x tend vers a : f est continue en a. □

Remarque 2.70. Le corollaire 2.69 s'étend tout naturellement aux dérivées à gauche et à droite : si une fonction $f : D \to \mathbb{R}$ est dérivable à gauche (resp. à droite) en $a \in D$, alors elle est continue à gauche (resp. à droite) en a.

Nous pouvons interpréter graphiquement la dérivée en un point. Soit $f : D \to \mathbb{R}$ une fonction, alors on note $\mathrm{Gr}(f)$ le graphe de f dans l'espace $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. C'est un sous-ensemble de $\mathbb{R}^2$:

$$\mathrm{Gr}(f) = \{(x, f(x)),\ x \in D\} \subset \mathbb{R}^2$$

Pour tout $x \in D$, on définit M_x le point du graphe de coordonnées $(x, f(x))$. Supposons que f est dérivable en $a \in D$ et soit $M_a = (a, f(a))$ le point du graphe de f correspondant à ce choix. On prend désormais $x \neq a$. La droite du plan qui passe par M_a et M_x a pour équation :

$$y(t) = \frac{t - a}{x - a}(f(x) - f(a)) + f(a) = \frac{f(x) - f(a)}{x - a} t + \frac{x f(a) - a f(x)}{x - a}$$

C'est bien une équation de droite car le membre de droite est linéaire en t. Cette équation est telle que le point d'abscisse $t = a$ (resp. $t = x$) a pour ordonnée $y(a) = f(a)$ (resp. $y(x) = f(x)$), c'est donc M_a (resp. M_x). On a donc bien l'équation de la droite $(M_a M_x)$. Son coefficient directeur (la pente) est $\frac{f(x)-f(a)}{x-a}$ et l'ordonnée à l'origine est $\frac{xf(a)-af(x)}{x-a}$. On voit que lorsque x tend vers a, le coefficient directeur de cette droite tend vers $f'(a)$.

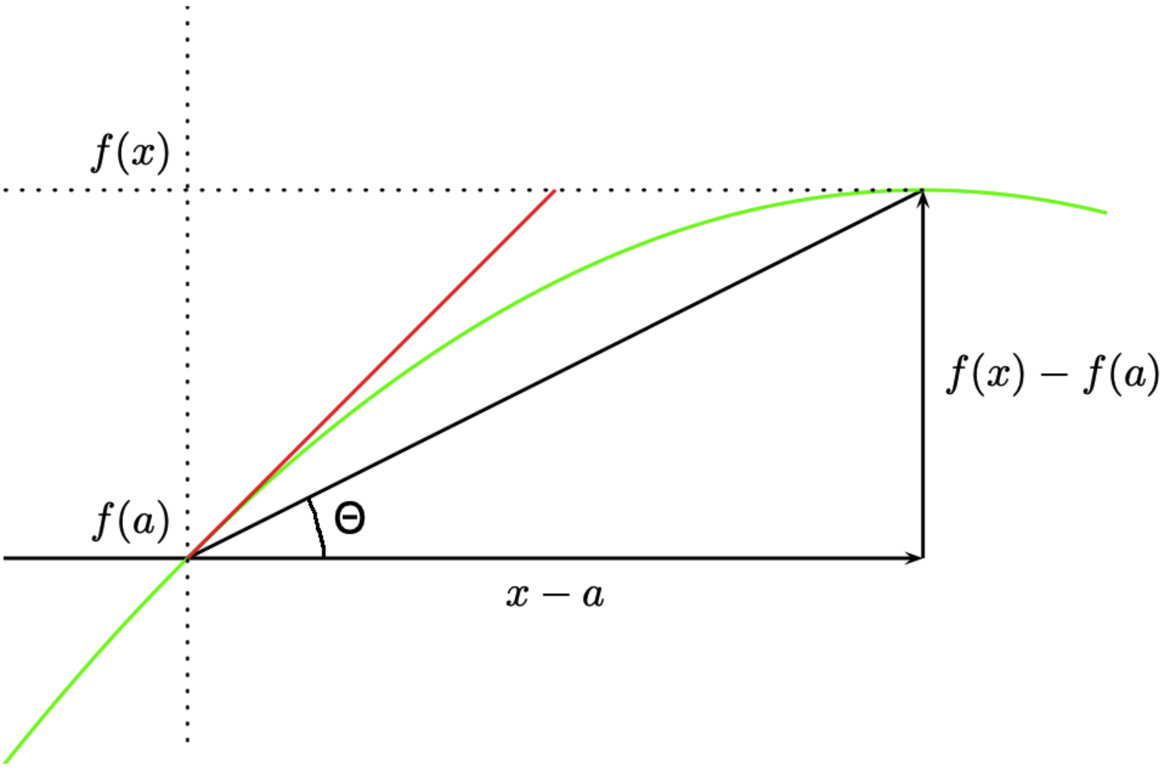


FIGURE 3 – Le coefficient directeur de la droite passant en M_a et M_x est la tangente de l'angle entre la droite et l'axe horizontale : $\tan(\Theta)$. C'est à dire c'est le quotient du côté opposé sur le côté adjacent, et on retrouve bien $\frac{xf(a)-af(x)}{x-a}$. La droite rouge marque la demi-tangente à droite au graphe de f au point a.

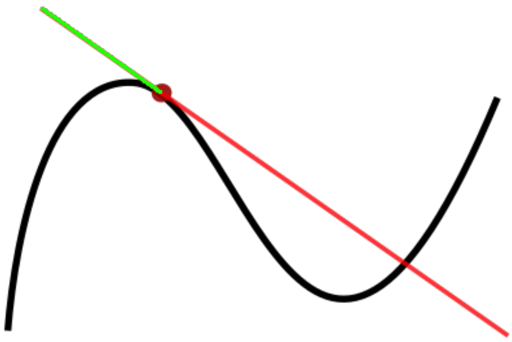
La *tangente* au graphe de f au point a est la droite obtenue comme limite de la droite (M_aM_x) lorsque le point M_x se rapproche du point M_a. Or lorsque x tend vers a, le coefficient directeur de la droite (M_aM_x) tend vers le nombre dérivé en a. Le nombre dérivé $f'(a)$ est donc le coefficient directeur de la droite tangente au graphe de f au point a. Lorsque la fonction $f : D \to \mathbb{R}$ est dérivable à gauche ou à droite, on dit qu'elle admet une demi-tangente à gauche ou à droite du point $M_a = (a, f(a))$. La demi-tangente à gauche est une demi-droite qui commence au point M_a, qui est tangente au graphe de f en M_a et qui se dirige vers l'arrière, c'est à dire vers les $x < a$. La demi-tangente à droite est une demi-droite qui commence au point M_a, qui est tangente au graphe de f en M_a et qui se dirige vers l'avant, c'est à dire vers les $x > a$.

Proposition 2.71. *Soit $f : D \to \mathbb{R}$ une fonction. Si f est dérivable en a, alors l'équation de la droite tangente au graphe de f au point $(a, f(a))$ a pour équation :*

$$y = f(a) + (x - a)f'(a) \tag{2.2}$$

Si f est dérivable à gauche (resp. à droite) en a, la demi-tangente à gauche (resp. à droite) au graphe de f au point $(a, f(a))$ est une demi-droite qui commence au point $(a, f(a))$ et qui part vers $x < a$ (resp. $x > a$), ayant pour équation (2.2).

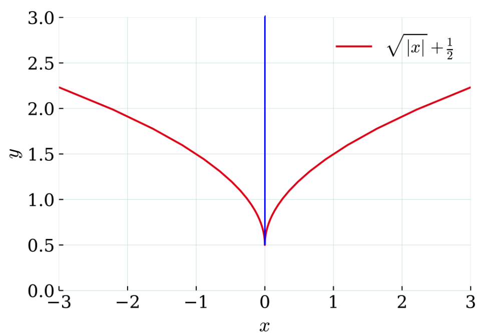
Si la fonction f est dérivable à gauche et à droite, et que la dérivée à gauche et à droite coincident, cela s'interprète géométriquement par le fait que les deux demi-tangentes à gauche et à droite sont exactement alignées et coincident uniquement au point M_a pour former une droite complète tangente au graphe de f au point M_a. Dans ce cas on retrouve bien géométriquement le résultat que la fonction est dérivable en a. Dans le graphique suivant, la demi-tangente à gauche est verte, la demi-tangente à droite est rouge.

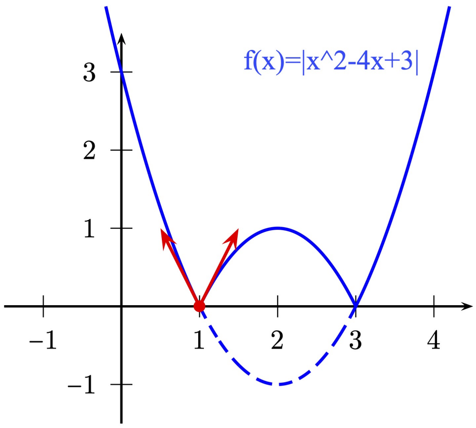


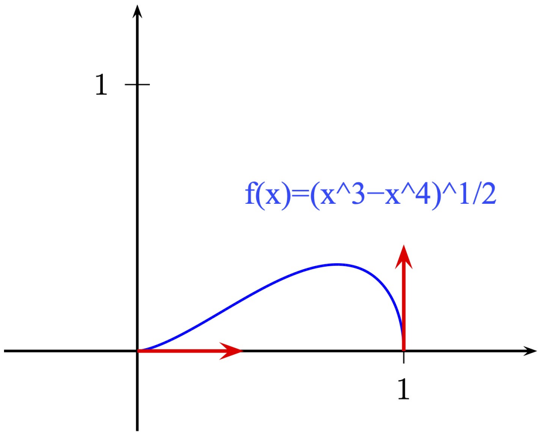
Exemple 2.72. Pour la fonction valeur absolue, la dérivée à gauche est $f'_g(0) = -1$ donc la demi-tangente à gauche est la demi-droite d'équation $y = -x$ commençant à (0,0) et s'éloignant dans le cadrant supérieur gauche (remontant le graphe vers l'arrière). La dérivée à droite est $f'_d(0) = 1$ donc la demi-tangente à droite est la demi-droite d'équation $y = x$ commençant à (0,0) et s'éloignant dans le cadrant supérieur droit (continuant le graphe vers l'avant).

