Chapitre 2 : Rappels sur la topologie de la droite réelle

La plus grande partie du contenu de ce chapitre a normalement déjà été vue pendant les années précédentes. Nous n'allons pas redonner toutes les preuves ni même toutes les définitions (par exemple pour les différentes limites des fonctions réelles). Notre but est :

- de faire des rappels pour se rafraîchir la mémoire. En particulier, le lecteur pourra en profiter pour se replonger dans ses cours précédents et revoir les points dont il n'est plus très sûr.
- d'énoncer quelques résultats fondamentaux des réels qui nous serviront pour des démonstrations des chapitres suivants.
- d'observer comment les différentes notions topologiques sont définies dans $\mathbb{R}$ pour pouvoir les généraliser à des espaces plus complexes dans la suite de ce cours.

1 Les structures de $\mathbb{R}$

1.1 Structure algébrique

Les opérations standards sur les réels sont l'addition et la multiplication ainsi que leurs inverses. Mathématiquement, on peut voir dans $\mathbb{R}$ plusieurs structures.

L'ensemble $\mathbb{R}$ muni de l'addition standard $+$ est un **groupe commutatif** :

- la somme de deux réels est un réel,
- le zéro est un élément neutre pour l'addition,
- pour tout $x \in \mathbb{R}$ il existe un inverse $-x$ pour l'addition,
- la somme est associative car $(a + b) + c = a + (b + c)$ et commutative dans le sens où $a + b = b + a$.

Si on y ajoute la multiplication standard $\times$, il s'agit d'un **corps commutatif** :

- $\mathbb{R}$ muni de $+$ est un groupe commutatif ayant 0 comme élément neutre,
- $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ muni de la multiplication est un groupe commutatif d'élément neutre 1,
- la multiplication est distributive par rapport à l'addition car $(a + b) \times c = a \times c + b \times c$.

Mais on peut le voir comme un cas très particulier d'**espace vectoriel réel** qui a juste une dimension :

- la somme de deux réels est réelle
- si $\lambda \in \mathbb{R}$ est un scalaire et $x \in \mathbb{R}$ un vecteur, alors $\lambda x \in \mathbb{R}$ est un vecteur.

Cette dernière structure peut paraître très artificielle mais nous allons étudier dans ce cours la topologie des espaces vectoriels réels. Il est donc intéressant de voir que $\mathbb{R}$ en est un cas particulier dont nous allons nous inspirer pour le cas général.

1.2 La notion de distance

Pour généraliser des concepts comme la continuité des fonctions à d'autres ensembles que les réels, il nous faut avoir une vision plus géométrique de ces définitions.

La **distance** entre deux réels x et y est $|x-y|$.

Un petit intervalle $]x-\varepsilon, x+\varepsilon[$ avec $\varepsilon > 0$ est un voisinage du point x. Il s'agit de tous les réels à distance au plus ε de x. C'est donc aussi la boule ouverte de centre x et de rayon ε. Plus généralement, on appelle **voisinage** de $x \in \mathbb{R}$ tout ensemble qui contient une boule $]x-\varepsilon, x+\varepsilon[$ avec $\varepsilon > 0$.

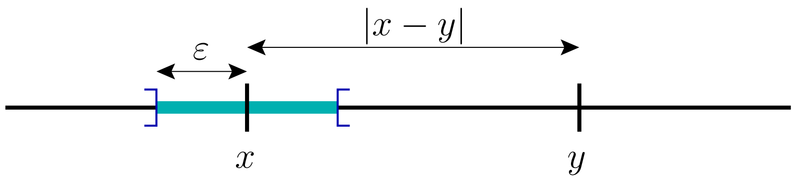


FIGURE 2.1 – *la distance $|x-y|$ entre deux nombres réels correspond bien à la notion usuelle. Un voisinage de x contient un petit intervalle autour de x.*

Il sera important de traduire toute phrase avec quantificateurs sous une forme plus géométrique et intuitive. Ainsi la convergence $u_n \longrightarrow \ell$ qui s'écrit

$$\forall \varepsilon > 0 \ , \ \exists n_0 \in \mathbb{N} \ , \ \forall n \geq n_0 \ , \ |u_n - \ell| < \varepsilon$$

peut se lire « *Quel que soit le petit voisinage de ℓ considéré, la suite (u_n) finit par entrer dans ce voisinage et y rester.* ». Vu comme cela, on comprend qu'on peut généraliser la notion de limite pour des suites dans des espaces plus généraux une fois qu'on a défini un concept de *voisinage* ou de *boule*.

Comme on l'a vu, la valeur absolue est une fonction primordiale pour faire de l'analyse sur $\mathbb{R}$. Voici quelques rappels concernant la valeur absolue :

- si $x \geq 0$, $|x| = x$ et si $x \leq 0$, $|x| = -x$. En particulier, $|x| = |-x|$.
- $|x| \leq a$ est équivalent à $-a \leq x \leq a$. En particulier, les boules sur $\mathbb{R}$ sont des intervalles car $\{y \in \mathbb{R}, |x-y| < \varepsilon\} =]x-\varepsilon\,, x+\varepsilon[$.
- $|x| \geq a$ est équivalent à $(x \leq -a$ ou $x \geq a)$.

- $|x \times y| = |x| \times |y|$

Une des propriétés fondamentales de la valeur absolue est la suivante.

Proposition 2.1

L'**inégalité triangulaire** s'énonce

$$|a+b| \leq |a| + |b|.$$

Il y a égalité si et seulement a et b ont même signe.

Pour $b = -c$, on obtient une majoration pour une différence $|a-c| \leq |a| + |c|$. C'est une inégalité triangulaire sur les distances si on l'écrit sous la forme

$$|x-y| = |x-z+z-y| \leq |x-z| + |y-z| \ .$$

On trouve bien qu'aller directement de x à y est toujours plus court que de passer par z. Il est aussi important de connaître une deuxième version de l'inégalité triangulaire.

Proposition 2.2

Pour tous réels a et b

$$\big||a| - |b|\big| \leq |a+b| \quad \text{et} \quad ||a| - |b|| \leq |a-b| \ .$$

Démonstration : On utilise la première version de l'inégalité triangulaire

$$|a| = |a+b-b| \leq |a+b| + |b|$$

et donc $|a+b| \geq |a| - |b|$. Mais les rôles de a et b sont symétriques, donc on trouve aussi $|a+b| \geq |b| - |a|$ (quitte à refaire l'argument en changeant les rôles de a et b). Pour la deuxième inégalité, il suffit de changer b en $-b$. $\square$

1.3 La structure d'ordre et les bornes supérieures

Nous avons l'habitude de munir $\mathbb{R}$ d'une structure d'ordre. Il est notable que cette structure est compatible avec les opérations naturelles sur $\mathbb{R}$. Par exemple, si $a > a'$ et $b > b'$, on a $a + b > a' + b'$. On a aussi les règles connues concernant la multiplication (attention aux changements de sens pour les négatifs). Ce n'est pas anodin puisqu'une telle structure d'ordre n'est pas toujours possible. Par exemple, on ne peut munir $\mathbb{C}$ d'une structure d'ordre qui se comporte raisonnablement bien par rapport aux opérations complexes (mais on peut munir $\mathbb{C}$ d'une structure d'ordre quand même, par exemple $a + ib > a' + ib'$ si $a > a'$ ou si $a = a'$ et $b > b'$).

On introduit les notions suivantes.

Définition 2.3

Soit $A \subset \mathbb{R}$ un sous-ensemble des nombres réels.
On dit que $M \in \mathbb{R}$ est **un majorant** de A si

$$\forall a \in A, \ , \quad a \leq M \ .$$

On dit que A est **majoré** s'il existe un réel M qui majore A.
On dit que $M \in \mathbb{R}$ est **la borne supérieure** de A, notée $\sup(A)$, si c'est le plus petit majorant de A, c'est-à-dire que c'est un majorant et que tout nombre plus petit n'est plus un majorant.
Si $M \in \mathbb{R}$ est la borne supérieure de A et que M appartient à A, alors on dit que M est **le maximum** de A et on le note $\max(A)$.

Exemples :

- Le segment $]0{,}1[$ est majoré par 2. Sa borne supérieure est 1 car d'une part tout $x \in]0{,}1[$ est plus petit que 1 et, d'autre part, tout nombre $1-\varepsilon$ avec $\varepsilon > 0$ n'est pas majorant car $x = \max(1-\varepsilon/2, 1/2)$ est dans $]0{,}1[$ et est plus grand que $1-\varepsilon$. Par contre 1 n'appartient pas à $]0{,}1[$ donc 1 n'est pas le maximum de $]0{,}1[$ et écrire $\max(]0{,}1[)$ n'a pas de sens.
- $\{x \in \mathbb{Q}, \, x \leq 2\}$ a 2 pour maximum.
- $\{x \in \mathbb{Q}, \, x^2 \leq 2\}$ a $\sqrt{2}$ pour borne supérieure mais n'a pas de maximum.

Si $A = \emptyset$ est vide, alors A est majoré par tous les réels car une proposition concernant tous les éléments de $\emptyset$ est trivialement vérifiée (il n'y a aucun élément à considérer !). L'ensemble des majorants de l'ensemble vide est donc $\mathbb{R}$ tout entier et il n'y a pas de plus petit majorant. Donc $A = \emptyset$ n'a pas de borne supérieure. De même, un ensemble non majorée n'a aucun majorant et donc pas de borne supérieure. Mais pour simplifier les notations, il peut être agréable de quand même faire un abus de notation utilisant les infinis.

Définition 2.4

Si $A = \emptyset$ est vide, on pose $\sup(A) = -\infty$.
Si $A \subset \mathbb{R}$ n'est pas majorée, on pose $\sup(A) = +\infty$.

! Attention : il ne s'agit que d'une notation. C'est commode pour énoncer des propositions sans avoir à faire plusieurs cas, mais il faut se méfier si on veut utiliser ces définitions comme des nombres concrets puisque des calculs comme $\infty - \infty$ n'ont pas de sens.

On procède de même pour la borne inférieure.

Définition 2.5

Soit $A \subset \mathbb{R}$ un sous-ensemble des nombres réels.
On dit que $m \in \mathbb{R}$ est **un minorant** de A si

$$\forall a \in A, \ , \quad m \leq a \ .$$

On dit que A est **minoré** s'il existe un réel m qui minore A.
On dit que $m \in \mathbb{R}$ est **la borne inférieure** de A, notée $\inf(A)$, si c'est le plus grand minorant de A, c'est-à-dire que c'est un minorant et que tout nombre plus grand n'est plus un minorant.
Si $m \in \mathbb{R}$ est la borne inférieure de A et que m appartient à A, alors on dit que m est **le minimum** de A et on le note $\min(A)$.

De nouveau, pour simplifier les notations, on peut écrire

Définition 2.6

Si $A = \emptyset$ est vide, on pose $\inf(A) = +\infty$.
Si $A \subset \mathbb{R}$ n'est pas minorée, on pose $\inf(A) = -\infty$.

Et naturellement

Définition 2.7

Un ensemble à la fois majoré et minoré est dit **borné**.

Exemples :

- Remarquons que, par définition, l'ensemble vide est borné.
- Un intervalle du type $]a,b]$ est borné. Il admet un maximum qui est b et une borne inférieure qui est a mais n'a pas de minimum.
- Un intervalle du type $[a,+\infty[$ est minoré mais pas majoré, il n'est donc pas borné et n'admet pas de borne supérieur (ni de maximum). Il admet a comme minimum (et donc aussi comme borne inférieure).
- Supposons que $A \subset \mathbb{R}$ est un ensemble pour lequel on a trouvé une borne $M > 0$ telle que pour tout $a \in A$, $|a| \leq M$. Alors A est borné et contenu dans $[-M,M]$.
- L'ensemble $\{x \in \mathbb{R},\ x^2 \leq 1\}$ est borné (c'est le segment $[-1,1]$). Mais l'ensemble $\{x \in \mathbb{R},\ 1 \leq x^2\}$ n'est ni majoré, ni minoré (il est égal à $]-\infty,1] \cup [1,+\infty[$). Attention donc à bien décrypter ce qu'est l'ensemble avant de décider s'il est majoré ou minoré (ne pas se fier juste aux sens des inégalités le définissant).

De ces définitions découlent des propriétés classiques. Nous n'allons pas toutes les énoncer et plusieurs seront vues en TD. Voici quelques exemples dont l'intérêt réside aussi dans la preuve. Dans tous les cas, on admet l'existence des bornes supérieures qui sera énoncée dans le théorème 2.12 plus loin (et qui sera de toute façon admise).

Proposition 2.8

Soit A un ensemble non vide majoré de réel. L'ensemble des majorants de A est exactement $[\sup A, +\infty[$.

Démonstration : Soit $M \in \mathbb{R}$. Si $M < \sup A$, alors M n'est pas un majorant de A puisque $\sup A$ est le plus petit d'entre eux. Si $M \geq \sup A$, alors comme $\sup A$ majore A, pout tout $a \in A$, $a \leq \sup A \leq M$. Donc M majore A. □

Proposition 2.9

Si A et B sont deux ensembles non vides majorés de réels avec $A \subset B$, alors $\sup A \leq \sup B$.