Remarque 2.73. Si la limite à gauche $\lim\limits_{\substack{x\to a\\ x<a}} \frac{f(x)-f(a)}{x-a}$ est infinie, alors on convient que la demi-tangente à gauche est verticale (le sens de la demi-droite dépend de quel infini). C'est le même raisonnement pour la limite à droite.

Prenons l'exemple de la fonction $f : x \mapsto \sqrt{|x|} + \frac{1}{2}$ et $a = 0$. Comme ici $\frac{f(x)-f(0)}{x-0} = \frac{\sqrt{|x|}}{x}$, la limite à gauche de zéro est $-\infty$ donc la demi-tangente à gauche est verticale ascendante (car la demi-tangente à gauche remonte le graphe de f vers les $x < 0$), tandis que la limite de zéro à droite est $+\infty$ donc la demi-tangente à droite est verticale ascendante (car la demi-tangente à droite suit le graphe de f vers les $x > 0$) Les deux demi-tangentes (en bleu) sont donc confondues.







Avec l'interprétation graphique de la dérivée, on peut comprendre le lien avec l'existence de minimum et de maximum d'une fonction continue.

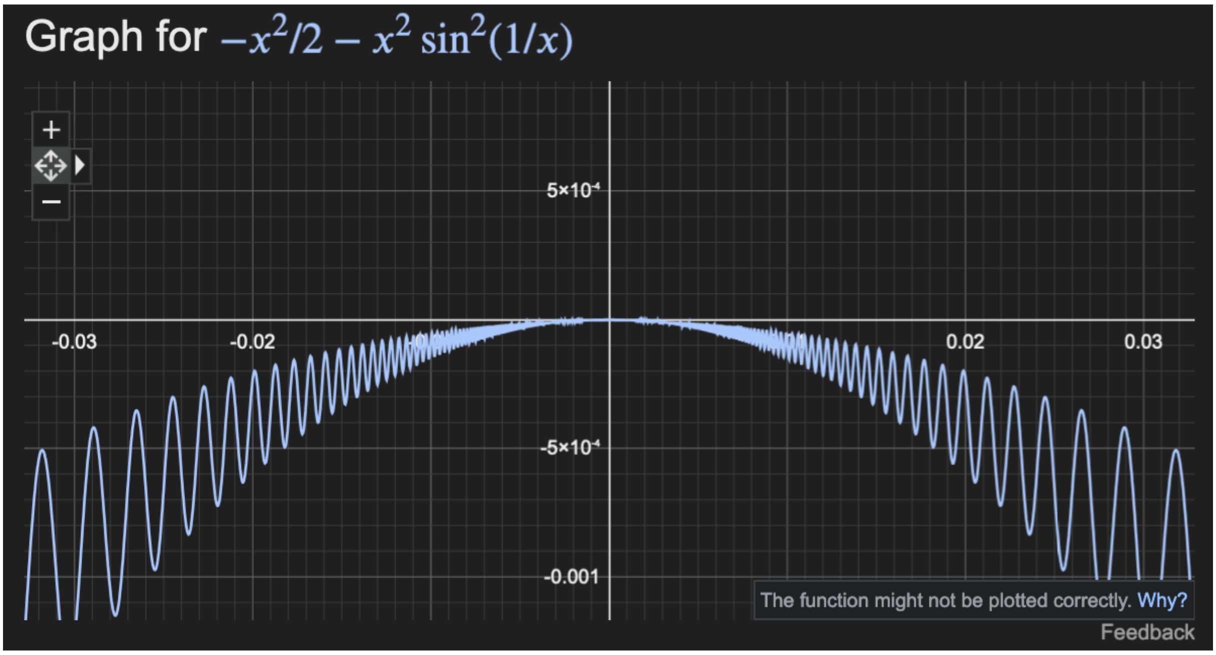
Définition 2.74. *Soit $f : D \to \mathbb{R}$ une fonction (pas nécessairement continue globalement), et soit $a \in D$. On dit que :*

- *f a un maximum (resp. minimum) global en a si, pour tout $x \in D$, $f(x) \leq f(a)$ (resp. $f(x) \geq f(a)$) ;*
- *f a un maximum (resp. minimum) local en a s'il existe un voisinage V de a tel que, pour tout $x \in V \cap D$, $f(x) \leq f(a)$ (resp. $f(x) \geq f(a)$) ;*
- *f a un extremum global (resp. local) en a si f a un maximum ou un minimum global (resp. local) en a.*

Remarque 2.75. Attention au pluriel la terminaison est -a plutôt que -um.

Exemple 2.76. Prenons la fonction définie sur $\mathbb{R}$ par $f(x) = |1 - x^2|$. Les minima locaux (et globaux) sont -1 et 1, et un maximum local est en 0. Pas de maxima globaux.

Remarque 2.77. Il est faux de croire que, si a est un maximum, alors f est croissante à gauche de a, puis décroissante après. $f(a)$ est certes la valeur maximale, mais f peut ne pas être monotone, ni à gauche, ni à droite. Ce genre de faits contre-intuitifs viennent du fait que notre intuition se base sur les dessins que l'on fait au tableau à main levée, or ceux-ci ne représentent qu'une infime partie des types de fonctions. Les exemples contre-intuitifs viennent des fonctions non dessinables justement : les fonctions nulle-part continues (la fonction de Dirichlet), les fonctions continues nulle-part dérivables (fonction de Weierstrass), et ici les fonctions majorée nulle-part monotones. Prendre par exemple la fonction définie pour tout $x \in \mathbb{R}$ par $f(x) = -\frac{x^2}{2} - x^2 \sin^2\left(\frac{1}{x}\right)$. C'est une fonction qui oscille entre les deux polynômes d'équation $x \mapsto -\frac{x^2}{2}$ et $x \mapsto -\frac{3x^2}{2}$. 0 est un maximum global de la fonction mais elle oscille infiniment dans tout voisinage de 0.



L'interprétation géométrique des extremas d'une fonction se fait simplement : avoir un maximum local en a veut dire que le graphe de f est *sous* la tangente à la courbe dans un voisinage de a assez petit, tandis qu'avoir un minimum local en a veut dire que le graphe de f est *au dessus* de la tangente à la courbe dans un voisinage de a assez petit. Nous en déduisons directement le résultat suivant lorsque f est dérivable :

Proposition 2.78. *Soit $f : D \to \mathbb{R}$ une fonction. Supposons que f admette un extremum local au point $a \in D$, et que f est dérivable en a. Alors $f'(a) = 0$.*

Démonstration. Disons que l'extremum est un maximum, et que le voisinage de a sur lequel $f(x) \le f(a)$ est de la forme $]a-\epsilon, a+\epsilon[$, pour un certain $\epsilon > 0$. Alors, on a les inégalités suivantes :

$$\text{pour tout } x \in]a-\epsilon, a[\cap D \text{ on a } f(x) \le f(a) \text{ et } x - a < 0 \text{ donc } \frac{f(x)-f(a)}{x-a} \ge 0$$
$$\text{pour tout } x \in]a, a+\epsilon[\cap D \text{ on a } f(x) \le f(a) \text{ et } x - a > 0 \text{ donc } \frac{f(x)-f(a)}{x-a} \le 0$$

Donc, par passage à la limite (dont on sait qu'elle existe car f est dérivable en a), on a d'une part en haut $f'_g(a) \ge 0$ et d'autre part en bas $f'_d(a) \le 0$. Comme $f'_g(a) = f'_d(a) = f'(a)$, on en déduit que $f'(a) = 0$. Si l'extremum de f est un minimum, alors $-f$ a un maximum en ce point et on prend le même argument. □

Remarque 2.79. Attention, la réciproque de la Proposition 2.78 est fausse car on peut avoir $f'(a) = 0$ sans que f n'atteigne un extremum en a. C'est par exemple le cas de la fonction $x \mapsto x^3$ dont la dérivée s'annule en $a = 0$ mais qui n'a pas de maximum ou de minimum nulle part. On appelle un tel point un *point d'inflexion*

Remarque 2.80. La Proposition 2.78 justifie qu'on cherche à calculer les dérivées des fonctions pour trouver leurs extremas, et elle est en fait au fondement de toute la physique moderne depuis le XVIII ème siècle. Le *principe de moindre action* de Maupertuis établit en effet que lorsqu'un changement physique advient dans un système, la quantité d'action nécessaire à ce changement est la plus petite possible. Le principe de Maupertuis (1744) généralise à tout système physique le principe de Fermat (1657) qui ne concernait que la lumière et qui disait que la lumière se

$\mathcal{D}_f$	$f(x)$	$\mathcal{D}_{f'}$	$f'(x)$
$\mathbb{R}$	$x^n \quad (n \in \mathbb{N})$	$\mathbb{R}$	$\begin{cases} nx^{n-1} & \text{si } n \geqslant 1 \\ 0 & \text{si } n = 0 \end{cases}$
$\mathbb{R}^*$	$x^n \quad (n \in \mathbb{Z})$	$\mathbb{R}^*$	nx^{n-1}
$\mathbb{R}_+^*$	$x^\alpha \quad (\alpha \in \mathbb{R})$	$\mathbb{R}_+^*$	$\alpha x^{\alpha-1}$
$\mathbb{R}_+$	$\sqrt[n]{x} = x^{\frac{1}{n}} \quad (n \in \mathbb{N}^*)$	$\mathbb{R}_+^*$	$\frac{1}{n} x^{\frac{1}{n}-1}$
$\mathbb{R}$	e^x	$\mathbb{R}$	e^x
$\mathbb{R}_+^*$	$\ln x$	$\mathbb{R}_+^*$	$\frac{1}{x}$
$\mathbb{R}^*$	$\ln \lvert x \rvert$	$\mathbb{R}^*$	$\frac{1}{x}$
$\mathbb{R}$	$\cos x$	$\mathbb{R}$	$-\sin x$
$\mathbb{R}$	$\sin x$	$\mathbb{R}$	$\cos x$
$\mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right)$	$\tan x$	$\mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right)$	$1 + \tan^2 x = \frac{1}{\cos^2 x}$
$\mathbb{R} \setminus \pi\mathbb{Z}$	$\operatorname{cotan} x$	$\mathbb{R} \setminus \pi\mathbb{Z}$	$-\left(1 + \operatorname{cotan}^2 x\right) = -\frac{1}{\sin^2 x}$

FIGURE 4 – Tableau des fonctions dérivées des fonctions usuelles, à connaitre par coeur.

propage d'un point à un autre de façon à minimiser son temps de trajet. Ce principe de moindre action a ensuite été formalisé mathématiquement par Lagrange quelques années plus tard, qui introduit une fonction – le Lagrangien – qui dépend des variables physiques du système et grâce auquel on peut calculer une autre quantité – l'*action* – qui, lorsqu'on la minimise en calculant où la dérivée s'annule, nous donne la solution physique du changement du système, c'est à dire le changement de moindre action. Le calcul des dérivées est donc très important dans tous les domaines appliqués.

2.4 Propriétés des fonctions dérivables

Définition 2.81. *On dit que f est dérivable sur D si f est dérivable en tout point a de D. Dans ce cas, la fonction :*

$$\begin{aligned} f' : D &\longrightarrow \mathbb{R} \\ x &\longmapsto f'(x) \end{aligned}$$

s'appelle la fonction dérivée *de f (elle n'est pas forcément continue).*

Proposition 2.82. *Soit $f, g : D \to \mathbb{R}$, et $h : D' \to \mathbb{R}$ telle que $h(D') \subset D$. On rappelle les formules suivantes pour la dérivée :*

$$(f+g)' = f' + g', \quad (f \times g)' = f' \times g + f \times g', \quad (f \circ h)' = h' \times (f' \circ h)$$

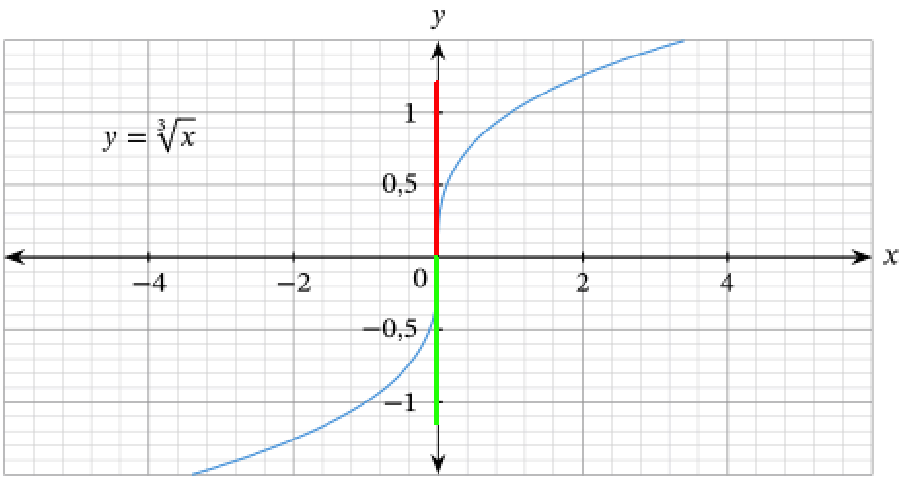
En particulier, on déduit que $\left(\frac{f}{g}\right)' = \frac{f' \times g - f \times g'}{g^2}$ et que, si f admet une fonction réciproque f^{-1}, alors sa dérivée est $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$.

Démonstration. Pour la fonction réciproque, nous avons que pour tout $x \in D$, $f \circ f^{-1}(x) = x$, c'est à dire en dérivant à gauche et à droite : $(f \circ f^{-1})'(x) = 1$. Par la formule de la dérivation de la composée, on a donc $(f^{-1})'(x) \times f'(f^{-1}(x)) = 1$, ce qui donne le résultat. $\square$

Exemple 2.83. Soit $n \in \mathbb{N}^*$. On rappel que la fonction racine n-ième est la fonction réciproque de la fonction puissance n-ième $x \mapsto x^n$. Si n est paire, la bijection se fait entre $[0, +\infty[$ et lui-même, tandis que si n est impair, la bijection est sur $\mathbb{R}$ entier. Dans tous les cas, calculons la dérivée à droite de la fonction racine n-ième, quel que soit n. Notons $f : x \mapsto x^n$ la fonction puissance, donc $f' : x \mapsto nx^{n-1}$, et $f^{-1} : x \mapsto x^{\frac{1}{n}}$ la fonction racine. D'après la formule de la dérivée de la réciproque, nous avons, pour tout $x > 0$:

$$(f^{-1})'(x) = \frac{1}{f' \circ f^{-1}(x)} = \frac{1}{nx^{\frac{n-1}{n}}} = \frac{1}{n} x^{\frac{1}{n}-1}$$

C'est bien la formule à laquelle on s'attend. Par contre on voit que dès que $n \geq 2$, la dérivée (à droite) tend vers $+\infty$ en 0. La demi-tangente est donc verticale ascendante. Cela se voit effectivement sur le graphe de la fonction racine n-ième (pair ou impair). La particularité du graphe de la fonction impaire est qu'il admet une demi-tangente à gauche en zéro : une demi-droite verticale descendante, donc le graphe admet en zéro une tangente verticale. L'image suivante illustre ce cas avec la racine cubique : la demi-tangente à droite est en rouge, la demi-tangente à gauche est en vert.



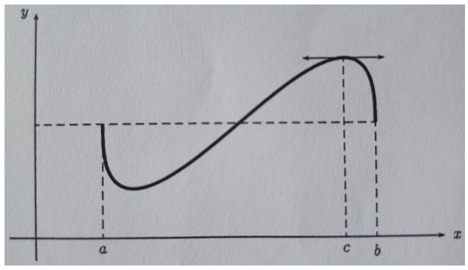
Exemple 2.84. Le même argument s'applique à la arctangente. Comme on sait que $\tan' = 1 + \tan^2$, on obtient que pour tout $x \in \mathbb{R}$, $\arctan'(x) = \frac{1}{1+\tan^2(\arctan(x))} = \frac{1}{1+x^2}$. Cette fois-ci en zéro, la tangente au graphe est la droite diagonale d'équation $y = x$.

Les deux principaux théorèmes sont les théorèmes de Rolle et celui des accroissements finis (le premier est un cas particulier du second). Le mathématicien Rolle a vécu au XVIIème siècle (l'analyse moderne n'était donc absolument pas encore bien définie) et il a énoncé le résultat suivant : entre deux racines d'un polynôme P, il y a une racine de sa dérivée P'. C'est donc dans le cadre de l'algèbre que ce résultat a été utilisé pendant plus de 150 ans.

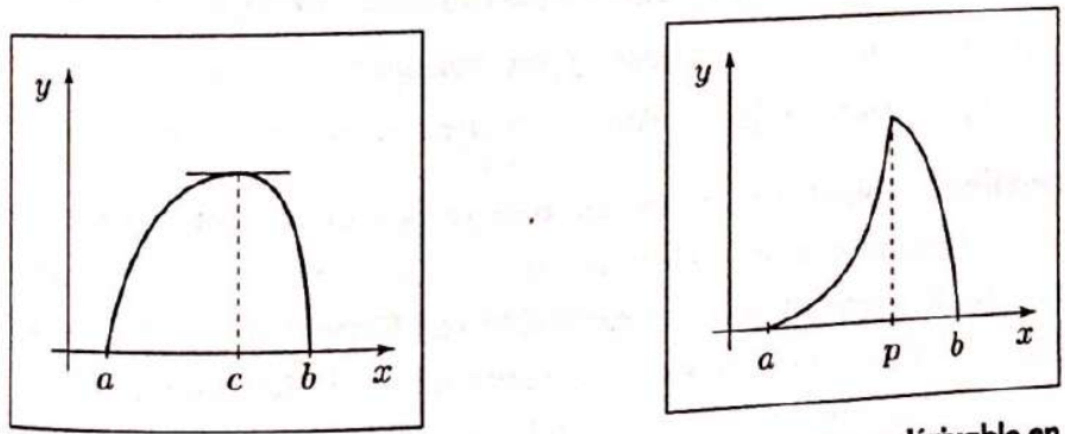
L'extension de ce résultat au champ de l'analyse et la reconnaissance du lien entre le cas particulier de Rolle et le théorème éponyme qui ne lui a été attribué que bien après est directement liée à l'évolution de la façon dont on démontre le théorème des accroissements finis. Joseph-Louis Lagrange et Augustin Louis Cauchy démontrent ce théorème en montrant l'inégalité des accroissements finis, puis en appliquant le théorème des valeurs intermédiaires à la dérivée pour obtenir une égalité. Ils supposent pour cela que la dérivée est continue. Un changement important a lieu en 1868 avec le Cours de calcul différentiel et intégral de Joseph-Alfred Serret. Sur une idée de Pierre-Ossian Bonnet, ce cours adopte la présentation moderne : ramener le théorème des accroissements finis au cas où la fonction s'annule aux bornes de l'intervalle et tirer parti du fait qu'une telle fonction admet un extremum à l'intérieur de l'intervalle où la dérivée s'annule. Cette démarche nouvelle ne manquera pas d'être remarquée, notamment par Gaston Darboux en 1874, qui s'intéresse alors aux fonctions dérivables à dérivée non continue. L'intérêt de la démonstration de Bonnet est en effet d'être à la fois plus simple que les démonstrations de Lagrange ou Cauchy, mais aussi plus générale, la continuité de la dérivée n'étant pas nécessaire. Cependant, dans l'ouvrage de Serret, le lien avec le théorème algébrique de Rolle n'est pas encore affirmé. Ce lien apparaîtra dans les ouvrages de mathématiques qui suivront, et peu d'années après, la démarche de Bonnet est définitivement adoptée, le nom de Rolle étant alors également attribué au lemme permettant de prouver l'égalité des accroissements finis. Le nom du Théorème est donc un hommage *a posteriori* accordé à Rolle.