Démonstration : Par définition, $\sup B$ est un majorant de B. Soit $a \in A$, comme $A \subset B$, a est aussi dans B et donc $a \leq \sup B$. Comme a est quelconque, on vient de montrer que $\sup B$ est un majorant de A. Comme $\sup A$ est le plus petit d'entre eux, on a $\sup A \leq \sup B$. □

Proposition 2.10

Si A et B sont deux ensembles non vides majorés de réels, alors $\sup(A \cup B) = \max(\sup A, \sup B)$.

Démonstration : Soit $M = \max(\sup A, \sup B)$. Pour tout $x \in A \cup B$, soit $x \in A$ et $x \leq \sup A \leq M$ (car $\sup A$ majore A), soit $x \in B$ et $x \leq \sup B \leq M$. Dans les deux cas, $x \leq M$ et donc M majore $A \cup B$.
Soit $m < M$. Supposons que $M = \sup A$ (sinon $M = \sup B$ et le raisonnement est symétrique). Comme $m < \sup A$, m n'est pas majorant de A (car par définition, $\sup A$ est le plus petit d'entre eux). Donc il existe $a \in A$ tel que $a > m$. Comme a est aussi dans $A \cup B$, m ne majore pas $A \cup B$ et M est bien le plus petit des majorants. □

Proposition 2.11

Soit $A \subset \mathbb{R}$ un ensemble non vide et majoré. Les propriétés suivantes sont équivalentes :

(i) le nombre M est la borne supérieure de A,

(ii) le nombre M majore A et il existe une suite (a_n) d'éléments de A qui converge vers M.

Démonstration : Si (ii) est vérifiée, on sait que M majore A donc il reste à montrer que c'est le plus petit majorant. Soit $x < M$ et soit $\varepsilon = M - x > 0$. Comme (a_n) tend vers M, il existe un rang n_0 à partir duquel $a_n > M - \varepsilon = x$. Donc $a_{n_0} > x$ et x ne peut pas majorer A. Donc M est bien la borne supérieure de A.
Supposons maintenant que (i) est vraie, i.e. que M est la borne supérieure de A. Par définition, M majore A. Nous allons construire la suite (a_n) comme suit. Pour tout $n \in \mathbb{N}$, on considère $M - 2^{-n}$. Comme M est le plus petit majorant, $M - 2^{-n}$ ne majore pas A et il existe au moins un élément a_n de A tel que $a_n > M - 2^{-n}$. Par ailleurs, comme M majore A, on a aussi $a_n \leq M$. Donc par théorème des gendarmes, on a bien $a_n \to M$. □

2 Les propriétés fondamentales de $\mathbb{R}$

Dans cette partie, nous allons voir plusieurs propriétés importantes des réels. Il s'agit de propriétés fondamentales dans le sens où elles proviennent plus ou moins directement de la construction même de $\mathbb{R}$. Pour les démontrer, il faudrait donc avoir défini concrètement les réels selon une des constructions possibles et utiliser les axiomes de la construction. Comme nous n'avons pas fait le travail de construction précisément, nous ne pourrons donc pas démontrer ces propriétés que nous admettrons.

• **Propriété de la borne supérieure**
Une des propriétés fondamentales des réels est l'existence d'une borne supérieure.

Théorème 2.12

Tout ensemble $A \subset \mathbb{R}$ non vide et majoré admet une borne supérieure.
Tout ensemble $A \subset \mathbb{R}$ non vide et minoré admet une borne inférieure.

Exemple :

L'ensemble
$$A = \{x \in \mathbb{R}\ ,\ \ x^5 - 3x - 1 \leq 0\}$$
est majoré et contient $x = 0$. Il admet donc une borne supérieure, mais qu'on ne peut pas écrire à l'aide des fonctions usuelles. C'est donc le théorème 2.12 qui donne l'existence de cette borne supérieure même si on ne peut pas écrire ce qu'elle vaut exactement.

Même si elle peut paraître naturelle, il s'agit d'une propriété *topologique* fondamentale de $\mathbb{R}$. Prenons l'ensemble $A = \{r \in \mathbb{Q},\ r^2 < 2\} \subset \mathbb{Q}$ qui est un ensemble de rationnels dont la définition n'a utilisé que des rationnels. Clairement, A est non vide car $0 \in A$ et A est majoré car si $x \in A$ alors $x \leq 2$ et pourtant A n'admet pas de borne supérieure dans $\mathbb{Q}$. Mais si on étend notre vue à tous les réels, c'est-à-dire qu'on regarde A comme sous-ensemble de $\mathbb{R}$, alors A admet une borne supérieure qu'on note $\sqrt{2}$. Donc le théorème 2.12 n'est plus vrai si on remplace $\mathbb{R}$ par $\mathbb{Q}$ et c'est la prise en compte de tous les réels qui permet d'obtenir toutes les bornes supérieures.

Il s'agit d'une propriété par nature assez élémentaire selon la construction de Dedekind, mais plus délicate à obtenir par la construction selon les suites de Cauchy.

• Complétude de $\mathbb{R}$

Rappelons la définition des suites de Cauchy.

Définition 2.13

Une suite de réels $(u_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ est dite **de Cauchy** si

$$\forall \varepsilon > 0 \ , \ \ \exists n_0 \in \mathbb{R} \ , \ \ \forall p \geq q \geq n_0 \ , \ \ |u_p - u_q| < \varepsilon \ .$$

On dit que $\mathbb{R}$ est complet car il vérifie la propriété suivante.

Théorème 2.14

Toute suite de Cauchy de $\mathbb{R}$ admet une limite dans $\mathbb{R}$.

Ce résultat est important car il permet de montrer qu'une suite converge sans même savoir quelle pourrait en être la limite. Il est à la base de nombreux résultats. Notons encore que tous les ensembles n'ont pas cette propriété puisque $\mathbb{Q}$ possède des suites de Cauchy qui ne convergent pas dans $\mathbb{Q}$ (par exemple si on considère une suite d'approximations rationnelles de $\sqrt{2}$).

Les théorèmes 2.12 et 2.14 énoncent des résultats en fait très semblables puisqu'ils reviennent à dire que, intuitivement, $\mathbb{R}$ « n'a pas de trous ».

• Densité des rationnels

Nous introduirons plus tard une notion de densité plus générale. Dans $\mathbb{R}$, cette notion s'ecrit comme suit.

Définition 2.15

Un sous-ensemble $A \subset \mathbb{R}$ est dit **dense** s'il vérifie une des caractérisations équivalentes suivantes :

i) chaque voisinage d'un point x de $\mathbb{R}$ contient un point de A

$$\forall x \in \mathbb{R} \ , \ \ \forall \varepsilon > 0 \ , \ \ \exists a \in A \ , \ \ a \in \,]x - \varepsilon, x + \varepsilon[.$$

ii) A rencontre tout intervalle ouvert de $\mathbb{R}$

$$\forall (x,y) \in \mathbb{R}^2 \ , \ \ x < y \quad \Longrightarrow \quad (\exists a \in A \ , \ \ x < a < y) \, .$$

iii) tout réel peut être approché par une suite de A

$$\forall x \in \mathbb{R} \ , \ \ \exists (a_n)_{n\in\mathbb{N}} \subset A \ , \ \ a_n \xrightarrow[n\to+\infty]{} x.$$

La construction des réels nous donne assez facilement que $\mathbb{Q}$ est dense dans $\mathbb{R}$.

Il est un peu moins évident de voir que c'est aussi le cas des irrationnels.

Théorème 2.16

Les ensembles $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont tous les deux denses dans $\mathbb{R}$.

Cela implique par exemple qu'entre deux réels, il y a une infinité de rationnels et aussi une infinité d'irrationnels.

• Corps ordonné archimédien

La définition suivante peut paraître triviale, mais il s'agit d'une propriété importante des réels. Les grecs l'avaient déjà identifiée puisqu'on trouve dans le livre V des Éléments d'Euclide la définition « *des grandeurs sont dites avoir une raison entre elles lorsque ces grandeurs, étant multipliées, peuvent se surpasser mutuellement* ».

Définition 2.17

Un corps ordonné K est **archimédien** si pour tout x et y dans K avec $y > x > 0$, il existe un entier n tel que $n.x > y$.

On peut utiliser le résultat suivant comme une définition des réels.

Théorème 2.18

L'ensemble $\mathbb{R}$ muni de ses structures habituelles est l'unique corps ordonné archimédien vérifiant la propriété de la borne supérieure.

Par « unique corps », il faut comprendre qu'un autre corps qui aurait les mêmes propriétés peut être mis en bijection avec $\mathbb{R}$ avec une bijection transportant exactement toutes les structures en question (l'image de la somme est la somme des images, l'ordre des images est le même que celui des antécédents etc.).

Exemple :

On considère la suite définie par $u_n = 1/n$ pour $n \in \mathbb{N}^*$. Soit $\varepsilon > 0$. Comme $\mathbb{R}$ est archimédien, il existe $n_0 \in \mathbb{N}^*$ tel que $n_0\varepsilon > 1$. Mais alors $0 < 1/n_0 < \varepsilon$ et donc $|u_{n_0}| < \varepsilon$. Les règles liant relation d'ordre et opérations sur $\mathbb{R}$, qui viennent du fait que $\mathbb{R}$ est un corps ordonné, montrent que pour $n \geq n_0$, $0 < u_n \leq u_{n_0}$. Donc pour tout $n \geq n_0$, $|u_n| < \varepsilon$. On vient de montrer que la suite $u_n = 1/n$ tend vers 0 grâce aux propriétés fondamentales de $\mathbb{R}$.

3 Suites réelles

Une suite réelle $(u_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ est une fonction de $\mathbb{N}$ dans $\mathbb{R}$. Il est plus agréable de la voir comme une liste ordonnée u_0, u_1, $u_2 \ldots$ de réels, l'indice $n \in \mathbb{N}$ étant un temps discret qui s'écoule. Le sujet principal d'étude des suites est de comprendre leur comportement quand $n \to +\infty$. Rappelons la définition classique.

Définition 2.19

Une suite réelle $(u_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ admet le réel $\ell \in \mathbb{R}$ comme **limite** si

$$\forall \varepsilon > 0 \ , \ \ \exists n_0 \in \mathbb{N} \ , \ \ \forall n \geq n_0 \ , \ \ |u_n - \ell| < \varepsilon \ .$$

On note $\lim\limits_{n\to+\infty} u_n = \ell$ ou $u_n \xrightarrow[n\longrightarrow+\infty]{} \ell$.
On dit aussi que (u_n) **converge** ou **tend vers** ℓ.
On dit (u_n) est **convergente**, ou bien converge, s'il existe $\ell \in \mathbb{R}$ tel que (u_n) tend vers ℓ. Dans le cas contraire, on dit que (u_n) est **divergente**.

On retrouve ci-dessus une transcription dans le formalisme des quantificateurs de Cauchy de ce que Gottfried Wilhelm Leibniz (1646-1716, Allemagne) tentait d'écrire avec des phrases comme

> *« à mesure qu'on la considère de plus en plus loin, l'erreur sera moindre que toute grandeur donnée. »*

On peut aussi définir une notion de convergence vers $\pm\infty$.

Définition 2.20

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle. On dit que (u_n) a pour **limite** $+\infty$ si

$$\forall M \in \mathbb{R} \ , \ \ \exists n_0 \in \mathbb{N} \ , \ \ \forall n \geq n_0, \ \ u_n \geq M \ .$$

Symétriquement, on dit que (u_n) a pour **limite** $-\infty$ si

$$\forall M \in \mathbb{R} \ , \ \ \exists n_0 \in \mathbb{N} \ , \ \ \forall n \geq n_0, \ \ u_n \leq M \ .$$

Attention, si on peut parler de limite infinie, on considère bien qu'une suite qui tend vers $\pm\infty$ est divergente.

Nous n'allons pas reprendre toutes les propriétés des limites réels ni leurs démonstrations. Nous allons reprendre seulement les propriétés suivantes, d'une part car elles nous seront utiles pour la suite et d'autre part pour bien montrer qu'elles proviennent des propriétés fondamentales de $\mathbb{R}$.

Théorème 2.21

Toute suite croissante majorée converge (vers une limite réelle finie). Symétriquement, toute suite décroissante minorée converge (vers une limite réelle finie).

Démonstration : Si (u_n) est croissante et majorée, alors l'ensemble $\{u_n, n \in \mathbb{N}\}$ est un ensemble majorée non vide de $\mathbb{R}$. Il admet donc une borne supérieure $\ell \in \mathbb{R}$. Par construction, comme ℓ est un majorant, $u_n \leq \ell$ pour tout $n \in \mathbb{N}$. Comme ℓ est le plus petit majorant, pour tout $\varepsilon > 0$, $\ell - \varepsilon$ n'est plus un majorant et il existe $n_0 \in \mathbb{N}$ tel que $u_{n_0} > \ell - \varepsilon$. Comme la suite est croissante,

on a que pour tout $n \geq n_0$, $u_n \geq u_{n_0} > \ell - \varepsilon$. On a donc que pour tout $n \geq n_0$, $u_n \in]\ell - \varepsilon, \ell]$ et donc que $|u_n - \ell| < \varepsilon$.
Le cas symétrique se démontre de la même façon. $\square$

Nous avons utilisé la propriété de la borne supérieure. Cette proposition n'est pas vraie dans $\mathbb{Q}$ car la suite des approximations décimales de π (écriture décimale tronquée à n chiffres) est une suite de rationnels croissante et majorée mais qui ne converge pas dans $\mathbb{Q}$.

Corollaire 2.22

Si $(u_n)_{n\in\mathbb{N}}$ est une suite croissante, alors on a l'alternative suivante :

i) soit (u_n) est majorée et converge vers une limite finie,
ii) soit (u_n) n'est pas majorée et diverge vers $+\infty$.

Démonstration : Si (u_n) est majorée, alors le théorème 2.21 conclut. Supposons donc que (u_n) n'est pas majorée : pour tout $M \in \mathbb{R}$, il existe forcément $n_0 \in \mathbb{N}$ tel que $u_{n_0} > M$ puisque M ne peut pas être un majorant. Mais comme la suite est croissante, on a alors pour tout $n \geq n_0$, $u_n \geq u_{n_0} > M$. Par définition, cela veut dire que (u_n) tend vers $+\infty$. $\square$

Un autre résultat important utilisant une des propriétés fondamentales de $\mathbb{R}$ est celui des suites adjacentes.

Théorème 2.23 (suites adjacentes)

Soient (u_n) et (v_n) deux suites que l'on suppose **adjacentes** c'est-à-dire que :

i) la suite (u_n) est croissante,
ii) la suite (v_n) est décroissante,
iii) on a $|u_n - v_n| \to 0$ quand n tend vers $+\infty$.

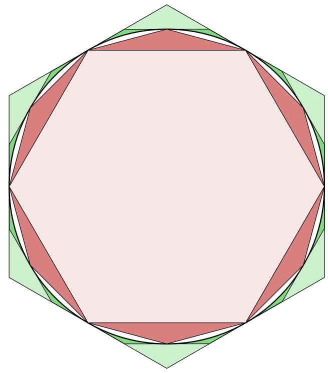
Alors (u_n) et (v_n) convergent toutes les deux vers une même limite $\ell \in \mathbb{R}$ et on a l'encadrement
$$\forall n \in \mathbb{N}\ , \quad u_n \leq \ell \leq v_n\ .$$

Démonstration : D'après la définition de la convergence, il existe un rang $n_0 \in \mathbb{N}$ tel que, pour tout $n \geq n_0$, $|u_n - v_n| \leq 1$. En particulier, à partir de ce rang, $v_n \geq u_n - 1$. D'après la monotonie des suites, si $n \geq n_0$, on a $v_n \geq u_n - 1 \geq u_{n-1} - 1 \geq \ldots \geq u_0 - 1$. Pour $n \leq n_0$, on a aussi $v_n \geq v_{n_0} \geq u_{n_0} - 1 \geq u_0 - 1$. La suite (v_n) est donc décroissante et minorée, elle converge vers une limite $\ell \in \mathbb{R}$ d'après le théorème 2.21. Par ailleurs, la preuve de ce théorème montre bien que ℓ est la borne inférieure de (v_n) et donc que $\ell \leq v_n$ pour tout n.
On peut faire le raisonnement symétrique pour (u_n) et montrer qu'elle converge vers une limite $\ell' \in \mathbb{R}$. Il reste à voir qu'il s'agit en fait de la même limite. Mais

comme les deux suites convergent, leur différence $u_n - v_n$ tend vers $\ell' - \ell$ et donc $|u_n - v_n|$ tend vers $|\ell' - \ell|$. Par hypothèse, ce nombre est égal à 0 donc $\ell = \ell'$. □

Exemples :

- Un exemple très visuel est donné par la méthode de quadrature du cercle qu'a utilisée Archimède au IIIème siècle avant J.C. pour donner l'encadrement $3+10/71 < \pi < 3+1/7$. On note u_n la surface du polygone à 6×2^n côtés inscrit dans le cercle et v_n la surface du polygone à 6×2^n côtés circonscrit au cercle. On part donc d'un hexagone puis on double le nombre de côtés comme sur la figure jointe. Il est clair que (u_n) est croissante et (v_n) décroissante et on peut se



  convaincre que $u_n - v_n \to 0$. Le calcul de u_n et v_n peut se faire par itération de formules trigos. On obtient deux suites adjacentes qui donnent des approximations de plus en plus proches de π. Archimède avait été jusqu'à $n = 5$, c'est-à-dire pour des polygones à 96 côtés, obtenant les trois premiers chiffres significatifs de π. Notons que cette construction de suites adjacentes nous fournit l'existence du nombre qu'on appelle π.

- On considère les suites (u_n) et (v_n) données par

$$u_n = \sum_{k=0}^{n} \frac{1}{k!} = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \ldots + \frac{1}{n!} \quad \text{et} \quad v_n = u_n + \frac{1}{n!} .$$

  Comme u_n est une somme de plus en plus grande de termes positifs, la suite (u_n) est clairement croissante. On peut montrer que (v_n) est décroissante (calcul laissé au lecteur ou aux TDs). Comme $|u_n - v_n| = 1/n! \to 0$, il s'agit de deux suites adjacentes. La puissance du théorème 2.23 est de montrer qu'elles convergent, même si on ne sait rien de cette limite. Il se trouve que leur limite est un nombre important en mathématique, on va donc lui donner un nom : c'est le nombre e. On peut ainsi s'autoriser à écrire

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \ldots$$

Le résultat suivant s'appuie sur les mêmes arguments mais avec un point de vue plus géométrique.

Théorème 2.24 (théorème des segments emboîtés)

Soient $I_n = [a_n, b_n]$ une suite de segments fermés bornés qui sont emboîtés dans le sens où $I_{n+1} \subset I_n$. Alors l'intersection $\cap_{n\in\mathbb{N}} I_n$ est non vide.

Démonstration : Comme les segments sont emboîtés, la suite (a_n) est croissante et majorée par b_0 et la suite (b_n) est décroissante et minorée par a_0. Toutes les deux convergent donc vers des limites a_∞ et b_∞. On vérifie facilement que $\cap_{n\in\mathbb{N}} I_n = [a_\infty, b_\infty]$. $\square$

Exemples :

- On procède à une recherche par dichotomie. On part du segment $I_0 = [0,1]$ que l'on coupe en deux et on choisit l'une des moitiés $I_1 = [0,1/2]$ ou $I_2 = [1/2,1]$. Puis I_1 est de nouveau coupé en deux et on choisit pour I_2 l'une des moitiés etc. Le théorème 2.24 nous assure qu'il restera au moins un point à l'intersection de tous ces choix (en fait juste un seul car la taille des segments tant vers 0).

- Revenons à l'ensemble de Cantor avec la vision géométrique. On part du segment $I_0 = [0,1]$ puis on coupe l'intervalle en deux en retirant le tiers central. On choisit pour I_1 l'un des tiers restants et on le redécoupe puis on choisit pour I_2 un des tiers restants etc. De nouveau, on est assuré qu'il reste un point à l'intersection de tous ces choix. Non seulement cela montre que l'ensemble de Cantor est non vide, mais cela montre qu'on peut exactement le mettre en bijection avec les suites infinies de choix du type $DGGGDDGDGG\ldots$. Par l'argument diagonal de Cantor, on montre que l'ensemble de ces suites est indénombrable.

- On fait attention que, même si la propriété peut paraître évidente, elle repose sur des propriétés topologiques. Par exemple, elle n'est plus vraie sur $\mathbb{Q}$ puisque $\cap_{n\in\mathbb{N}}[a_n,b_n] = \emptyset$ si (a_n) et (b_n) sont des approximations décimales de $\sqrt{2}$ par en-dessous et par au-dessus à 10^{-n} près. Si on reste sur $\mathbb{R}$, il faut faire attention à la topologie des intervalles. Ainsi $\cap_{n\in\mathbb{N}}]0,2^{-n}] = \emptyset$ et $\cap_{n\in\mathbb{N}}[n,+\infty[= \emptyset$.

Une suite convergente est forcément bornée, mais il existe des suites bornées qui ne converge pas comme $(-1)^n$. La propriété fondamentale suivante correspond à ce que nous appellerons plus tard la *compacité*. Elle permet d'obtenir quand même une convergence, quitte à ne considérer qu'une partie de la suite. Son nom vient des mathématiciens Bernard Bolzano (1781-1848, Hongrie) et Karl Weierstrass (1815-1897, Allemagne).

Théorème 2.25 (théorème de Bolzano-Weierstrass)

Soit $(u_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ une suite réelle bornée. Alors il existe une fonction $\varphi : \mathbb{N} \to \mathbb{N}$ strictement monotone telle que la suite extraite $(u_{\varphi(n)})_{n\in\mathbb{N}}$ converge vers une limite $\ell \in \mathbb{R}$.

Démonstration : Notre raisonnement consiste à appliquer une méthode de dichotomie. Comme (u_n) est bornée, il existe un segment $[a_0,b_0]$ tel que $u_n \in [a_0,b_0]$ pour tout $n \in \mathbb{N}$. On prend $\varphi(0) = 0$. Soit $m = (a_0 + b_0)/2$ le milieu de $[a_0,b_0]$. Comme il y a une infinité de termes de la suite dans $[a_0,b_0]$, il faut bien qu'il y en ait une infinité dans au moins une des moitié $[a_0,m]$ et/ou $[m,b_0]$ (principe des tiroirs). On note $[a_1,b_1]$ une des moitiés qui convient et on pose $\varphi(1)$ comme le premier $n > \varphi(0) = 0$ tel que u_n soit dans $[a_1,b_1]$. De nouveau, on prend le milieu $m = (a_1 + b_1)/2$. Comme il y a une infinité de termes de la suite dans $[a_1,b_1]$, il faut bien qu'il y en ait une infinité dans au moins une des moitié $[a_0,m]$ et/ou $[m,b_0]$, que l'on note $[a_2,b_2]$. On pose $\varphi(2)$ comme le premier $n > \varphi(1)$ tel que u_n soit dans $[a_2,b_2]$. De proche en proche, on peut extraire ainsi une sous-suite. À partir du rang n_0, tous les termes de la sous-suite se retrouvent dans un intervalle de taille $2^{-n_0}|b_0 - a_0|$ puisqu'on divise l'intervalle considéré par deux à chaque étape. Cela montre que cette sous-suite est une suite de Cauchy. Le théorème 2.14 nous dit qu'elle converge donc. □

Quand on parle de compacité, on pense plutôt à prendre une suite de points dans un ensemble. La version topologique du théorème précédent s'énoncera plutôt comme suit.

Corollaire 2.26 (compacité simplifiée dans $\mathbb{R}$)

Soient $a < b$ deux réels. Soit $(x_n) \subset [a,b]$ une suite de points de l'intervalle fermé borné $[a,b]$, alors on peut en extraire une sous-suite $(x_{\varphi(n)})$ qui converge vers une limite $\ell \in [a,b]$.

Démonstration : Comme $a \leq x_n \leq b$ pour tout n, la suite est bornée et on peut utiliser le théorème précédent pour en extraire une sous-suite $(x_{\varphi(n)})$ qui converge vers une limite $\ell \in \mathbb{R}$. Comme $x_{\varphi(n)} \geq a$ pour tout n, les théorèmes de comparaison nous disent que $\ell \geq a$. De même, on obtient que $\ell \leq b$, ce qui conclut. □

Les limites possibles des suites extraites forment ce qu'on appelle les *valeurs d'adhérence* de la suite.

Définition 2.27

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle. On dit que $\ell \in \mathbb{R}$ est une **valeur d'adhérence** de (u_n) s'il existe une sous-suite $(u_{\varphi(n)})$ de (u_n) qui converge vers ℓ.

Exemples :

- La suite $(-1)^n$ a deux valeurs d'adhérence qui sont ± 1.
- Les valeurs d'adhérence de la suite $u_n = \cos n$ forment exactement le segment $[-1,1]$, mais ce n'est pas facile à montrer (il faut utiliser que $\mathbb{Z} + \pi\mathbb{Z}$ est dense dans $\mathbb{R}$).
- La suite définie par $u_n = 0$ si n pair et $u_n = n$ si n impair est une suite non bornée mais qui a 0 pour valeur d'adhérence.
- Le théorème 2.25 dit donc que toute suite bornée a au moins une valeur d'adhérence.

Le concept de valeur d'adhérence se généralise à des espaces autres que $\mathbb{R}$ (voir chapitres suivants). Il existe par contre un concept généralisant celui de limites qui est fortement lié à l'ordre sur $\mathbb{R}$.