Théorème 2.85. Théorème de Rolle. *Soit* $f : [a, b] \to \mathbb{R}$ *une fonction continue, dérivable sur* $]a, b[$*, telle que* $f(a) = f(b)$*. Alors il existe* $c \in]a, b[$ *tel que* $f'(c) = 0$.

Démonstration. La fonction f est continue sur un segment donc par la Proposition 2.42, elle admet au moins un maximum et un minimum. Si l'un se trouve en a et l'autre en b, alors la fonction est constante (car $f(a) = f(b)$) donc sa dérivée est nulle partout donc n'importe quel choix de point dans $]a, b[$ convient. Sinon, le maximum ou le minimum se trouve dans l'intervalle ouvert. Disons par exemple que cet extremum de la fonction se trouve en $c \in]a, b[$. Alors, $f'(c) = 0$. □



Remarque 2.86. Attention il est nécessaire de supposer que la fonction est dérivable partout, car on peut se retrouver dans des situations bizarres où la fonction a un extremum mais que la dérivée en ce point n'existe pas (donc n'est pas nulle) :



les hypothèses sont vérifiées

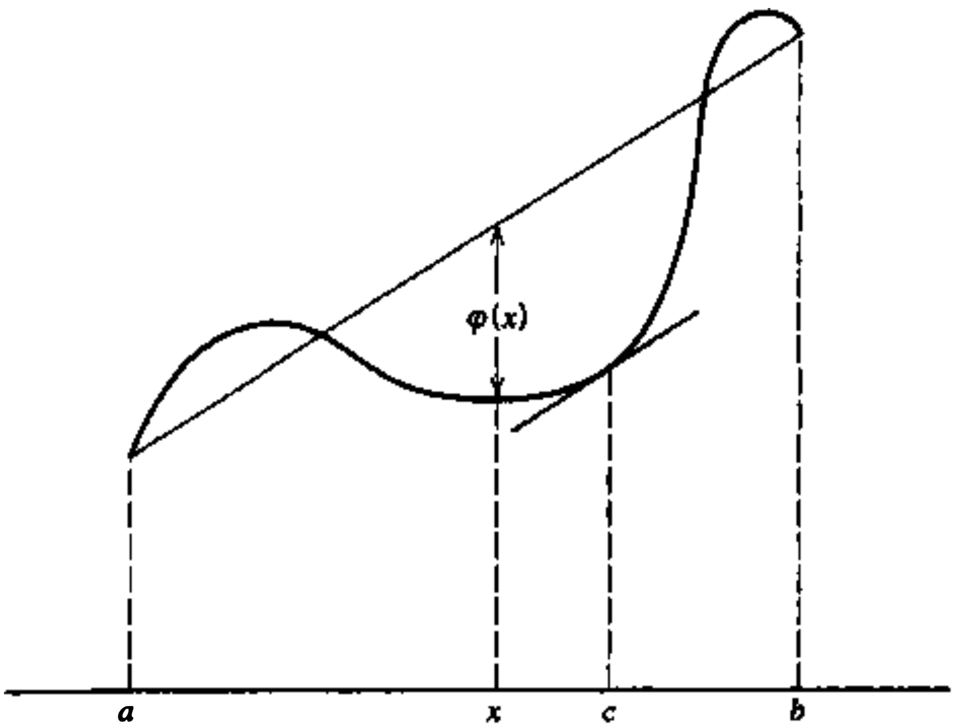
la fonction n'est pas dérivable en p, il n'y a pas de point à tangente horizontale

Le théorème de Rolle est un cas particulier du théorème suivant, plus général. Le théorème dit qu'il existe un point du graphe de f dont la tangente a la même pente que la droite passant par $M_a = (a, f(a))$ et $M_b = (b, f(b))$. Les conséquences du théorème sont importantes car elles nous donnent le sens de variation de n'importe quelle fonction dérivable.

Théorème 2.87. Théorème des accroissements finis. *Soit* $f : [a, b] \to \mathbb{R}$ *continue et dérivable sur* $]a, b[$*. Alors il existe* $c \in]a, b[$ *tel que :*

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

Démonstration. On pose $\varphi(x) = f(x) - \left[f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right]$. C'est une fonction continue sur $[a,b]$, dérivable sur $]a,b[$ et telle que $\varphi(a) = \varphi(b) = 0$. Donc par Rolle, il existe $c \in]a,b[$ telle que $\varphi'(c) = 0$. C'est à dire, le résultat demandé. $\square$



Corollaire 2.88. *Soit $f : [a,b] \to \mathbb{R}$ continue et dérivable sur $]a,b[$, alors on a :*

- *f est constante ssi $f' = 0$;*
- *f est croissante ssi $f' \geq 0$;*
- *f est décroissante ssi $f' \leq 0$.*

Démonstration. Dans les trois cas, le sens direct est facile, et découle de passage à la limite des taux d'accroissements de signe constant. Montrons l'implication "$f' \geq 0$ implique f croissante". Soit $u < v$ deux éléments de $[a,b]$. Donc f est continue sur $[u,v]$ et dérivable sur $]u,v[$. Donc par le théorème des accroissements finis, il existe $w \in]u,v[$ tel que :

$$f(w) = \frac{f(v) - f(u)}{v - u} \text{ c'est à dire } f(v) - f(u) = f'(w)(v-u)$$

Le membre de droite est positif ou nul car $f' \geq 0$ et $v - u > 0$, donc le membre de gauche l'est aussi, c'est à dire $f(v) \geq f(u)$. Ceci étant vrai pour tous points $u < v$ dans le segment $[a,b]$, la fonction f est croissante sur $[a,b]$. Pour l'implication "$f' \leq 0$ implique f décroissante", on prend la même preuve pour $-f$. $\square$

Remarque 2.89. L'équivalence ne fonctionne pas pour l'inégalité strictement croissante. Seule les implications "$f' > 0$ implique f est strictement croissante" et "$f' < 0$ implique f est strictement décroissante" sont valides (cela se voit dans la preuve). Par contre, on peut avoir une fonction strictement croissante et pourtant la dérivée f' s'annule en un point : par exemple $f : x \mapsto x^3$. La condition nécessaire et suffisante valide en toute généralité est donc :

$$f \text{ est strictement croissante sur } [a,b] \text{ ssi } f' \geq 0 \text{ et } f' > 0 \text{ sur une partie dense de } \mathbb{R}$$

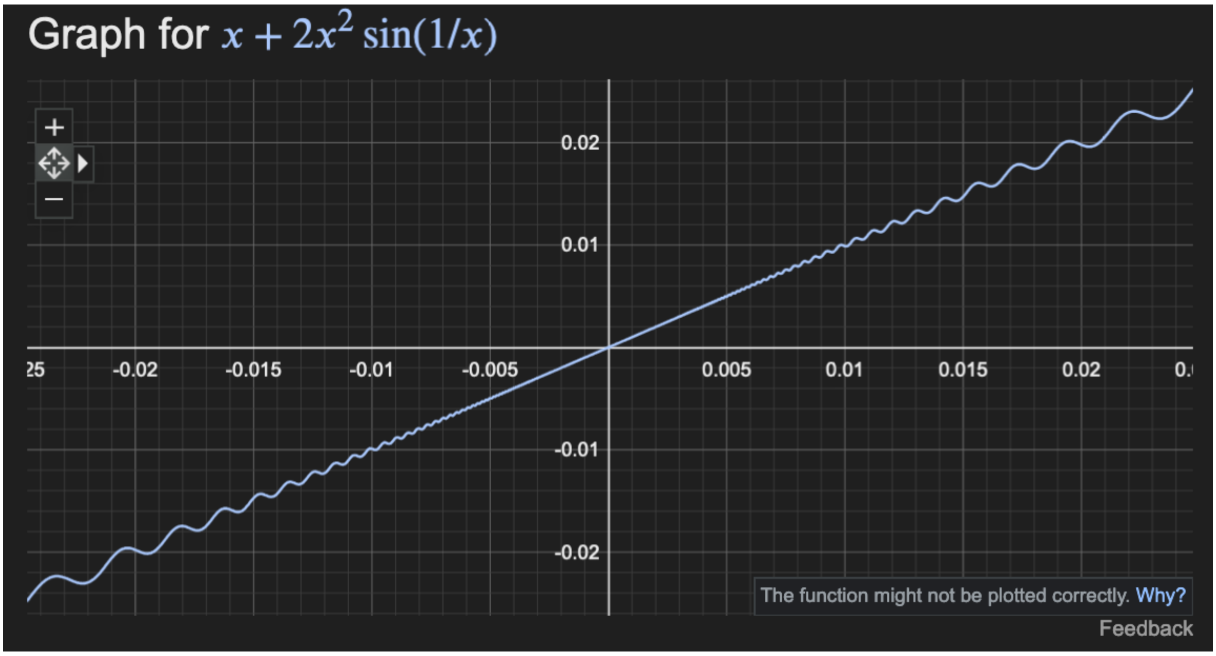
Remarque 2.90. Enfin, il est faux de croire que "$f'(a) > 0$ implique que f' est strictement positive sur un intervalle contenant a". C'est le cas bien sûr si la fonction dérivée f' est continue en a, et ce n'est pas le cas si f' est discontinue en a. Par exemple prenons la fonction

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = \begin{cases} 0 & \text{si } x = 0 \\ x + 2x^2 \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \end{cases} \end{aligned}$$

qui admet pour fonction dérivée :

$$\begin{aligned} f' : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = \begin{cases} 1 & \text{si } x = 0 \\ 1 + 4x \sin\left(\frac{1}{x}\right) - 2\cos\left(\frac{1}{x}\right) & \text{si } x \neq 0 \end{cases} \end{aligned}$$

Le calcul de la dérivée en $x = 0$ se fait avec la limite du taux d'accroissement (la définition classique), tandis que pour $x \neq 0$, on peut dériver classiquement somme, produit et composée de fonctions. On voit que $f'(0) = 1$, mais que f' n'est de signe constant sur aucun voisinage de 0 car le terme $4x\sin\left(\frac{1}{x}\right)$ tend vers 0 mais pas le dernier terme qui oscille entre -2 et $+2$.