Définition 2.28

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle. Si (u_n) est majorée, on appelle **limite supérieure** ou plus simplement « limsup » la limite

$$\limsup_{n\to+\infty} u_n \;:=\; \lim_{n\to+\infty} \left(\sup_{k\geq n} u_k\right) .$$

Si (u_n) est minorée, on appelle **limite inférieure** ou plus simplement « liminf » la limite

$$\liminf_{n\to+\infty} u_n \;:=\; \lim_{n\to+\infty} \left(\inf_{k\geq n} u_k\right) .$$

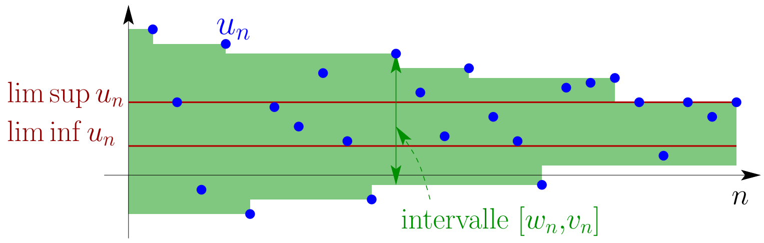


FIGURE 2.2 – *L'exemple d'une suite qui ne converge pas mais qui possède une limite supérieure et une limite inférieure. En vert, on voit l'intervalle dans lequel se trouve tous les u_k pour $k\geq n$, qui se réduit petit à petit pour converger vers l'intervalle* $[\liminf u_n, \limsup u_n]$.

La définition ci-dessus semble assurer que les limites existent bien. C'est en effet le cas comme le dit la proposition suivante.

Proposition 2.29

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle. Si (u_n) est majorée par $M\in\mathbb{R}$, alors $\limsup u_n$ existe et appartient à $[-\infty,M]$. Si (u_n) est minorée par $m\in\mathbb{R}$, alors $\liminf u_n$ existe et appartient à $[m,+\infty]$.
La suite (u_n) est bornée si et seulement si les limsup et liminf existent et sont finies. On a alors toujours $\liminf u_n \leq \limsup u_n$, avec égalité si et seulement si (u_n) converge.

Démonstration : Si (u_n) est majorée par M, l'ensemble $\{u_k,\ k\geq n\}$ est majoré par M et non vide. On peut donc poser $v_n=\sup_{k\geq n} u_k$ et $v_n\leq M$ par définition de la borne supérieure. Comme $\{u_k,\ k\geq n+1\}\subset\{u_k,\ k\geq n\}$, on a $v_{n+1}\leq v_n$. La suite (v_n) étant décroissante, soit elle tend vers $-\infty$, soit elle converge vers une limite finie, forcément inférieure à M puisque $v_n\leq M$ pour tout n.

Le cas de la liminf se traite pareil en posant $w_n = \inf_{k\geq n} u_k$, qui est une suite croissante.
Par définition, on a $w_n \leq u_n \leq v_n$. Si les limsup et liminf existent et sont finies, (w_n) et (v_n) sont deux suites convergentes dans $\mathbb{R}$. En particulier, elles sont bornées et donc (u_n) est aussi bornée. Réciproquement, si $(u_n) \subset [m,M]$, alors on a $m \leq w_n \leq u_n \leq v_n \leq M$. La suite (v_n) est décroissante minorée et (w_n) est croissante majorée. Donc ces deux suites convergent vers des limites finies. De $w_n \leq u_n \leq v_n$ on obtient que $\liminf u_n \leq \limsup u_n$ et que s'il y a égalité, alors (u_n) converge vers la même limite (théorème des gendarmes).
Il reste à montrer que si (u_n) converge vers $\ell \in \mathbb{R}$, alors $\liminf u_n = \limsup u_n = \ell$. Soit $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que si $k \geq n_0$, alors $\ell - \varepsilon \leq u_k \leq \ell + \varepsilon$. Donc pour tout $n \geq n_0$, on a $v_n \geq u_n \geq \ell - \varepsilon$ (la borne sup est un majorant) et aussi $v_n \leq \ell + \varepsilon$ (car $\ell + \varepsilon$ est aussi un majorant et la borne sup est le plus petit de tous). D'où $|\ell - v_n| \leq \varepsilon$. On en déduit que (v_n) tend vers ℓ et donc $\limsup u_n = \ell$. Le cas de la limite inférieure est symétrique. □

Il est possible que $\limsup u_n = -\infty$, ce qui correspond à $u_n \to -\infty$. La définition n'est pas alors abusive car elle correspond à une suite réelle (v_n) qui tend vers $-\infty$. Mais dans ce cas, la suite (u_n) n'est pas minorée et on a $\inf_{k\geq n} u_k = -\infty$. La suite (w_n) serait alors toujours égale à $-\infty$ et ce n'est plus une suite réelle. On utilisera alors l'abus de notation $\liminf u_n = -\infty$ même si cela ne correspond pas à la définition de la limite inférieure. De même, on posera $\limsup u_n = +\infty$ pour une suite non majorée.

Exemples :

- Si $u_n = (-1)^n$, on a $\liminf u_n = -1$ et $\limsup u_n = 1$.
- On pose $u_n = (-1)^n + 1/n$. On a $v_{2n-1} = v_{2n} = (-1)^n + 1/n$ et $w_n = -1$. Donc $\liminf u_n = -1$ et $\limsup u_n = 1$.
- On regarde de nouveau la suite définie par $u_n = 0$ si n pair et $u_n = n$ si n impair. On a $\liminf u_n = 0$ et $\limsup u_n = +\infty$.

Proposition 2.30

Soit $(u_n)_{n\in\mathbb{N}}$ une suite bornée. Alors $\limsup u_n$ et $\liminf u_n$ sont respectivement les plus grandes et plus petites valeurs d'adhérence de (u_n).

Démonstration : Soit ℓ une valeur d'adhérence de (u_n), c'est-à-dire la limite d'une sous-suite $(u_{\varphi(n)})$. On a $w_{\varphi(n)} \leq u_{\varphi(n)} \leq v_{\varphi(n)}$ et à la limite on trouve $\liminf u_n \leq \ell \leq \limsup u_n$. Il reste à montrer que les limsup et liminf sont des valeurs d'adhérence.
Soit $\ell = \limsup u_n$, qui est la limite de $v_n = \sup_{k\geq n} u_k$. Comme v_n tend vers ℓ en décroissant, à partir d'un rang n_1, on a $\ell \leq v_n \leq \ell + 1$. Par définition, on a $u_k \leq v_{n_1}$ pour tout $k \geq n_1$. Comme $v_{n_1} - 1$ ne majore pas $\sup_{k\geq n_1} u_k$,

il existe $k_1 \geq n_1$ tel que $v_{n_1} - 1 \leq u_{k_1}$. On pose $\varphi(1) = k_1$ et on note que $\ell - 1 \leq u_{\varphi(1)} \leq \ell + 1$.
On prend maintenant un rang n_2 à partir duquel $\ell \leq v_n \leq \ell + 1/2$. Quitte à le prendre plus grand encore, on peut supposer que $n_2 > \varphi(1)$. Par définition, on a $u_k \leq v_{n_2}$ pour tout $k \geq n_2$. Comme $v_{n_2} - 1/2$ ne majore pas $\sup_{k \geq n_2} u_k$, il existe $k_2 \geq n_2$ tel que $v_{n_2} - 1/2 \leq u_{k_2}$. On pose $\varphi(2) = k_2$ et on note que $\ell - 1/2 \leq u_{\varphi(2)} \leq \ell + 1/2$ et aussi que $\varphi(2) > \varphi(1)$.
On continue ainsi de suite et on construit une suite extraite $(u_{\varphi(n)})$ telle que $\ell - 1/n \leq u_{\varphi(n)} \leq \ell + 1/n$. Ceci montre que ℓ est une valeur d'adhérence de (u_n).
La preuve pour la liminf est évidemment symétrique. □

4 Fonctions réelles

Nous allons ici revoir certaines définitions et propriétés concernant les fonctions réelles, c'est-à-dire les fonctions du type

$$f \ : \ \mathcal{D}_f \subset \mathbb{R} \longrightarrow \mathbb{R} \ .$$

Nous allons le faire avec le même esprit que la partie sur les suites : nous n'allons pas tout reprendre mais nous concentrer sur certains aspects qui seront utiles pour les prochains chapitres, soit comme propriété fondamentale, soit comme modèle pour une généralisation. Notons qu'on ne parlera que très peu de dérivation dans ce cours. En effet, la topologie est surtout liée aux concepts de limites et de continuité. L'aspect dérivation et intégration relève plutôt de ce qu'on appelle la *géométrie différentielle* et sera vue au second semestre. Cela n'empêchera évidemment pas d'utiliser les concepts d'intégrale ou de dérivée, mais ils ne seront pas l'objet principal de notre étude. En particulier, ils n'apparaîtront pas dans cette partie.

La définition de la limite dans le cas des fonctions suit exactement les mêmes principes que dans le cas des suites. Il y a beaucoup de cas différents donc plutôt que de les apprendre tous par cœur, l'important est de comprendre comment ils sont construits :

- les voisinages d'un point $x \in \mathbb{R}$ sont du type $]x - \varepsilon, x + \varepsilon[$,
- les voisinages de $+\infty$ sont du type $]M, +\infty[$,
- les voisinages de $-\infty$ sont du type $]-\infty, M[$.

et on peut retenir le principe général :

La définition de
« *$f(x)$ tend vers $\ell \in [-\infty, +\infty]$ quand x tend vers $\ell' \in [-\infty, +\infty]$* »
s'écrit
« *pour tout voisinage $\mathcal{V}$ de ℓ, il existe un voisinage $\mathcal{V}'$ de ℓ' tel que $f(\mathcal{V}') \subset \mathcal{V}$* »
ce qui peut se comprendre comme
« *si x est suffisamment proche de ℓ', alors $f(x)$ reste aussi proche de ℓ que voulu* ».

Nous n'allons pas écrire les 9 cas différents de limites (voire plus si on compte aussi les limites à droite et à gauche). De toute façon, ce concept de limite sera revu dans un cadre bien plus général ensuite. Voyons juste quelques exemples.

Définition 2.31

Soit $f : \mathcal{D}_f \to \mathbb{R}$ une fonction réelle définie autour de x_0 dans le sens où pour tout $\delta > 0$, $]x_0 - \delta, x_0 + \delta[\cap \mathcal{D}_f$ est non vide. Alors,

- on dit que f tend vers $\ell \in \mathbb{R}$ quand x tend vers x_0 si

$$\forall \varepsilon > 0 \ , \ \ \exists \delta > 0 \ , \ \ \forall x \in]x_0 - \delta, x_0 + \delta[\cap \mathcal{D}_f \ , \ \ |f(x) - \ell| < \varepsilon \ .$$

- On dit que f tend vers $+\infty$ quand x tend vers x_0 si

$$\forall M \in \mathbb{R} \ , \ \ \exists \delta > 0 \ , \ \ \forall x \in]x_0 - \delta, x_0 + \delta[\cap \mathcal{D}_f \ , \ \ f(x) \geq M \ .$$

Définition 2.32

Soit $f : \mathcal{D}_f \to \mathbb{R}$ une fonction réelle définie près de $+\infty$, c'est-à-dire que pour tout $x_0 > 0$, $[x_0, +\infty[\cap \mathcal{D}_f$ est non vide. On dit que f tend vers $-\infty$ quand x tend vers $+\infty$ si

$$\forall M \in \mathbb{R} \ , \ \ \exists x_0 \in \mathbb{R} \ , \ \ \forall x \in [x_0, +\infty[\cap \mathcal{D}_f \ , \ \ f(x) \leq M \ .$$

Ces définitions sont un peu lourdes pour pouvoir inclure le cas où f n'est pas définie partout, ni même dans un voisinage de $+\infty$. Par exemple, on notera que si f est définie de $\mathbb{N}$ dans $\mathbb{R}$, alors la notion de limite de f en $+\infty$ retombe bien sur la notion de limite pour la suite $(f(n))_{n\in\mathbb{N}}$. C'est naturel puisqu'une suite n'est rien d'autre qu'une fonction définie sur $\mathbb{N}$.

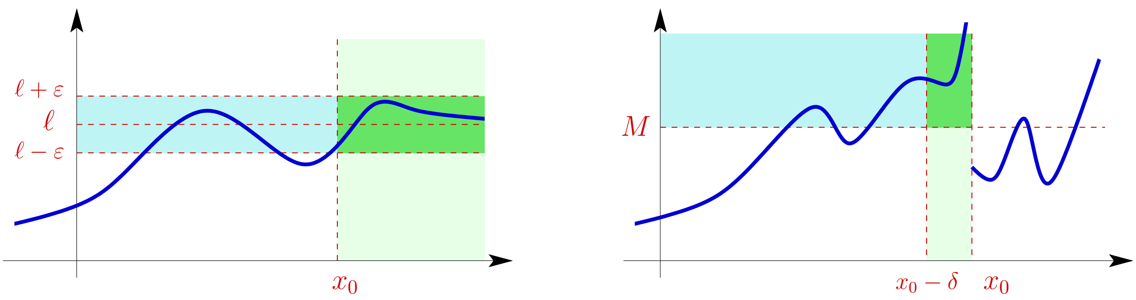


Figure 2.3 – *Des illustrations d'autres exemples de limites. À gauche un exemple où $f(x) \to \ell$ quand $x \to +\infty$: pour tout $\varepsilon > 0$, il existe un point x_0 à partir duquel, pour $x \geq x_0$, $f(x)$ est toujours dans l'intervalle $]\ell - \varepsilon, \ell + \varepsilon[$. À droite un exemple où $f(x) \to +\infty$ quand $x \to x_0^-$: pour tout $M > 0$, il existe un intervalle $]x_0 - \delta, x_0[$ tel que pour tout $x \in]x_0 - \delta, x_0[$, $f(x)$ est toujours plus grand que M.*

À partir de ces définitions, on retrouvera toutes les propriétés des limites qu'on connaît déjà. Comme on fera le cas général plus loin, nous laissons au lecteur la démonstration dans le cas de $\mathbb{R}$. Voyons un exemple d'énoncé pour s'entraîner. Notons qu'il s'agit d'une propriété fortement liée à $\mathbb{R}$ dans le sens où la preuve utilise la propriété de la borne supérieure.

Proposition 2.33

Soit $f : [0,+\infty[\to \mathbb{R}$ une fonction décroissante et minorée. Alors $f(x)$ converge vers une limite finie $\ell \in \mathbb{R}$ quand x tend vers $+\infty$.

Démonstration : Comme f est minorée, son image $\{f(x), x \geq 0\}$ est non vide et minorée et admet donc une borne inférieure $\ell := \inf_{x\geq 0} f(x)$. Comme ℓ est un minorant de l'image de f, pour tout $x \geq 0$, $f(x) \geq \ell$. Soit $\varepsilon > 0$. Comme $\ell + \varepsilon$ n'est plus un minorant de l'image de f, il existe $x_0 \geq 0$ tel que $f(x_0) < \ell + \varepsilon$. Mais comme f est décroissante, on a $f(x) < \ell + \varepsilon$ pour tout $x \geq x_0$. Au total, on a $\ell \leq f(x) < \ell + \varepsilon$ pour tout $x \geq x_0$, ce qui montre que f tend vers ℓ quand x tend vers $+\infty$. $\square$

Exemple :

On considère les étirements $y(t)$ d'un ressort de raideur k qui est soumis à un frottement d'intensité γ. La longueur $y(t)$ est solution de l'équation différentielle $my''(t) + \gamma y'(t) = -ky(t)$. Si on considère l'énergie $E(t) = \frac{1}{2}(m|y'(t)|^2 + k|y(t)|^2)$, on a $E'(t) = my'(t)y''(t) + ky(t)y'(t) = -\gamma|y'(t)|^2 \leq 0$. On trouve que $E(t)$ est décroissante et elle est clairement positive, donc $E(t)$ admet une limite finie positive quand $t \to +\infty$. Ceci est la première étape pour montrer que l'énergie se dissipe jusqu'à ce que le ressort revienne à l'équilibre.

Pour les fonctions réelles, il y a plusieurs façons équivalentes de définir la continuité. On peut donc en choisir une comme définition de base et les autres comme caractérisations équivalentes.

Définition 2.34

Soit $f : \mathcal{D}_f \to \mathbb{R}$ une fonction réelle. On dit que f est **continue en** $x_* \in \mathcal{D}_f$ si on a

$$\forall \varepsilon > 0 \ , \ \exists \delta > 0 \ , \ \forall x \in \mathcal{D}_f \ , \ |x - x_*| < \delta \implies |f(x) - f(x_*)| < \varepsilon \ .$$

On dit que f est **continue sur un ensemble** $E \subset \mathcal{D}_f$ si f est continue en tout point de E.
On note $\mathcal{C}^0(E,F)$ l'ensemble des fonctions continues sur $E \subset \mathbb{R}$ à valeurs dans $F \subset \mathbb{R}$.

Exemples :

- Là où elles sont définies, les fonctions usuelles sont continues, sauf la partie entière. Donc toute fonction définie par une formule sera continue là où la formule fait sens (sauf dans le cas rare où la partie entière entre en jeu).

- Beaucoup de grandeurs physiques sont en général considérées comme continues, comme la température, la position, la vitesse... Si bien qu'on pourrait penser qu'il n'y a pas à s'embêter avec les fonctions discontinues. Mais dans beaucoup de modèles, il est intéressant de prendre des fonctions discontinues. Par exemple, si on modélise un circuit électronique dont on allume le courant avec un interrupteur au temps $t = 0$, il est plus simple de penser que l'intensité est la fonction I définie par $I(t) = 0$ si $t \leq 0$ et $I(t) = 1$ si $t > 0$ qui est discontinue en 0. En effet, même si la vraie intensité est possiblement continue à cause d'un passage de courant progressif quand l'interrupteur se ferme, cela est trop compliqué à modéliser et il est très probablement non pertinent de s'embêter avec cela. On préférera donc une fonction discontinue. De la même façon, si on regarde une bille qui fait un rebond parfait sur un mur, on supposera le choc élastique. Si la position varie continûment par rapport au temps, sa vitesse sera réfléchie instantanément lors du rebond et ne sera pas continue. Là encore, si on regarde tout en détail, le changement n'est pas immédiat, mais alors la conservation du moment cinétique obligerait à prendre en compte les déformations du mur, ce qui est trop difficile à faire.

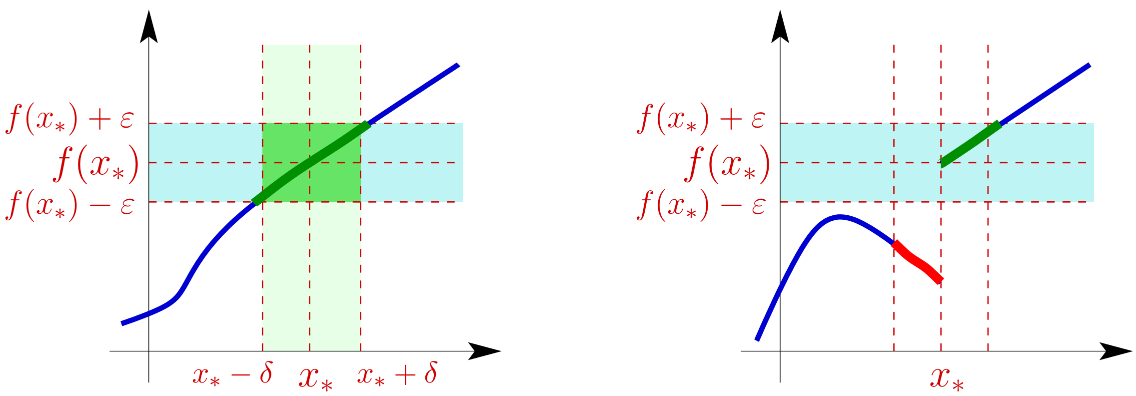


FIGURE 2.4 – *À gauche, une fonction continue en* x_* *: pour tout écart* $\varepsilon > 0$*, on peut trouver un petit intervalle* $]x_* - \delta, x_* + \delta[$ *autour de* x_* *dont l'image reste à distance moins de* ε *de* $f(x_*)$*. À droite, la fonction n'est pas continue en* x_* *: quand* $\varepsilon > 0$ *est suffisamment petit, il y a toujours des points aussi proches que l'on veut de* x_* *dont l'image est plus loin que* ε *de* $f(x_*)$*.*

On pourra utiliser à tout moment les caractérisations équivalentes suivantes.

Théorème 2.35 (critères équivalents pour la continuité)

Soit $f : \mathcal{D}_f \to \mathbb{R}$ une fonction réelle et $x_* \in \mathcal{D}_f$. Les propositions suivantes sont équivalentes

i) f est continue en x_*, i.e. pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $x \in \mathcal{D}_f$ vérifie $|x - x_*| < \delta$ alors $|f(x) - f(x_*)| < \varepsilon$,

ii) pour toute suite $(x_n) \subset \mathcal{D}_f$ qui tend vers x_*, $f(x_n)$ tend vers $f(x_*)$,

iii) les limites à gauche et à droite de f en x_* existent, sont finies et égales à la valeur de f en x_*, c'est-à-dire, si ces limites ont un sens,

$$\lim_{x \to x_*^-} f(x) = \lim_{x \to x_*^+} f(x) = f(x_*) .$$

Démonstration : Nous allons procéder par une boucle d'implications.

Commençons par supposer que i) est vérifiée. Soit une suite (x_n) tendant vers x_* et soit $\varepsilon > 0$. Par continuité, il existe $\delta > 0$ tel que si $x \in]x_* - \delta, x_* + \delta[$ alors $|f(x) - f(x_*)| < \varepsilon$. Comme (x_n) tend vers x_*, il existe un rang $N \in \mathbb{N}$ tel que $|x_n - x_*| < \delta$ pour $n \geq N$. Donc pour $n \geq N$, $|f(x_n) - f(x_*)| < \varepsilon$ et donc ii) est vérifiée.

Montrons que ii) implique iii) par contraposée. Imaginons que la limite à droite de f en x_* n'existe pas ou bien est différente de $f(x_*)$, c'est-à-dire que la phrase

$$\forall \varepsilon > 0 \ , \ \ \exists \delta > 0 \ , \ \ \forall x \in]x_*, x_* + \delta[\cap \mathcal{D}_f \ , \ \ |f(x) - f(x_*)| < \varepsilon$$

est fausse. On a donc qu'il existe $\varepsilon > 0$ tel que

$$\forall \delta > 0 \ , \ \ \exists x \in]x_*, x_* + \delta[\ , \ \ |f(x) - f(x_*)| \geq \varepsilon \tag{2.1}$$

En appliquant (2.1) à $\delta = 1$, on trouve un point x_1 dans $]x_*, x_* + 1[$ tel que $|f(x_1) - f(x_*)| \geq \varepsilon$. Puis en appliquant (2.1) à $\delta = 1/2$, on trouve un point x_2 dans $]x_*, x_* + 1/2[$ tel que $|f(x_2) - f(x_*)| \geq \varepsilon$ et on recommence ainsi : pour $\delta = 1/n$, on trouve un point x_n dans $]x_*, x_* + 1/n[$ tel que $|f(x_n) - f(x_*)| \geq \varepsilon$. On a ainsi une suite (x_n) qui tend vers x_* et telle que $f(x_n)$ reste à distance plus grande que $\varepsilon > 0$ de $f(x_*)$. Ceci contredit ii). La démonstration est similaire si le problème vient de la limite à gauche.

Supposons finalement que iii) est vraie. Pour tout $\varepsilon > 0$, d'après les définitions des limites à gauche et à droite, il existe δ_+ et $\delta_- > 0$ tels que pour tout $x \in]x_* - \delta_-, x_*[$ et pour tout $x \in]x_*, x_* + \delta_+[$, $|f(x) - f(x_*)| < \varepsilon$. On pose $\delta = \min(\delta_-, \delta_+)$, on a donc $|f(x) - f(x_*)| < \varepsilon$ pour tout $x \in]x_* - \delta, x_* + \delta[$ (le cas $x = x_*$ s'incluant de façon triviale). $\square$

Le théorème des valeurs intermédiaires correspond à l'idée simple de la continuité comme « le tracé sans lever le crayon ». Dans ce sens, il peut paraitre simpliste mais c'est un théorème fondamental qui est plus profond qu'il paraît.

Théorème 2.36 (Théorème des valeurs intermédiaires dit T.V.I.)

Soit $[a,b]$ un intervalle de $\mathbb{R}$ et $f \in \mathcal{C}^0([a,b],\mathbb{R})$ une fonction continue sur $[a,b]$. Soit y une valeur strictement comprise entre les images de a et b, c'est-à-dire que soit $f(a) < y < f(b)$, soit $f(b) < y < f(a)$. Alors, il existe $x \in]a,b[$ tel que $y = f(x)$.

Démonstration : La fonction $g : x \mapsto f(x) - y$ est aussi continue sur $[a,b]$. Le problème revient alors à trouver un point $x \in [a,b]$ où g s'annule en sachant que $g(a)$ et $g(b)$ sont de signes opposés.
On pose $u_0 = a$ et $v_0 = b$. Soit $m_0 = (a + b)/2$ le milieu du segment. Comme $g(a)$ et $g(b)$ sont de signes opposés, on a soit que $g(m_0)$ est du même signe que $g(u_0) = g(a)$ et on pose alors $u_1 = m_0$ et $v_1 = b$, soit $g(m_0)$ est du même signe que $g(b)$ est on pose alors $u_1 = a$ et $v_1 = m_0$. Puis on reprend $m_1 = (u_1 + v_1)/2$ le milieu du nouveau segment. Si $g(m_1) = 0$, on a trouvé notre x tel que $g(x) = 0$ et on peut s'arrêter. Si $g(m_1)$ est du même signe que $g(u_1)$ et on pose alors $u_2 = m_1$ et $v_2 = v_1$, si $g(m_1)$ est du même signe que $g(v_1)$ on pose alors $u_2 = u_1$ et $v_2 = m_1 \ldots$ On continue ainsi en coupant chaque segment en deux et en gardant le morceaux pour lequel les images des bords sont de signes opposés. Soit le processus s'arrête car on a trouvé un point x où g s'annule, soit il se poursuit infiniment. Mais dans ce dernier cas, cela nous construit deux suites (u_n) et (v_n) qui sont par construction adjacentes car (u_n) est croissante, (v_n) est décroissante et $|u_n - v_n|$ est la taille de l'intervalle à l'étape n qui vaut $2^{-n}(b - a)$. Donc (u_n) et (v_n) convergent vers un même limite x. Comme $a = u_0 \leq u_n \leq v_n \leq v_0 = b$, $x \in [a,b]$. Par ailleurs, $g(u_n)$ et $g(v_n)$ sont de signe opposés. Comme g est continue, $g(u_n)$ converge vers $g(x)$ et $g(x)$ est du même signe que $g(u_n)$ au sens large. Mais de même, $g(x)$ est du même signe que $g(v_n)$ au sens large. Le seul nombre que a les deux signes au sens large est $y = 0$. Donc $g(x) = y$ et on a trouvé le point cherché.
Il reste juste à remarquer que x n'est pas seulement dans $[a,b]$ mais en fait dans $]a,b[$. En effet, $g(a)$ et $g(b)$ sont supposés non nuls, donc x ne peut être ni a, ni b. $\square$

Exemples :

- Soit $P : x \mapsto ax^3 + bx^2 + cx + d$ un polynôme de degré 3, c'est-à-dire que $a \neq 0$. C'est une fonction continue sur $\mathbb{R}$. Supposons $a > 0$. Quand x tend vers $+\infty$, $P(x)$ tend vers $+\infty$ donc pour x assez grand $P(x) > 0$: il existe $b \in \mathbb{R}$ tel que $P(b) > 0$. Quand x tend vers $-\infty$, $P(x)$ tend vers $-\infty$ donc il existe a assez négatif pour que $P(a) < 0$. Le théorème des valeurs intermédiaires nous dit qu'il existe $x \in]a,b[$ tel que $P(x) = 0$. Le cas $a < 0$ est symétrique. On obtient donc le résultat que tout polynôme réel de degré 3 admet au moins une racine réelle. Notons qu'il existe des polynômes de degré 2 sans racines réelles (comme $P(x) = x^2 + 1$).