On fini cette section par deux résultats intéressants de culture générale. Le premier est la caractérisation des fonctions dérivées. On appelle *fonction de Darboux* toute fonction réelle qui a la propriété des valeurs intermédiaires. Les fonctions continues sont des fonctions de Darboux (c'est le Théorème des valeurs intermédiaires), mais il existe des fonctions de Darboux non-continues (par exemple, la fonction $x \mapsto \sin\left(\frac{1}{x}\right)$). Darboux a montré que les fonctions dérivées possèdent la propriété des valeurs intermédiaires, c'est à dire que les fonctions dérivées sont des fonctions de Darboux. Ce théorème est important de par son corollaire.

Théorème de Darboux. *Soit $f : [a, b] \to \mathbb{R}$ une fonction dérivable, alors f' possède la propriété des valeurs intermédiaires.*

Corollaire 2.91. *Soit $f : [a, b] \to \mathbb{R}$ une fonction. Si elle ne possède pas la propriété des valeurs intermédiaires, alors elle n'est pas intégrable.*

La deuxième résultat est un résultat sur les suites, définies par récurrence à l'aide d'une fonction.

Théorème 2.92. ***Théorème du point fixe.*** *Soit $f : D \to \mathbb{R}$ une fonction dérivable. Supposons qu'il existe un point fixe $\ell \in D$ de f, c'est à dire tel que $f(\ell) = \ell$, et qu'il existe un intervalle $I =]\ell - \epsilon, \ell + \epsilon[$ et un réel $0 < \lambda < 1$ tel que pour tout $x \in I$, $|f'(x)| \leq \lambda$. Alors toute suite $(u_n)_n$ définie par $u_0 \in I$ et la relation de récurrence $u_{n+1} = f(u_n)$ converge vers ℓ.*

Démonstration. Montrons par récurrence que $u_n \in I$ pour tout n. C'est vrai pour $n = 0$. Supposons que cela soit vrai jusqu'au rang n. Supposons que $u_n < \ell$ (puisque sinon la preuve est finie), et donc comme I est un intervalle, $]u_n, \ell[\subset I$. Alors on peut appliquer le théorème des accroissements à $[u_n, \ell]$. Il existe $c \in]u_n, \ell[$ tel que :

$$f'(c) = \frac{f(\ell) - f(u_n)}{\ell - u_n} = \frac{u_{n+1} - \ell}{u_n - \ell}$$

En valeur absolue et en rappelant que $|f'(c)| \leq \lambda$, cela donne donc :

$$|u_{n+1} - \ell| \leq \lambda |u_n - \ell| < |u_n - \ell|$$

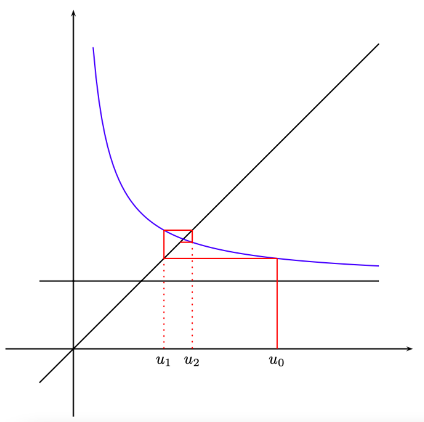
Donc la distance de u_{n+1} à ℓ est inférieure à la distance de u_n à ℓ, donc $u_{n+1} \in I$. La relation ci dessus tient donc pour tout $n \in \mathbb{N}$. On a donc :

$$|u_n - \ell| \leq \lambda^n |u_0 - \ell|$$

donc la suite $(|u_n - \ell|)_n$ tend vers 0, autrement dit u_n tend vers ℓ. □

Remarque 2.93. Si la dérivée de f est continue, on n'a pas besoin de supposer l'existence de I car elle est une conséquence de l'hypothèse $|f'| \leq \lambda < 1$.

Exemple 2.94. Prenons la fonction $f(x) = 1 + \frac{1}{x}$ et soit $\ell = \varphi = \frac{1+\sqrt{5}}{2}$ le nombre d'or, c'est à dire qu'il satisfait $\varphi^2 = \varphi + 1$, autrement dit $\varphi = 1 + \frac{1}{\varphi}$, c'est un point fixe de la fonction. On a $f'(x) = -\frac{1}{x^2}$ donc la valeur absolue de la dérivée est décroissante et pour $x = \frac{\sqrt{5}}{2}$, on a $|f'(x)| = 4/5 < 1$. Donc pour tout $x \in]\varphi - \frac{1}{2}, \varphi + \frac{1}{2}[$, on a $|f'(x)| \leq 4/5 < 1$. Donc par le théorème du point fixe, toute suite qui commence dans cet intervalle et définie par récurrence grâce à $u_{n+1} = A + \frac{1}{u_n}$ converge vers $\ell = \varphi$.



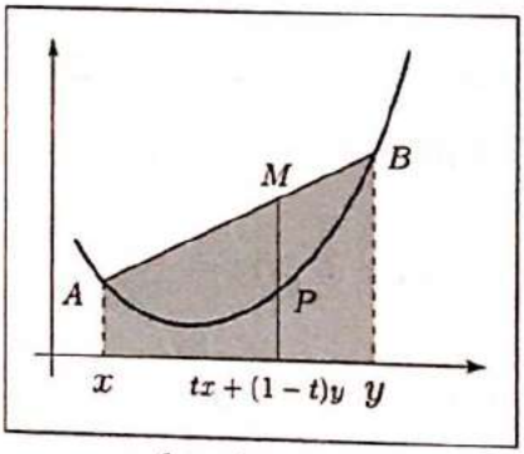
2.5 Convexité et dérivées d'ordre supérieur

Dans la suite on va prendre $D = I$ un intervalle de $\mathbb{R}$.

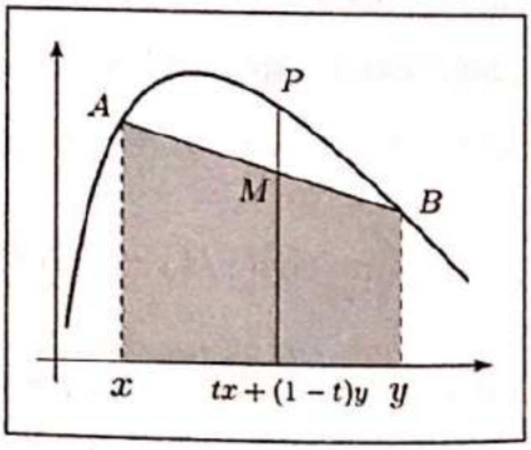
Définition 2.95. *Soit $f : I \to \mathbb{R}$ une fonction. On dit que f est* convexe *si quels que soient les nombres $a < b \in I$, on a :*

$$f(ta + (1-t)b) \leq tf(a) + (1-t)f(b) \qquad \textit{pour tout } t \in [0, 1].$$

On dit que f est concave *si la fonction $-f$ est convexe.*



une fonction convexe



une fonction concave

Donnons l'interprétation géométrique de cette définition. Soit $f : I \to \mathbb{R}$ une fonction, et soit $a < b \in I$. On pose $A = (a, f(a))$ et $B = (b, f(b))$ les points du graphe de f d'abscisse a et b, respectivement. Lorsque la variable t parcourt le segment $[0, 1]$, le point $x_t = ta + (1-t)b$ parcourt le segment $[a, b]$ depuis b (pour $t = 0$) vers a (atteint en $t = 1$). La droite passant par A et B a pour équation

$$y = \frac{x-a}{b-a}(f(b) - f(a)) + f(a) = \frac{f(b)-f(a)}{b-a}x + \frac{bf(a) - af(b)}{b-a}$$

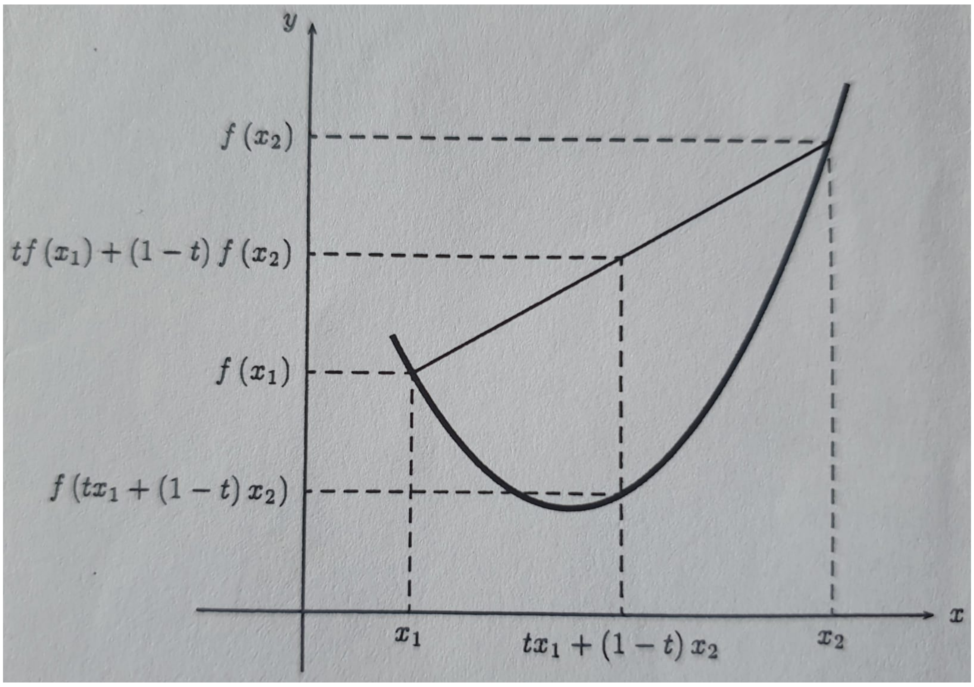
Soit $t \in [0, 1]$, et $x_t = ta + (1-t)b$. Le point de la droite d'abscisse x_t a pour ordonnée :

$$\begin{aligned} y_t &= \frac{ta + (1-t)b - a}{b-a}(f(b) - f(a)) + f(a) \\ &= \frac{(1-t)(b-a)}{b-a}(f(b) - f(a)) + f(a) \\ &= tf(a) + (1-t)f(b) \end{aligned}$$

L'ensemble des points de la droite qui ont pour abscisse x_t et ordonnée y_t est le segment $[A, B]$:

$$[A, B] = \Big\{ M_t = \big(ta + (1-t)b, tf(a) + (1-t)f(b)\big),\ t \in [0, 1] \Big\}$$

Maintenant soit P_t le point du graphe de f d'abscisse $x_t = ta + (1-t)b$ (la même que M_t). Le point a logiquement pour ordonnée $f(ta + (1-t)b)$. La condition de convexité, nous dit donc que le point P_t est *en dessous* du point M_t, pour tout $t \in [0, 1]$.