- Soit $d(t)$ la distance d'un solide à un point de référence, il est naturel de considérer $d(t)$ comme une fonction continue du temps. Si à $t = 0$ le solide était sur le point de référence et si à $t = T > 0$, il était à distance $d(T) = 100\,\text{m}$, alors à un moment entre 0 et T, il a été à distance $d(t) = 10\,\text{m}$.

- Un récipient contient une quantité de liquide que l'on vide progressivement dans un autre récipient qui était vide au départ. Il existe un moment où les deux récipients contiennent exactement le même volume de liquide. En effet, si $V(t)$ est la quantité de liquide dans le récipient d'origine, alors on a au départ $V(0) > 0$ et à la fin $V(T) = 0$. Comme $V(t)$ est naturellement une quantité physique continue, il existe un temps $t \in]0,T[$ tel que $V(t) = V(0)/2$.

- La fonction f définie sur $\mathbb{R}$ par $f(x) = 0$ si $x < 0$ et $f(x) = 1$ si $x \geq 0$ n'est pas continue en $x = 0$. Les valeurs entre $f(-1) = 0$ et $f(1) = 1$ ne sont pas prises par la fonction. Celle-ci ne vérifie pas la propriété des valeurs intermédiaires.

- Une bille de vitesse $V > 0$ subit un choc élastique contre un mur et rebondit en repartant à vitesse $-V < 0$. Pourtant la bille n'a jamais été au repos car son énergie cinétique étant conservée, elle vaut toujours $\frac{1}{2}mV^2 \neq 0$. C'est parce que dans cette modélisation, la vitesse passe brutalement de V à $-V$: elle est discontinue et ne vérifie pas forcément le T.V.I.

Le deuxième résultat théorique important concernant la continuité est lié à ce qu'on appelle la compacité. Il permet non seulement de borner une fonction mais il garantit l'existence d'extrema. On lui associe parfois le nom de Karl Weierstrass (1815-1897, Allemagne).

Théorème 2.37 (théorème des valeurs extrêmes)

Soit $[a,b] \subset \mathbb{R}$ un intervalle fermé et borné et $f \in \mathcal{C}^0([a,b],\mathbb{R})$ une fonction réelle continue sur $[a,b]$. Alors f est bornée et atteint ses bornes sur $[a,b]$. Autrement dit, il existe $x_{\max}$ et $x_{\min}$ tels que

$$f(x_{\max}) = \max_{x\in[a,b]} f(x) = \sup_{x\in[a,b]} f(x) \quad \text{et} \quad f(x_{\min}) = \min_{x\in[a,b]} f(x) = \inf_{x\in[a,b]} f(x)\ .$$

Démonstration : On va montrer que f est majorée sur $[a,b]$ et atteint son maximum. Le cas du minimum est symétrique.
Supposons que f ne soit pas majorée sur $[a,b]$, alors par définition, pour tout $n \in \mathbb{N}$, il existe un point $x_n \in [a,b]$ tel que $f(x_n) \geq n$ et en particulier $f(x_n) \to +\infty$. D'après le théorème de Bolzano-Weierstrass (corollaire 2.26), on peut extraire de (x_n) une sous-suite $(x_{\varphi(n)})$ qui converge vers une limite $\ell \in [a,b]$. On a que $f(x_{\varphi(n)})$ tend vers $+\infty$ car c'est une sous-suite de $f(x_n)$ mais aussi que $f(x_{\varphi(n)})$ tend vers $f(\ell)$ par continuité de f. Comme $f(\ell)$ est un nombre fini, c'est contradictoire et donc f est majorée sur $[a,b]$.
On pose $M = \sup_{x\in[a,b]} f(x)$. Pour tout $n \in \mathbb{N}$, $M - 2^{-n}$ n'est plus un majorant et il existe donc $x_n \in [a,b]$ tel que $M - 2^{-n} < f(x_n) \leq M$. Donc $f(x_n)$ tend vers M. Mais comme précédemment, on peut extraire de (x_n) une sous-suite

$(x_{\varphi(n)})$ qui converge vers une limite $x_{\max} \in [a,b]$. Et par continuité $f(x_{\varphi(n)})$ tend vers $f(x_{\max})$. Donc $f(x_{\max}) = \sup_{x\in[a,b]} f(x)$ est le maximum cherché. $\square$

! Comme on le voit dans les exemples ci-dessous, il est important que f soit continue mais aussi que $[a,b] \subset \mathbb{R}$ soit un intervalle fermé et borné, c'est-à-dire qu'il inclut ses bornes a et b qui sont des réels (finis). On fera plus tard la généralisation de ce résultat. Le point important est que $[a,b]$ est un *compact* qui vérifie le corollaire 2.26.

Exemples :

- Soit $d(t)$ la distance d'un solide à un point de référence, il est naturel de considérer $d(t)$ comme une fonction continue du temps. Pendant un intervalle de temps $[0,T]$, le solide s'est éloigné au maximum d'une distance D et il existe un temps $t_0 \in [0,T]$ où il était pile à distance D.
- La fonction $f : x \mapsto x$ est continue sur $[0,+\infty[$ mais n'est pas majorée dessus. On ne peut pas appliquer le théorème précédent car $[0,+\infty[$ n'est pas un intervalle borné.
- La fonction $f : x \mapsto 1/x$ est continue sur $]0,1]$ mais n'est pas majorée dessus. On ne peut pas appliquer le théorème précédent car $]0,1]$ n'est pas un intervalle fermé.
- La fonction f définie sur $[0,1]$ par $f(0) = 0$ et $f(x) = 1/x$ si $x \in]0,1]$ n'est pas majorée sur $[0,1]$. Même si $[0,1]$ est fermé et borné, on ne peut pas appliquer le théorème précédent car f n'est pas continue.

En rassemblant les deux énoncés de cette partie, on obtient ce qu'on pourra appeler dans pas très longtemps « l'image d'un connexe compact par une fonction continue est un connexe compact ».

Corollaire 2.38

L'image d'un intervalle $[a,b] \subset \mathbb{R}$ par une fonction réelle continue est un intervalle $[\alpha,\beta] \subset \mathbb{R}$.

Démonstration : Le théorème 2.37 nous dit que l'image par f continue d'un intervalle $[a,b]$ est bornée et que les bornes sont atteintes. Donc $\alpha = \min_{[a,b]} f$ et $\beta = \max_{[a,b]} f$ sont dans l'image de f, atteints aux points $x_{\min}$ et $x_{\max}$ respectivement. Par définition de ces extrema, l'image de f est incluse dans $[\alpha,\beta]$. Mais le théorème des valeurs intermédiaires appliqué sur l'intervalle $[x_{\min},x_{\max}]$ (ou $[x_{\max},x_{\min}]$) nous dit que toutes les valeurs de $[\alpha,\beta]$ sont atteintes. $\square$

Finissons ce chapitre par introduire l'uniforme continuité dans le cas de $\mathbb{R}$.

Définition 2.39

Une fonction $f : \mathcal{D}_f \to \mathbb{R}$ est dite **uniformément continue** si

$$\forall \varepsilon > 0 \ , \ \exists \delta > 0 \ , \ \forall x,y \in \mathcal{D}_f \ , \ |x-y| < \delta \implies |f(x) - f(y)| < \varepsilon \ .$$

! Si on écrit la continuité de f sous la même forme, on écrirait

$$\forall \varepsilon > 0 \ , \ \forall x \in \mathcal{D}_f \ , \ \exists \delta > 0 \ , \ \forall y \in \mathcal{D}_f \ , \ |x-y| < \delta \implies |f(x)-f(y)| < \varepsilon \ .$$

On voit alors que l'uniforme continuité est strictement plus forte que la continuité : si la fonction est continue, δ est choisi après x et peut dépendre de x ; alors que si la fonction est uniformément continue, δ est choisi avant x et est donc le même pour tout x, i.e. il est uniforme en x.

Le résultat le plus important concernant l'uniforme continuité dans $\mathbb{R}$ est le suivant. Son nom vient du mathématicien Eduard Heine (1821-1881, Allemagne). De nouveau, la notion importante cachée derrière est celle de *compacité* et c'est pour cela qu'il est important de considérer un intervalle $[a,b]$ fermé et borné.

Théorème 2.40 (théorème de Heine)

Soit $a < b$ et $f : [a,b] \to \mathbb{R}$ une fonction continue. Alors f est uniformément continue sur $[a,b]$.

Démonstration : Nous raisonnons par l'absurde. Si f n'est pas uniformément continue, c'est qu'il existe $\varepsilon > 0$ tel que pour tout $\delta = 1/n$, il existe x_n et y_n dans $[a,b]$ tels que $|x_n - y_n| < 1/n$ et $|f(x_n) - f(y_n)| \geq \varepsilon$. Par le théorème de Bolzano-Weierstrass (corollaire 2.26), on peut extraire une sous-suite $(x_{\varphi(n)})$ telle que $x_{\varphi(n)} \to x_* \in [a,b]$. Comme $|x_n - y_n| < 1/n$, on doit aussi avoir $y_{\varphi(n)} \to x_*$. Par continuité, il faut que $f(x_{\varphi(n)})$ et $f(y_{\varphi(n)})$ aient la même limite $f(x_*)$. C'est-à-dire qu'on devrait avoir $|f(x_{\varphi(n)}) - f(y_{\varphi(n)})| \to 0$, ce que contredirait l'inégalité $|f(x_n) - f(y_n)| \geq \varepsilon$. □

Exemples :

- Le résultat précédent nous donne beaucoup d'exemples : toutes les fonctions usuelles sur un intervalle $[a,b]$ sont uniformément continue.
- La fonction $x \in [0,+\infty[\mapsto 2x$ est uniformément continue. En effet, pour tout $\varepsilon > 0$, on peut prendre $\delta = \varepsilon/2$.
- La fonction $x \in [0,+\infty[\mapsto x^2$ n'est uniformément continue. En effet, pour $\varepsilon = 1$ et pour tout $\delta > 0$, on peut prendre $x = \frac{2}{\delta}$ et $y = x + \frac{\delta}{2}$. On a alors
$$|x^2 - y^2| \ = \ \frac{\delta}{4}(4x + \delta) \ > \ \delta x \ = \ 2 \ > \ \varepsilon.$$

Ceci contredit la définition de l'uniforme continuité. Visuellement, on voit que la distance séparant deux images de points proches x et $x+\delta$ est de plus grande quand x va vers $+\infty$. C'est donc que l'uniforme continuité n'est pas possible.

- La fonction $x \in]0,1] \mapsto \frac{1}{x}$ n'est uniformément continue. En effet, pour $\varepsilon = 1/2$ et pour tout $\delta > 0$, on pose $\alpha = \min(\delta,1/2)$ et on peut prendre $x = \alpha$ et $y = 2\alpha$. On a alors
$$|\frac{1}{x} - \frac{1}{y}| \;=\; \frac{1}{2\alpha} \;\geq\; 1 \;>\; \varepsilon.$$
Ceci contredit la définition de l'uniforme continuité. Visuellement, on voit que la distance séparant deux images de points proches x et $x+\delta$ est de plus grande quand x tend vers 0. C'est donc que l'uniforme continuité n'est pas possible.

L'uniforme continuité peut paraître une définition technique, mais elle est centrale pour beaucoup de preuves de propriétés qui sont fausses si la fonction est seulement continue.