Une fonction $f : I \to \mathbb{R}$ est donc convexe si et seulement si, pour tous points distincts A et B du graphe de f, le graphe de f entre A et B est *en dessous* du segment $[A, B]$. Pour une fonction concave c'est l'inverse : le graphe est *au dessus* du segment $[A, B]$. On utilise le mot convexe car une fonction est convexe si et seulement si la partie du plan qui se situe au dessus du graphe de f est *convexe* dans le sens topologique. En topologie, une partie d'un espace est convexe si tout segment entre deux points de cette partie est contenu dans cette partie.

Exemple 2.96. Prenons la fonction polynomiale $f : x \mapsto x^2 + 1$ définie sur $\mathbb{R}$. Prenons $a < b$ et définissons $A = (a, f(a))$ et $B = (b, f(b))$. Nous avons donc, pour tout $t \in [0,1]$:

$$f\big(ta + (1-t)b\big) = \big(ta + (1-t)b\big)^2 + 1 = t^2a^2 + (1-t)^2b^2 + 2t(1-t)ab + 1$$

Ajoutons $ta^2 - ta^2 = 0$ et $(1-t)b^2 - (1-t)b^2 = 0$ au membre de droite. En notant que $1 = t + (1-t)$, on obtient alors :

$$f\big(ta + (1-t)b\big) = tf(a) + (1-t)f(b) - t(1-t)(a-b)^2$$

Comme $t \in [0,1]$, $(1-t) \in [0,1]$, et comme $(a-b)^2 > 0$ le dernier terme $t(1-t)(a-b)^2$ est positif. Comme il est précédé d'un signe $-$, on a bien la condition de convexité comme on pouvait le voir sur un dessin.

Exemple 2.97. En France, l'impôt sur le revenu est une fonction convexe du revenu (dessin).

Les propositions suivantes caractérisent une fonction convexe, selon qu'elle est ou non dérivable, ou non doublement dérivable.

Lemme 2.98. *Une fonction $f : I \to \mathbb{R}$ est convexe si et seulement si pour tous $a < b$ éléments de I, on a :*

$$\forall\, x \in]a, b[\qquad \frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x}$$

Démonstration. Soit $a < b$ deux tels nombres. Ce sont les abscisses des points A et B. Quand t parcourt le segment $[0,1]$, le point $x_t = ta + (1-t)b$ parcourt le segment $[a, b]$ (de b à a). On a

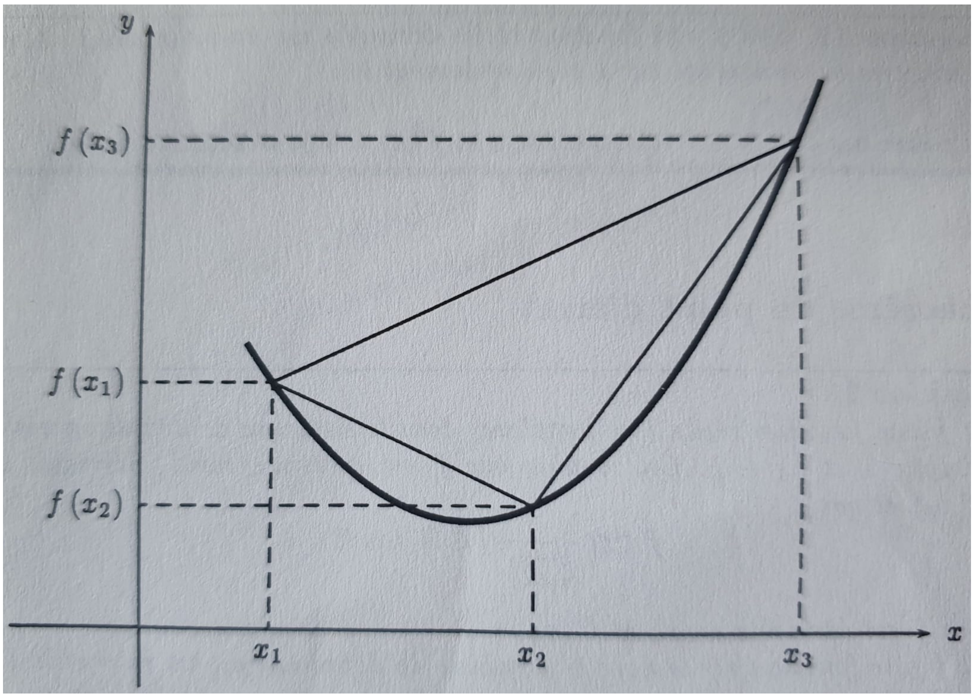


FIGURE 5 – Pour tout $x_1 < x_2 < x_3$ on a $\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_1)}{x_3-x_1} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2}$.

alors $x_t - a = (1-t)(b-a)$, ce qui nous donne que l'inégalité de gauche de la proposition peut se récrire :

$$\forall t \in [0,1] \qquad f(x_t) \leq f(a) + (x_t - a)\frac{f(b)-f(a)}{b-a} = tf(a) + (1-t)f(b)$$

tandis que l'inégalité de droite se récrit :

$$\forall t \in [0,1] \qquad f(x_t) \leq f(b) + (x_t - b)\frac{f(b)-f(a)}{b-a} = tf(a) + (1-t)f(b)$$

Les deux inégalités sont la condition de convexité de la fonction f. □

Remarque 2.99. On ne le montrera pas mais le lemme dit de manière équivalente que f est convexe si et seulement si pour tout $a \in I$, la fonction $x \mapsto \frac{f(x)-f(a)}{x-a}$ est croissante sur I.

Corollaire 2.100. *Soit I un intervalle ouvert. Si $f : I \to \mathbb{R}$ est convexe sur I alors elle est dérivable à gauche et à droite en tout point de I, donc continue sur I.*

Démonstration. Comme I est un intervalle ouvert, pour tout $a \in I$, il existe x et y tel que $x < a < y$. On a alors, d'après la remarque précédente :

$$\frac{f(x)-f(a)}{x-a} \leq \frac{f(y)-f(a)}{y-a}$$

En outre, pour $x < a$, le taux d'accroissement de gauche est une fonction croissante de x, majorée par le membre de droite, donc admet une limite en a, qui n'est autre que la dérivée à gauche de f en a. D'autre part, le taux d'accroissement de droite est une fonction croissante de $y > a$. Il diminue donc lorsque y tend vers a. Comme il est minoré par le membre de gauche, il admet une limite en a qui est la dérivée à droite de f en a. On a donc :

$$f'_g(a) \leq f'_d(a)$$

Etant dérivable à gauche et à droite, f est continue à gauche et à droite, donc continue. □

Proposition 2.101. *Soit $f : I \to \mathbb{R}$ une fonction convexe (resp. concave). Si a est un point critique de f alors c'est un minimum (resp. maximum) local de f.*

Soit I un intervalle ouvert et soit $f : I \to \mathbb{R}$ une fonction deux fois dérivable. Cela veut dire que la fonction dérivée f' définie par :

$$\begin{aligned} f' : \ I & \longrightarrow \mathbb{R} \\ x & \longmapsto f'(x) \end{aligned}$$

est elle-même dérivable. On note f'' sa fonction dérivée et on l'appelle la *dérivée seconde* de f. En particulier, cela veut dire que f' est continue (mais pas forcément f'').

Proposition 2.102. *Soit $f : I \to \mathbb{R}$ une fonction deux fois dérivable. f est convexe ssi $f'' \geq 0$.*

Démonstration. Comme f est dérivable sur I, elle est continue sur I. Montrons que si f' est croissant alors f est convexe. On pose $g(x) = f(x) - \left[f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right]$. C'est une fonction continue sur $[a, b]$, dérivable sur $]a, b[$ et telle que $g(a) = g(b) = 0$. Donc par le Théorème de Rolle, il existe $c \in]a, b[$ telle que $g'(c) = 0$. Or la fonction g' est croissante, puisqu'on a $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$. Cela veut donc dire que pour tout $x \in]a, c]$, $g'(x) \leq 0$, et pour tout $x \in [c, b[$, $g'(x) \geq 0$. Ce qui veut dire que g est décroissante sur $[a, c]$ et croissante sur $[c, b]$. Comme $g(a) = g(b) = 0$, on déduit que $g(x) \leq 0$ pour tout $x \in [a, b]$. Autrement dit, pour tout $x \in [a, b]$

$$f(x) \leq f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \qquad \Longleftrightarrow \qquad \frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a}$$

ce qui montre que f est convexe.