Exemple :

Si $f :]0,1] \to \mathbb{R}$ est définie par $f(x) = \cos(1/x)$, il s'agit d'une fonction continue et bornée qui n'a pas de limite quand $x \to 0$. En effet, $f(1/2n\pi) \to 1$ et $f(1/(2n+1)\pi) \to -1 \neq 1$. Notons qu'elle n'est pas uniformément continue (ce qui pourra se démontrer par la contraposée de ce qui suit).
Mais si on prend $f :]0,1] \to \mathbb{R}$ quelconque qui est uniformément continue sur $]0,1]$ alors on va montrer que f admet une limite quand $x \to 0$. On pose $u_n = f(1/n)$ et on prétend que (u_n) est de Cauchy. En effet, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que $|f(x) - f(y)| < \varepsilon$ si $|x - y| < \delta$. Donc si p et q sont assez grands, par exemple si $p,q \geq 1/\delta$, alors $|1/p - 1/q| < \delta$ et donc $|u_p - u_q| = |f(1/p) - f(1/q)| < \varepsilon$. Par le théorème 2.14, la suite (u_n) converge vers une limite $\ell \in \mathbb{R}$. Montrons maintenant que cette limite est bien celle de toute la fonction (nous n'avons que le cas $f(1/n)$ pour le moment). Soit $\varepsilon > 0$. On applique l'uniforme continuité : il existe $\delta > 0$ tel que si $|x - y| < \delta$ alors $|f(x) - f(y)| < \varepsilon/2$. On a aussi un rang n_0 tel que si $n \geq n_0$, $|\ell - f(1/n)| < \varepsilon/2$. Soit $x \in]0,\delta[$ et soit $n \geq \max(n_0,1/\delta + 1)$. On a alors $1/n \in]0,\delta[$ et donc $|x - 1/n| < \delta$. On obtient au final
$$|\ell - f(x)| \leq |\ell - f(1/n)| + |f(1/n) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \;.$$
Ceci montre bien que $f(x) \to \ell$ quand $x \to 0$.

5 Quelques applications

5.1 La méthode de Héron

Il y a plus de 3.000 ans, les savants de l'Antiquité mésopotamienne et égyptienne savaient extraire des racines carrées. On trouve ainsi des tablettes d'argile avec la valeur de la diagonale d'un carré montrant qu'ils savaient calculer $\sqrt{2}$ à la précision du millionième. Les scribes n'ont pas expliqué leur méthode, mais on pense qu'il s'agit du même algorithme que celui utilisé par les grecs plusieurs siècles plus tard. Il est appelé *méthode de Héron* ou *méthode babylonienne*. L'idée est géométriquement simple. Supposons que l'on veuille calculer $\sqrt{a}$, cela revient à trouver quel est le côté d'un carré de surface a. On part d'un rectangle de surface a, par exemple de côtés $u_0 = a$ et $v_0 = 1$ (en pratique, il est plus rapide de partir d'une meilleure approximation du bon carré, par exemple $u_0 = 4$ et $v_0 = 2{,}5$ pour calculer $\sqrt{10}$). Ce premier rectangle a la bonne surface mais n'est pas un carré. Pour le rendre « plus carré », on prend comme nouveau côté la moyenne des côtés en posant $u_1 = (u_0+v_0)/2$ et donc $v_1 = a/u_1$. Notre nouveau rectangle est encore d'aire a mais il est proche du carré. On itère ainsi le procédé et on se rapproche de plus en plus de la racine carrée cherchée.

Un calcul de $\sqrt{2}$ *datant de plus de 3.500 ans*

Mathématiquement parlant, comme $v_n = a/u_n$, cela revient à partir de $u_0 = a$ et itérer la récurrence

$$u_{n+1} = \frac{1}{2}\left(u_n + \frac{a}{u_n}\right) . \tag{2.2}$$

L'étude de la fonction $f : x \mapsto (x+a/x)/2$ montre qu'elle est décroissante sur $]0,\sqrt{a}]$ et croissante sur $[\sqrt{a},+\infty[$ et que $f(\sqrt{a}) = \sqrt{a}$. En particulier, f envoie $[\sqrt{a},+\infty[$ sur lui-même et pour tout n, $u_n \geq \sqrt{a}$. On a alors que $v_n = a/u_n \leq \sqrt{a}$ et donc que $u_{n+1} = (u_n + v_n)/2 \leq (u_n + u_n)/2 = u_n$. La suite (u_n) est donc décroissante et minorée par $\sqrt{a}$ et elle converge vers un réel $\ell \geq \sqrt{a}$. En passant à la limite dans (2.2), on obtient que $\ell = (\ell + a/\ell)/2$ ce qui donne $\ell^2 = a$. Comme ℓ est positive, on a donc $\ell = \sqrt{a}$.

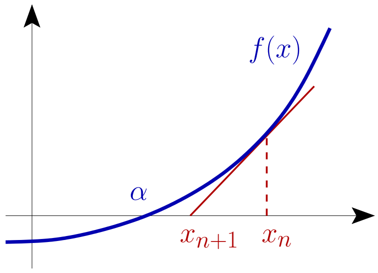
Pour obtenir une bonne approximation de $\sqrt{a}$, il suffit donc de partir d'une valeur raisonnable et d'itérer la formule (2.2) qui n'est composée que d'opérations élémentaires. On peut prouver que cette méthode est très efficace car elle converge quadratiquement : à chaque étape, on double le nombre de chiffres exacts de notre approximation.

5.2 La méthode de dichotomie

Si on regarde bien la preuve du théorème des valeurs intermédiaires, elle fournit une méthode simple, constructive et explicite pour chercher les zéros d'une fonction continue, c'est-à-dire pour trouver une solution à une équation $f(x) = 0$. Supposons que f soit continue sur un intervalle $[a,b]$ et que l'on sache que $f(a)$ et $f(b)$ sont de signes opposés. Alors on coupe l'intervalle $[a,b]$ en deux et on garde l'intervalle où f admet des valeurs de signes opposés au bord. Puis on recoupe cette intervalle en deux etc. On sait par le T.V.I. qu'il existe bien une solution de $f(x) = 0$ dans chacun des intervalles et plus on avance, plus ces intervalles sont petits et plus on obtient une bonne approximation de cette solution. Attention toutefois que cette méthode ne garantit pas de trouver toutes les solutions même si elle permet d'en trouver au moins une. L'algorithme est schématiquement :

```
Tant que (a-b)/2 > précision faire boucle
       pose m=(a+b)/2
       si f(a)f(m)<0
             alors pose b=m
             sinon pose a=m
       fin si
fin boucle
Écrit “la solution vaut” (a+b)/2 “avec la précision ±” précision
```

5.3 L'algorithme de Newton



La méthode de Newton est utilisée pour approcher les zéros de fonctions. Considérons une fonction f dont on veut trouver un zéro α, c'est-à-dire une solution de l'équation $f(\alpha) = 0$. L'idée géométrique est de partir d'une valeur assez proche x_0, puis d'approcher la fonction f par sa tangente en x_0 et de prendre comme nouvelle position x_1 le zéro de cette tangente. Puis on recommence le procédé et on espère que la suite (x_n) tende bien vers α. Ce n'est pas toujours le cas et l'algorithme peut échouer, ne serait-ce que si la tangente que l'on regarde est horizontale et donc n'a pas de zéro. La méthode a été introduite par Isaac Newton en 1669 mais celui-ci ne l'a utilisée que pour les polynômes. Il a fallu un peu de temps avant que l'on comprenne la généralité de l'algorithme et qu'il prenne sa forme moderne.

Concrètement, si f est dérivable, sa tangente au point x_n est la droite d'équation $y = f'(x_n)(x - x_n) + f(x_n)$. Comme x_{n+1} est le zéro de cette tangente, on a

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} . \qquad (2.3)$$

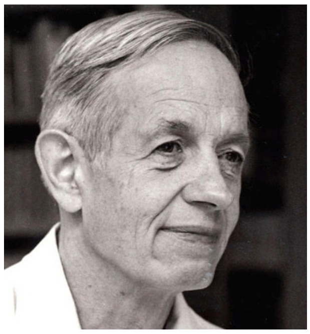
Supposons que sur l'intervalle $[\alpha,x_0]$, la fonction f soit croissante. Elle est donc positive et f' est positive aussi. On obtient alors que $x_{n+1} \leq x_n$. Supposons en outre que f soit convexe sur $[\alpha,x_0]$, elle est alors au-dessus de ses tangentes et $\alpha \leq x_{n+1}$ (voir la figure ci-dessus). Sous ces hypothèses, la suite (x_n) est donc décroissante et minorée par α et elle converge vers une limite $\ell \in [\alpha,x_0]$. En passant à la limite dans (2.3), on a que $\ell = \ell - f(\ell)/f'(\ell)$ et donc que $f(\ell) = 0$. Comme α est le seul zéro de f dans l'intervalle considéré, $\ell = \alpha$. Dans ce cas, la suite (x_n) partant de x_0 et définie par la récurrence (2.3) converge bien vers un zéro f. Dans d'autre cas, la convergence n'a pas forcément lieu mais on peut montrer qu'elle est très rapide si elle a lieu.

Une application simple de l'étude précédente est le cas du plus grand zéro des polynômes scindés. Prenons par exemple $f(x) = x^2 - a$ avec $a \in \mathbb{R}$. La fonction f est croissante pour $x \geq 0$ et convexe sur $\mathbb{R}$. Si on part d'un point $x_0 \geq \sqrt{a}$ et qu'on itère la méthode de Newton, la suite (x_n) va donc converger vers $\sqrt{a}$. Dans le cas $f(x) = x^2 - a$, (2.3) devient

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} = x_n - \frac{x_n}{2} + \frac{a}{2x_n} = \frac{1}{2}\Big(x_n + \frac{a}{x_n}\Big) .$$

On retrouve la méthode de Héron !

5.4 Un théorème de point fixe



John Nash
1928-2015
États-Unis

Les points fixes d'une fonction f, c'est-à-dire les solutions de l'équation $f(x) = x$, jouent un rôle important dans beaucoup de théories et applications. Par exemple, leur interprétation en tant qu'équilibres d'une dynamique $u_{n+1} = f(u_n)$ se retrouve en théorie des jeux et donc en économie. En se limitant aux fonctions à une seule variable réelle, cela revient souvent à prendre des modèles simplistes, mais essayons d'en comprendre le principe général. Imaginons un agent économique qui peut produire une quantité x de biens et en tirer un profit $h(x,y)$ qui dépend de sa production x et de l'état du marché y. Il va tenter de maximiser ses gains et donc produire la quantité $x_* = g(y)$ telle que $h(x_*,y) = \max_x h(x,y)$. Mais l'état du marché y dépend aussi de la production x de l'agent (une surproduction peut baisser les prix etc.). Si l'agent produit une quantité x, que le marché est dans $y(x)$, que son gain est $h(x,y(x))$ mais que produire la quantité $x_* = g(y(x))$ lui apporterait un meilleur gain, il sera insatisfait et changera sa production en x_*. Mais cela va changer l'état $y(x)$ du marché en $y(x_*)$ et la production optimale ne sera pas forcément toujours x_* etc. L'idée est que si le marché est stable, c'est qu'on a fini dans un état tel que $x_* = g(y(x_*))$. On est amené à chercher un point fixe de la fonction $f : x_* \mapsto g(y(x_*))$. C'est un *équilibre de Nash* du nom du mathématicien John Nash, lauréat du prix dit Nobel d'économie et du prix Abel.

Théorème 2.41

Soit $[a,b] \subset \mathbb{R}$ et $f \in \mathcal{C}^0([a,b],[a,b])$ une fonction continue envoyant $[a,b]$ sur lui-même. Alors f a au moins un point fixe, c'est-à-dire qu'il existe $x_* \in [a,b]$ tel que $f(x_*) = x_*$.

Démonstration : Si $f(a) = a$, c'est gagné, donc supposons que l'on a $f(a) \neq a$, ce qui implique $f(a) > a$ car $f(a) \in [a,b]$. De même, si $f(b) = b$, c'est gagné, donc supposons que l'on a $f(b) \neq b$, ce qui implique $f(b) < b$. On considère la fonction $g : x \mapsto f(x) - x$. La fonction g est continue sur $[a,b]$ comme somme de fonctions continues. Par ailleurs, les hypothèses ci-dessus nous donnent que $g(a) = f(a) - a > 0$ et $g(b) = f(b) - b < 0$. En appliquant le théorème des valeurs intermédiaires, on voit qu'il existe $x_* \in]a,b[$ tel que $g(x_*) = 0$. Mais cela veut dire que $f(x_*) - x_* = 0$ et donc x_* est un point fixe de f. $\square$

5.5 Le minimum d'un puits de potentiel

On considère un potentiel continu $V : \mathbb{R} \to \mathbb{R}$. On suppose que $\lim_{x\to\pm\infty} V(x) = +\infty$. Cela veut dire que ce potentiel est un *potentiel puits* qui a tendance à ramener notre système vers une zone bornée. Par les résultats ci-dessus, nous pouvons

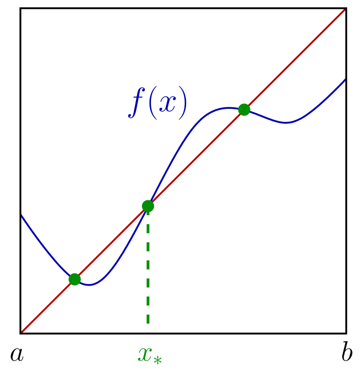


FIGURE 2.5 – *Une illustration du théorème 3.41 : le graphe de la fonction continue f doit forcément intersecter la droite $y = x$ et donc f a forcément au moins un point fixe (ici elle en a trois).*

montrer que V admet un minimum global, c'est-à-dire un point x_* où l'énergie potentielle est la plus petite possible et donc pour lequel, un système dans cet état x_* resterait stable proche de cette énergie minimale. En effet, soit $M = V(0) + 1$. Comme $\lim_{x\to-\infty} V(x) = +\infty$, il existe x_- tel que si $x \le x_-$ alors $V(x) \ge M$. De même, comme $\lim_{x\to+\infty} V(x) = +\infty$, il existe x_+ tel que si $x \ge x_+$ alors $V(x) \ge M$. Donc pour tout x en dehors de $[x_-,x_+]$, $V(x)$ est minoré par $V(0) + 1$, et en particulier $0 \in [x_-,x_+]$ puisque $V(0) < V(0) + 1$. Par ailleurs, comme V est continue sur l'intervalle fermé borné $[x_-,x_+]$, V y est minoré et atteint son minimum en un point x_*. On a donc que pour tout $x \in [x_-,x_+]$, $V(x) \ge V(x_*)$. Mais comme $0 \in [x_-,x_+]$, pour tout $x \notin [x_-,x_+]$, $V(x) \ge V(0) + 1 > V(0) \ge V(x_*)$. Au total, on a bien que $V(x) \ge V(x_*)$ pour tout $x \in \mathbb{R}$ et V admet un minimum global en x_*.