Inversement, montrons que si f est convexe, alors f' est croissante. Soit $a < b \in I$, on rappelle que nous avons deux fonctions croissantes :

$$x \longmapsto \frac{f(x) - f(a)}{x - a} \qquad \text{et} \qquad y \longmapsto \frac{f(y) - f(b)}{y - a}$$

Donc, pour tout $x, y \in]a, b[$, on a :

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a} = \frac{f(a) - f(b)}{a - b} \leq \frac{f(y) - f(b)}{y - b}$$

Quand x tend vers a par la droite, le membre de gauche tend vers $f'_d(a) = f'(a)$ car f est dérivable, tandis que quand y tend vers b par la gauche, le membre de droite tend vers $f'_g(b) = f'(b)$. Donc on a $f'(a) \leq f'(b)$ pour tous $a < b$ de I donc f' est croissante. Pour le deuxième énoncé, si $f'' \geq 0$ alors f' est croissante. $\square$

Remarque 2.103. Une version affaiblie de cette proposition dit que si la fonction $f : I \to \mathbb{R}$ est dérivable, alors f est convexe ssi f' est croissante. Dans ce cadre, nous pouvons montrer que pour tout x et a dans l'intervalle I, on a :

$$f(x) \geq f(a) + f'(a)(x - a) \tag{2.3}$$

Le membre de droite est l'équation de la tangente au graphe de f au point $(a, f(a))$. Ainsi, si f' est croissante (et donc convexe), cela veut dire que le graphe de f est au dessus de chacune de ses tangentes. C'est une autre caractérisation géométrique de la convexité.

Exemple 2.104. Puisque la fonction exponentielle est sa propre dérivée (et est strictement positive), alors la fonction exponentielle est convexe. D'autre part en appliquant l'équation (2.3) à $a = 0$ on obtient :

$$e^x \geq 1 + x \qquad \text{pour tout nombre réel } x.$$

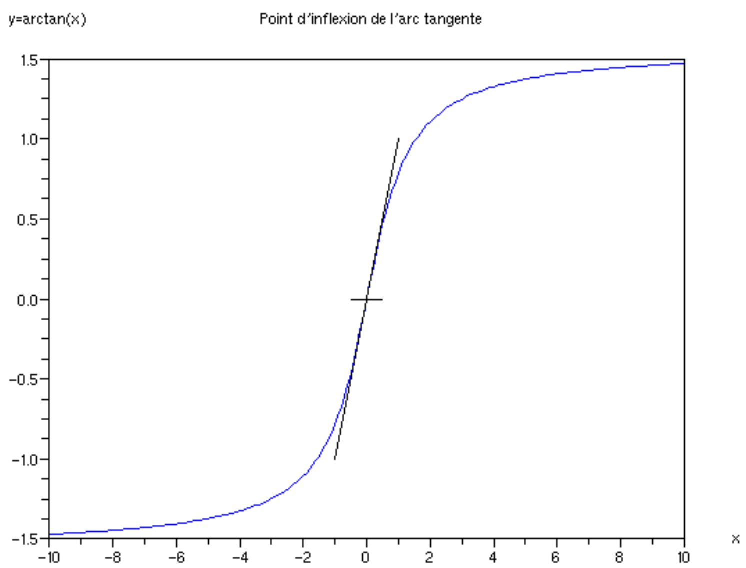
Exemple 2.105. Pour tout $x > 0$, on a que la dérivée seconde du logarithme est $\ln''(x) = -\frac{1}{x^2} < 0$ donc la fonction logarithme est concave (car $f = -\ln$ est convexe). La formule (2.3) évaluée en $a = 1$ donne alors

$$-\ln(x) \geq 0 - (x - 1) \qquad \Longleftrightarrow \qquad \ln(x) \leq x - 1$$

Remarque 2.106. Avec la caractérisation de la convexité/concavité par rapport à la dérivée seconde, on peut montrer que f est convexe ssi f^{-1} est concave. En effet on exprime $(f^{-1})''$ en fonction de f et f' par : $(f^{-1})'' = -\frac{f''\circ f^{-1}}{(f'\circ f^{-1})^3}$.

Définition 2.107. *Soit $f : I \to \mathbb{R}$ une fonction. On appelle* point d'inflexion *tout point $a \in I$ où f change de convexité. Plus précisément il existe $\epsilon > 0$ tel que f est concave (resp. convexe) sur $[a - \epsilon, a]$ et convexe (resp. concave) sur $[a, a + \epsilon]$.*

Exemple 2.108. La fonction $f : x \mapsto x^3$, de dérivée seconde $f''(x) = 2x$, admet un point d'inflexion en 0. La courbe passe de concave à convexe. La fonction arctangente admet comme dérivée première $\arctan'(x) = \frac{1}{1+x^2}$ et comme dérivée seconde $\arctan''(x) = -\frac{2x}{(1+x^2)}$. Cette dernière change de signe en $x = 1$. La courbe passe de convexe à concave.



Proposition 2.109. *Soit $f : I \to \mathbb{R}$ une fonction dérivable en un point $a \in I$. Le point a est un point d'inflexion de f si l'une des deux conditions équivalentes sont satisfaites :*

1. *le graphe de f traverse sa tangente au point $(a, f(a))$;*
2. *la fonction $h \mapsto f(a + h) - f(a) - f'(a)h$ s'annule en changeant de signe en 0.*

Si f est deux fois dérivable, a est un point d'inflexion si f'' s'annule et change de signe en a.

Maintenant, approfondissons l'étude des fonctions plusieurs fois dérivables. Soit $f : D \to \mathbb{R}$ une fonction dérivable sur le domaine D (qu'on suppose ouvert, et très souvent un intervalle ouvert ou une union d'intervalles ouverts). La fonction dérivée $f' : D \to \mathbb{R}$ est donnée par :

$$\begin{aligned} f' :\ D &\longrightarrow \mathbb{R} \\ x &\longmapsto f'(x) \end{aligned}$$

On suppose que la fonction dérivée est elle aussi dérivable, en tout point de D. Cela définit la dérivée seconde de la fonction f :

$$\begin{aligned} f'' : D &\longrightarrow \mathbb{R} \\ x &\longmapsto (f')'(x) \end{aligned}$$

Si jamais la fonction dérivée seconde est dérivable sur D, on peut continuer le processus en la dérivant, et cela définit la dérivée troisième f''', aussi notée $f^{(3)}$, et ainsi de suite.

Définition 2.110. *Soit $f : D \to \mathbb{R}$ une fonction définie sur un domaine ouvert D. Pour tout $n \in \mathbb{N}^*$ on définit, si elle existe, la* dérivée n-ième de f *– notée $f^{(n)}$ – par la récurrence suivante :*

$$f^{(0)} = f \quad \text{et} \quad f^{(k)} = (f^{(k-1)})' \quad \text{pour tout } 1 \le k \le n.$$

En particulier, $f^{(1)} = f'$ et $f^{(2)} = f''$. Dans ce cas on dit que f est n fois dérivable*.*

Remarque 2.111. La définition implique que si la fonction n'est pas dérivable à un certain rang N – c'est à dire si $f^{(N-1)}$ est bien définie mais n'est pas dérivable – alors $f^{(m)}$ n'existe pas, pour tout $m \ge N$. D'autre part, il est direct de voir qu'une fonction dérivable n fois a toutes ses dérivées k-ièmes continues, pour $0 \le k \le n-1$. En effet, si $f^{(k+1)}$ existe, cela veut dire que $f^{(k)}$ est dérivable. Cela veut dire que $f^{(k)}$ est continue.

Proposition 2.112. *La dérivée n-ième d'un produit de deux fonctions f, g au moins n fois dérivables est :*

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$$

où on rappelle la valeur du coefficient binomial $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Remarque 2.113. La preuve utilise l'astuce suivante : si on a une situation de la forme $\sum_{k=0}^{n} A_{k+1} + B_k$, il faut séparer en deux $\sum_{k=0}^{n} A_{k+1} + \sum_{k=0}^{n} B_k$, et réindexer la première somme pour avoir :

$$\sum_{k=1}^{n+1} A_k + \sum_{k=0}^{n} B_k = A_{n+1} + B_0 + \sum_{k=1}^{n} A_k + B_k$$

Démonstration. On prouve notre formule par récurrence. Pour $n = 0$, on a $(fg)^{(0)} = fg$. Pour $n = 1$ aussi c'est vrai puisque $(fg)' = f'g + fg'$ et les coefficients binomiaux sont $\binom{1}{0} = \binom{1}{1} = 1$. Maintenant supposons que la formule est vraie pour tout $0 \le k \le n$ et montrons là au rang $n+1$. Supposons que f et g sont dérivables au rang $n+1$. Prenons la formule au rang n et dérivons là une fois :

$$\begin{aligned} (fg)^{(n+1)} &= \sum_{k=0}^{n} \binom{n}{k} \left(f^{(k+1)} g^{(n-k)} + f^{(k)} g^{(n-k+1)} \right) \\ &= \sum_{k=1}^{n+1} \binom{n}{k-1} f^{(k)} g^{(n-(k-1))} + \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n+1-k)} \\ &= \sum_{k=0}^{n+1} \left(\binom{n}{k-1} + \binom{n}{k} \right) f^{(k)} g^{(n+1-k)} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)} g^{(n+1-k)} \end{aligned}$$

où on a utilisé la formule $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$ en supposant que $\binom{n}{-1} = \binom{n}{n+1} = 0$. □

Remarque 2.114. C'est la même preuve que celle du binôme de Newton ! Où l'on a :

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

On peut retrouver la formule de Newton à partir de celle de Leibniz pour la dérivée d'un produit de deux fonctions f et g en prenant $f(x) = e^{ax}$ et $g(x) = e^{bx}$ et en dérivant leur produit $e^{(a+b)x}$ n fois.

Dans la suite on se restreint au cas où $D = I$ est un intervalle ouvert. Rappelons la formule des accroissements finis ; si $f : I \to \mathbb{R}$ est dérivable et si $a < b \in I$ alors il existe $c \in]a, b[$ tel que :

$$f'(c) = \frac{f(b) - f(a)}{b - a} \quad \text{c'est à dire} \quad f(b) = f(a) + f'(c)(b - a)$$

Cette formule est en fait un cas particulier d'une formule plus générale pour les fonctions dérivables n fois.