5.6 Le théorème de Weierstrass

Cette dernière application concerne en fait un théorème important de l'analyse.

Théorème 2.42 (théorème de Weierstrass et polynômes de Bernstein)

Soient $a < b$. Pour toute fonction continue $f \in \mathcal{C}^0([a,b],\mathbb{R})$, il existe une suite de polynômes $(P_n) \subset \mathbb{R}[X]$ qui converge uniformément vers f sur $[a,b]$, c'est-à-dire que

$$\sup_{x\in[a,b]} \left|f(x) - P_n(x)\right| \xrightarrow[n\longrightarrow+\infty]{} 0 \,. \tag{2.4}$$

Plus précisément, si $[a,b] = [0,1]$, la suite

$$P_n = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} X^k (1-X)^{n-k} \tag{2.5}$$

satisfait la convergence uniforme demandée.

Karl Weierstrass
1815-1897
Allemagne

Sergueï Bernstein
1880-1968
Russie

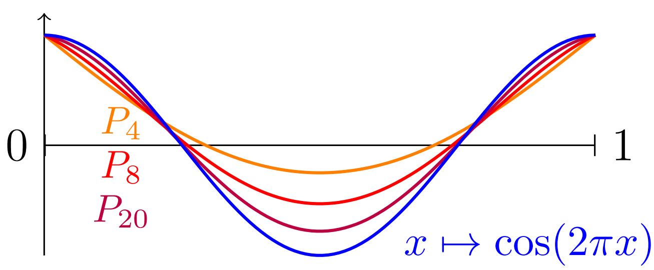
Pour comprendre l'importance du théorème de Weierstrass, il faut se rappeler que les polynômes et leur quotient sont les seuls fonctions que l'on peut calculer de façon exacte. Des opérations comme la racine carré ou le log ne sont en fait implémentées que de façon approchée. Par ailleurs, il est facile de coder et stocker un polynôme car il s'agit juste d'une suite de coefficients. Le théorème de Weierstrass nous dit qu'avec ces polynômes, on peut faire toutes les formes possibles (au moins de façon approchée). Ainsi, on peut utiliser les polynômes pour coder tous les dessins sur un ordinateur et la police de caractères de ce poly est elle-même codée à l'aide de polynômes.

Il existe plusieurs suites (P_n) satisfaisant (2.4). La suite (2.5) est due à Bernstein. La forme (2.5) est valable pour le segment $[0,1]$ mais la transformation $x \mapsto (x-a)/(b-a)$ permet de passer de $[a,b]$ à $[0,1]$ et donc de transposer cette formule à tout segment $[a,b]$.

Les polynômes de Taylor nous fournissent déjà des approximations de fonctions par développement limité. Mais il ne s'agit que d'une approximation proche d'un point et qui peut rester très mauvaise ailleurs, même quand on va de plus en plus loin dans le développement. L'interpolation de Lagrange fournit une interpolation polynomiale qui passe par un certain nombre de points. Mais là encore, l'approximation peut rester mauvaise ailleurs, même si on augmente le nombre de points.

Exemples :

- Les polynômes de Bernstein (2.5) fournissent une suite explicite. On peut faire le test ci-dessous avec la fonction $x \in [0,1] \longmapsto \sin(2\pi x)$. Les premières expressions sont simples, par exemples $P_4 = -4X^4 + 8X^3 - 4X + 1$ mais on obtient rapidement des polynômes peu pratiques mêmes si explicites. Par ailleurs, la convergence semble très lente.



Il se trouve que les polynômes de Bernstein ne sont pas les meilleurs pour faire des bonnes approximations, même s'ils suffisent à obtenir le théorème comme nous allons le voir dans la preuve. En pratique, on utilise d'autres approximations polynomiales plus adéquates.

- On rappelle que la limite uniforme d'une suite de fonctions continues est continue. Le théorème de Weierstrass échoue donc si f n'est pas continue.

- Il est important de considérer un intervalle *compact*, c'est-à-dire $[a,b]$ fermé et borné. Ainsi, on ne peut pas approcher uniformément toutes les fonctions par des polynômes sur $\mathbb{R}$. Par exemple, la fonction $x \in \mathbb{R} \mapsto \sin x$ est bornée mais pas constante. Elle n'est pas bien approchée par les constantes mais dès qu'on prend un polynôme non constant, celui-ci a des limites infinies en $\pm\infty$ et donc l'erreur dans (2.4) n'est même pas majorée sur $\mathbb{R}$ et $\sup_x |f(x) - P_n(x)| = +\infty$. De même, si on considère $x \in [0, +\infty[\mapsto e^x$, la croissance en $+\infty$ de l'exponentielle est trop forte pour pouvoir être approchée par des polynômes. Enfin, le même problème arrive si on considère un intervalle borné mais non fermé. Ainsi $x \in]0,1] \mapsto 1/x$ n'est pas bornée alors que tout polynôme est borné sur $]0,1]$. Donc la distance entre $1/x$ et un polynôme ne pourra pas être petite uniformément sur tout $]0,1]$.

Bernstein s'est appuyé sur les probabilités pour trouver la formule (2.5) qui a fourni une nouvelle preuve du théorème de Weierstrass. Imaginons une expérience aléatoire qui a une chance $p \in [0,1]$ de réussir. La variable aléatoire X comptant le nombre de succès de n expériences indépendantes suit la loi binomiale de paramètre p définie par

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} .$$

Il est connu (cf n'importe quel cours de probabilités de niveau L2) que

$$E(X) = np \quad \text{et } \mathrm{Var}(X) = np(1-p) .$$

On en déduit les résultats suivants.

Lemme 2.43

Pour tout $x \in \mathbb{R}$ et $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1 .$$

Démonstration : Il s'agit simplement de la formule du binôme $\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = \big(x + (1-x)\big)^n = 1$. Mais on peut aussi la voir comme le fait que la loi binomiale est une loi de probabilité et donc $\sum_{k=0}^{n} P(X = k) = 1$ pour $p = x \in [0,1]$. Comme les deux membres de l'égalité sont des polynômes en x, s'ils sont égaux sur $[0,1]$, ils le sont sur tout $\mathbb{R}$. □

Lemme 2.44

Pour tout $x \in \mathbb{R}$ et $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \left(x - \frac{k}{n}\right)^2 \binom{n}{k} x^k (1-x)^{n-k} = \frac{x(1-x)}{n} .$$

Démonstration : Comme ci-dessus, il suffit de montrer l'égalité pour tout $x = p \in [0,1]$. Notre connaissance de la variance de la loi binomiale nous dit que

$$\begin{aligned} np(1-p) = \mathrm{Var}(X) &= \sum_{k=0}^{n} P(X=k)(k - E(X))^2 \\ &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \big(k - np\big)^2 \end{aligned}$$

On obtient la formule cherchée en divisant par n^2 des deux côtés. $\square$

Nous allons utiliser ces lemmes dans une estimation astucieuse grâce à la continuité uniforme.

Démonstration du théorème 2.42 : Comme dit plus haut, la transformation $x \mapsto (x-a)/(b-a)$ permet de passer de $[a,b]$ à $[0,1]$ et elle transforme les polynômes en polynômes. Il suffit donc de montrer que la suite P_n donnée par (2.5) vérifie (2.4) dans le segment $[a,b] = [0,1]$.
Soit $f \in \mathcal{C}^0([0,1],\mathbb{R})$ et soit $\varepsilon > 0$, on doit montrer que pour n assez grand $\sup_x |f(x) - P_n(x)| < \varepsilon$. On commence par utiliser le lemme 2.43 pour écrire

$$f(x) \ - \ \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big) \binom{n}{k} x^k (1-x)^{n-k} \ = \ \sum_{k=0}^{n} \Big(f(x) - f\Big(\frac{k}{n}\Big)\Big) \binom{n}{k} x^k (1-x)^{n-k} .$$

Nous allons maintenant utiliser le théorème de Heine (théorème 2.40) : la fonction f est uniformément continue. Donc il existe $\delta > 0$ tel que pour tout x,y avec $|x-y| < \delta$, on a $|f(x) - f(y)| < \varepsilon/2$. Par ailleurs, le théorème 2.37 nous assure qu'il existe une borne $M > 0$ telle que $\sup_{x\in[0,1]} |f(x)| \leq M$.
Pour tout $n \in \mathbb{N}$ et $x \in [0,1]$, on introduit l'ensemble $K_{n,x,\varepsilon}$ des indices $k \in [\![0,1]\!]$ tels que $|x - k/n| < \delta$. On découpe alors la somme en deux en utilisant que $x \in [0,1]$

$$\begin{aligned} \left| f(x) \ - \ \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big) \binom{n}{k} x^k (1-x)^{n-k} \right| &\leq \sum_{k=0}^{n} \Big| f(x) - f\Big(\frac{k}{n}\Big)\Big| \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \sum_{k \in K_{n,x,\varepsilon}} \Big| f(x) - f\Big(\frac{k}{n}\Big)\Big| \binom{n}{k} x^k (1-x)^{n-k} \\ &\quad + \sum_{k \notin K_{n,x,\varepsilon}} \Big| f(x) - f\Big(\frac{k}{n}\Big)\Big| \binom{n}{k} x^k (1-x)^{n-k} \end{aligned}$$

Pour les indices dans $K_{n,x,\varepsilon}$, on utilise simplement le lemme 2.43 et la définition de cet ensemble : l'erreur entre $f(x)$ et $f(k/n)$ y est petite et donc (toujours en utilisant la positivité de x et $1-x$)

$$\sum_{k\in K_{n,x,\varepsilon}} \left|f(x) - f\left(\frac{k}{n}\right)\right| \binom{n}{k} x^k(1-x)^{n-k} \leq \frac{\varepsilon}{2} \sum_{k\in K_{n,x,\varepsilon}} \binom{n}{k} x^k(1-x)^{n-k} \leq \frac{\varepsilon}{2}$$

L'autre partie suit l'idée de l'inégalité de Bienaymé-Tchebychev. On se débarasse grossièrement de f par le théorème 2.37 qui nous assure qu'il existe une borne $M > 0$ telle que $\sup_{x\in[0,1]} |f(x)| \leq M$. Puis si k/n est à distance plus que δ de x, alors $|x - k/n|^2 > \delta^2$. Alors le lemme 2.44 permet d'obtenir que

$$\begin{aligned}
\sum_{k\notin K_{n,x,\varepsilon}} \left|f(x) - f\left(\frac{k}{n}\right)\right| \binom{n}{k} x^k(1-x)^{n-k} &\leq 2M \sum_{k\notin K_{n,x,\varepsilon}} \binom{n}{k} x^k(1-x)^{n-k} \\
&\leq \frac{2M}{\delta^2} \sum_{k\notin K_{n,x,\varepsilon}} |x-k/n|^2 \binom{n}{k} x^k(1-x)^{n-k} \\
&\leq \frac{2M}{\delta^2} \sum_{k=0}^{n} |x-k/n|^2 \binom{n}{k} x^k(1-x)^{n-k} \\
&\leq \frac{2M}{n\delta^2} x(1-x) \leq \frac{M}{2n\delta^2}
\end{aligned}$$

En regroupant les deux estimations précédentes, on obtient pour tout $\varepsilon > 0$, si $\delta > 0$ est le module de continuité uniforme défini plus haut et M une borne uniforme sur $|f|$, alors pour tout $x \in [0,1]$,

$$\left| f(x) \ - \ \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k(1-x)^{n-k} \right| \leq \frac{\varepsilon}{2} + \frac{M}{2n\delta^2} \ .$$

On note que cette estimation est indépendente de x. Il reste maintenant à prendre n assez grand tel que $M/n\delta^2 < \varepsilon$ et on obtient

$$\sup_{x\in[0,1]} \left| f(x) \ - \ \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k(1-x)^{n-k} \right| < \varepsilon$$

ce qui conclut. $\square$