Proposition 2.115. Formule de Taylor-Lagrange. *Si la fonction $f : [a, b] \to \mathbb{R}$ est dérivable n fois sur $]a, b[$ et telle que $f, f', f^{(2)}, \ldots, f^{(n-1)}$ sont continues sur $[a, b]$, alors il existe $c \in]a, b[$ tel que :*

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(a)}{2!}(b-a)^2 + \frac{f^{(3)}(a)}{3!}(b-a)^3 + \ldots + \frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} + \frac{f^{(n)}(c)}{n!}(b-a)^n$$

Démonstration. Si $n = 1$ on retrouve le théorème des accroissements finis. Sinon, $n \geq 2$, on définit la quantité A par la formule suivante :

$$A = \frac{n!}{(b-a)^n} \left[f(b) - f(a) - f'(a)(b-a) - \frac{f''(a)}{2!}(b-a)^2 - \frac{f^{(3)}(a)}{3!}(b-a)^3 - \ldots - \frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} \right]$$

Comme dans la démonstration du théorème des accroissements finis, on introduit une fonction auxiliaire définie pour tout $x \in [a, b[$:

$$\varphi(x) = f(b) - f(x) - f'(x)(b-x) - \frac{f''(x)}{2!}(b-x)^2 - \frac{f^{(3)}(x)}{3!}(b-x)^3 - \ldots - \frac{f^{(n-1)}(x)}{(n-1)!}(b-x)^{n-1} - \frac{(b-x)^n}{n!} A$$

Comme la fonction f est dérivable n fois, alors φ est dérivable une fois. La définition de A donne $\varphi(a) = 0$. D'autre part, on a directement que $\varphi(b) = 0$. On peut donc appliquer le théorème de Rolle et il existe $c \in]a, b[$ tel que $\varphi'(c) = 0$. Or la dérivée de φ sur l'intervalle $]a, b[$ vaut :

$$\varphi'(x) = \frac{(b-x)^{n-1}}{(n-1)!} \left(A - f^{(n)}(x) \right)$$

Comme $c \neq b$, l'égalité $\varphi'(c) = 0$ donne $f^{(n)}(c) = A$ d'où le résultat. □

Remarque 2.116. La formule de Taylor-Lagrange, qui est une généralisation à l'ordre n du théorème des accroissements finis, n'est en fait qu'une conséquence de celui-ci !

Exemple 2.117. Prenons comme fonction cosinus et fixons $x \in \mathbb{R}$. Soit $a = 0$ et $b = x > 0$, ce qui fait que $b - a = x$. On connait les dérivées successives du cosinus et on a donc $\cos^{(4p)}(t) = \cos(x), \cos^{(4p+1)}(t) = -\sin(t), \cos^{(4p+2)}(t) = -\cos(t), \cos^{(4p+3)}(t) = \sin(t)$, pour tout $t \in [0, x]$. Quand on évalue les dérivées en $a = 0$, seules les dérivées d'ordre pair survivent et on a le tableau suivant :

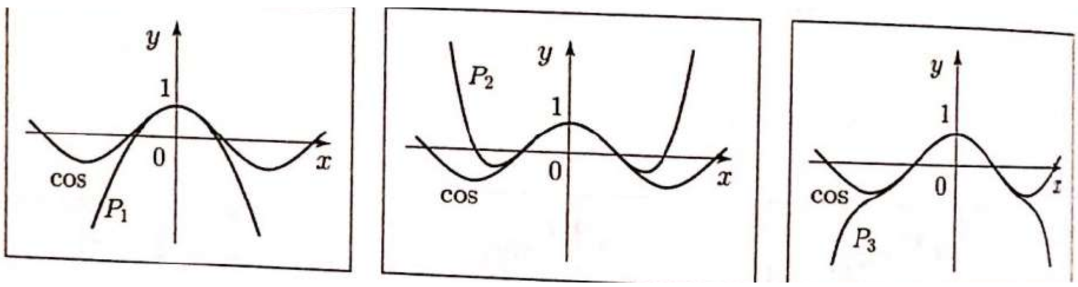
n	0	1	2	3	4	5	6
$\cos^{(n)}(0)$	1	0	-1	0	1	0	-1

D'après la Proposition 2.115, il existe un certain $0 < c < x$ tel que (par exemple, choix arbitraire) :

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} - \sin(c)\frac{x^7}{7!}$$

Comme la fonction est dérivable une infinité de fois et que ses dérivées n-ièmes sont continues partout, nous pouvons pousser le développement en puissance de x beaucoup plus loin. Par exemple, à n'importe quel ordre $n \in \mathbb{N}^*$, nous avons qu'il existe $c_n \in]0, x[$ tel que

$$\cos(x) = \sum_{k=0}^{E\left(\frac{n-1}{2}\right)} \frac{(-1)^k}{(2k)!} x^{2k} + \frac{\cos^{(n)}(c_n)}{n!} x^n$$



Exemple 2.118. Prenons la fonction exponentielle sur $\mathbb{R}$, c'est à dire $f(t) = e^t$ qu'on note aussi $\exp(t)$, et fixons $x > 0$ (le même raisonnement fonctionne pour $x < 0$). La fonction exponentielle est dérivable une infinité de fois sur $\mathbb{R}$. Comme les dérivées successives de l'exponentielle sont elle-même, lorsqu'elles sont évaluées en 0, elles donnent toujours la même valeur $\exp^{(k)}(0) = e^0 = 1$. D'autre part comme la fonction exponentielle est croissante, nous avons $1 \leq e^t \leq e^x$ pour tout $0 \leq t \leq x$, c'est à dire que $1 \leq \exp^{(n)}(t) \leq e^x$ pour tout $n \in \mathbb{N}$ et $0 \leq t \leq x$. Soit $a = 0$ et $b = x$, donc $b > a$. Comme l'exponentielle est dérivable une infinité de fois sur $\mathbb{R}$, les conditions de la Proposition 2.115 sont satisfaites sur le segment $[0, x]$. Il existe donc $c_n \in]a, b[$ tel que nous avons l'égalité suivante :

$$\left| e^x - \left(1 + x + \frac{x^2}{2} + \frac{x^3}{3!} \ldots + \frac{x^{n-1}}{(n-1)!}\right) \right| = e^{c_n} \frac{x^n}{n!}$$

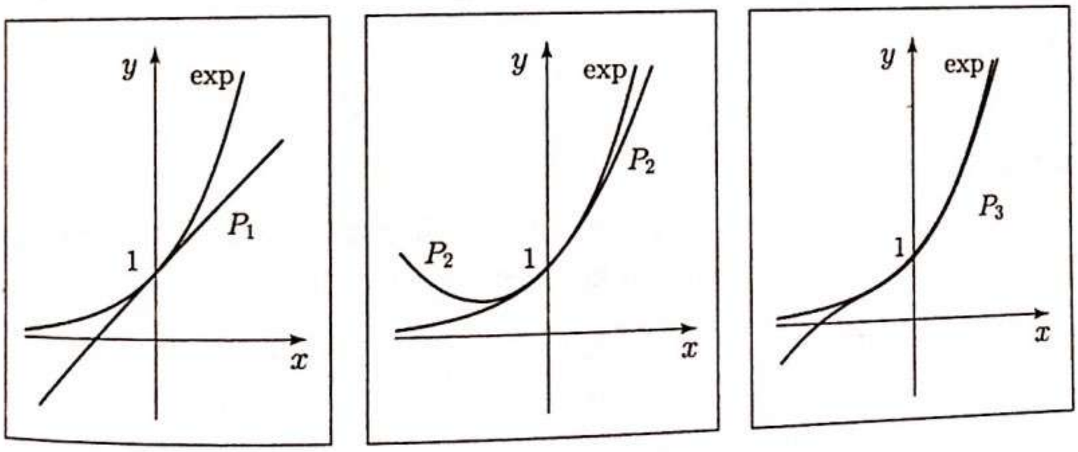
Mais comme $1 \leq e^{c_n} \leq e^x$ (on rappelle que x est fixé), on obtient la majoration suivante, qui ne dépend plus du nombre c_n :

$$\left| e^x - \left(1 + x + \frac{x^2}{2} + \frac{x^3}{3!} \ldots + \frac{x^{n-1}}{(n-1)!}\right) \right| \leq e^x \frac{x^n}{n!}$$

Posons $P_{n-1}(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} \ldots + \frac{x^{n-1}}{(n-1)!} = \sum_{k=0}^{n-1} \frac{x^k}{k!}$. Autrement dit nous avons l'inégalité suivante :

$$|e^x - P_{n-1}(x)| \leq e^x \frac{x^n}{n!} \tag{2.4}$$

On peut prendre un autre entier $n' \in \mathbb{N}^*$, et dans ce cas il existe un autre $c_{n'}$ dans $]0, x[$ tel que l'inégalité (2.4) est vraie pour n' : $|e^x - P_{n'-1}(x)| \leq e^x \frac{x^{n'}}{n'!}$. L'inégalité (2.4) est donc vraie pour tout $n \in \mathbb{N}^*$.



En admettant que la suite $\left(\frac{x^n}{n!}\right)_{n\in\mathbb{N}}$ converge vers 0, nous déduisons que le membre de droite de l'inégalité (2.4) tend vers 0 quand n tend vers l'infini (rappelons que x est fixé). Alors nous déduisons que $\lim\limits_{n\to\infty} e^x - P_n(x) = 0$. Nous déduisons que la suite $(P_n(x))_{n\in\mathbb{N}}$ tend vers le nombre réel e^x en l'infini. Maintenant, si $x < 0$, on effectue le même raisonnement sur $[x, 0]$ et on prend comme majorant des dérivées de l'exponentielle le nombre 1, car $e^x \leq e^t \leq 1$ pour tout $t \in [x, 0]$. Le reste du raisonnement est inchangé. Nous avons donc prouvé la proposition suivante :

Proposition 2.119. *Pour tout $x \in \mathbb{R}$, nous avons $\sum_{k=0}^{n} \frac{x^k}{k!} \underset{n\to+\infty}{\longrightarrow} e^x$.*

Remarque 2.120. On note donc, pour tout $x \in \mathbb{R}$, $e^x = \sum_{k=0}^{+\infty} \frac{x^k}{k!}$. En d'autres mots, on a approché la fonction exponentielle par un polynôme infini (une série entière).

Nous voyons que le développement de Taylor-Lagrange nous permet de remplacer une fonction au voisinage d'un point, en fonction d'un polynôme à l'ordre n, et de majorer la différence entre la fonction et le polynôme grâce à l'inégalité de Taylor. Cette propriété s'avère très puissante en analyse mathématique et en physique.