Notes de cours d'Analyse réelle I

Olivier Collin

Introduction

Si vous qui étudiez en Mathématiques vous premenez au hasard dans la rue et interpellez un passant pour lui demander ce que fait un mathématicien dans la vie, il est plus que probable que l'on vous réponde - après la classique introduction *"Vous savez, j'ai toujours été nul en Mathématiques... en fait j'ai toujours haï les Mathématiques !"* - qu'un mathématicien est quelqu'un qui compte avec des nombres très compliqués. Même des gens par ailleurs bardés de diplômes arriveront à prendre un mathématicien pour une espèce de comptable un peu étrange... D'ailleurs ne parle-t-on pas d'analphabétisme *numérique* dans les sociétés modernes ?

Beaucoup des Mathématiques sont aujourd'hui enseignées à l'école en mettant l'emphase sur l'application de règles ou de formules dont on ne demande pas une compréhension réelle. Ainsi la Géométrie n'est présentée qu'à travers des manipulations algébriques de coordonnées cartésiennes ou encore l'Algèbre linéaire est présenté comme du calcul avec des matrices. En mettant l'emphase sur la résolution de problèmes répétitifs, on présente les Mathématiques comme une espèce de soupe de formules dont on ne veut surtout pas connaître les ingrédients. C'est là une aberration puisqu'on néglige alors complèment deux des aspects fondamentaux de l'activité mathématique : la faculté de *raisonnement* et la *créativité*. Plus que toute autre activité humaine, les mathématiques relèvent essentiellement du domaine des idées. On peut expérimentalement trouver tous des nombres premiers aussi grands que l'on veut, mais ceci ne montre pas qu'il y en a une infinité. C'est seulement le raisonnement mathématique correct émis par Archimède il y a plus de deux millénaires qui permet d'*affirmer* qu'il y a une infinité de nombres premiers : si l'on suppose qu'il n'y en a qu'un nombre fini que l'on ordonne du plus petit au plus grand $p_1 < p_2 < ... < p_n$, alors le nombre $p = p_1\, p_2\, ...\, p_n + 1$ est à la fois premier et plus grand que tous les p_i $(1 \leq i \leq n)$, ce qui constitue une contradiction. Avec une idée mathématique bien trouvée, on résout en quelques lignes un problème que l'on aurait pu étudier indéfiniment à l'ordinateur sans réel succès.

Les objets rencontrés dans ce cours auront pour la plupart déjà été vus par des étudiants aujourd'hui inscrits dans un diplôme de premier cycle en sciences. Les notions de *suite*, de *série*, de *fonction continue*, de *fonction dérivable*, etc font déjà partie du bagage de tous ceux suivant ce cours. Par contre le point-de-vue que l'on aura sur ces objets est radicalement différent : on cherchera à raisonner de manière rigoureuse sur ces notions et ceci mènera à la construction d'un ensemble de connaissances assises sur des fondements solides, qui pourra être généralisé dans des cours plus avancés de Mathématiques ou encore appliqués à des domaines où l'on s'attend à ce que les mathématiques utilisées soient rigoureuses. Au lecteur ou à la lectrice qui se demande pourquoi les mathématiciens s'entêtent-ils à vouloir toujours tout démontrer, on peut rappeler les mots très justes du mathématicien français Roger Godement :

(1) Toute assertion qui n'est pas intégralement démontrée est potentiellement fausse.
(2) Utiliser une assertion non complètement démontrée pour en prouver d'autres augmente exponentiellement les risques d'erreurs.
(3) C'est à l'auteur d'une assertion qu'incombe la charge de la démontrer.

Comme dans toute introduction à un sujet aussi vaste, il y a une bonne part de choix tant pour ce qui est de la matière que la présentation. La contrainte principale en est une de temps : il y a seulement deux cours d'Analyse réelle à l'Université du Québec à Montréal, ce qui signifie que le cours *Analyse I* contient à toute fin pratique la matière de deux cours tels que prénsentés un peu partout ailleurs en Amérique du nord... Donner un cours où la rigueur est moindre sous prétexte d'une économie de temps n'est pas une option puisque ce ne serait alors qu'une répétition d'un cours pré-universitaire ! Donner un cours où certains chapitres seraient sautés complètement créerait des lacunes difficilement surmontables plus tard dans la formation des élèves et c'est donc très peu envisageable. La seule option satisfaisante est d'affronter les difficultés avec un soucis d'économie de temps et d'énergie dans la présentation de la matière. Comprendre une théorie mathématique ne se réduit pas à savoir en appliquer mécaniquement les résultats, mais plutôt en retenir les points essentiels au point de pouvoir, grosso modo, en reconstruire la structure logique, tout comme vraiment comprendre un théorème implique que l'on peut retracer les grandes lignes de sa preuve. Cette approche a le mérite de présenter les fondements de l'Analyse dans $\mathbb{R}$ en encourageant les étudiants à développer l'autonomie nécessaire pour des lectures complémentaires, ce qui leur servira dans la suite de leurs études quelque soit le cheminement emprunté.

En rédigeant ces notes de cours j'espère aider les étudiants à travailler de manière sérieuse et efficace. Ce travail implique une lecture critique des résultats et preuves présentées dans le cours, la recherche personnelle d'exemples illustrant une matière qui se veut forcément théorique, la résolution de problèmes proposés dans le texte et dans les séances de travaux pratiques ainsi que la consultation d'ouvrages complémentaires pour développer une perspective plus large sur la matière.

CHAPITRE 1

Les nombres réels et leurs suites

1. Quelques pré-requis et conventions

Un des faits saillants d'un cours d'Analyse mathématique - par opposition à un cours de Calcul différentiel par exemple - est le niveau de rigueur accru auquel l'on s'attend. Si plusieurs des objets qui définissent ce cours ont déjà été rencontrés auparavant (nombres réels, suites, fonction continues, etc) des questions fondamentales ont été laissées de côté et doivent maintenant être étudiées avec soin. Dans cette section nous résumons rapidement quelques notions pré-requises pour la suite du cours.
A la base de l'édifice mathématique, du moins dans sa pratique depuis plus de 100 ans, il y a le domaine de la Logique. Notre but ne sera certainement pas de donner ici un cours de Logique mathématique - ce qui prendrait en soi une session universitaire - mais plutôt de mettre en place certains fondements qui nous aiderons dans notre étude. Si ce qui suit semblera nettement insuffisant à un logicien, nous devrons nous en contenter par souci d'économie de temps et un syllabus déjà bien chargé, et un lecteur particulièrement intéressé par les fondements pourra toujours consulter des ouvrages qui finiront par le convaincre que ce que nous faisons ici est parfaitement rigoureux. Une des utilisations que nous ferons de la Logique mathématique sera d'écrire de façon très brève des énoncés mathématiques qu'il serait long à écrire en mots. Nous donnerons des exemples au fur et à mesure.

La première notion que nous prendrons pour acquis est celle d'*ensemble*. Pour ce cours un ensemble sera tout simplement une collection d'éléments que l'on considère comme ayant quelque chose en commun. De façon générale, un ensemble sera noté par une lettre majuscule comme A, B ou C et ainsi de suite. Un ensemble particulièrement banal mais qui reviendra très souvent dans le cours est l'*ensemble vide* qui, par définition, ne contient pas d'élément... On note cet ensemble par le symbole $\varnothing$.

Etant donné un ensemble A et un élément a, si cet élément appartient à l'ensemble A on écrit $a \in A$, alors que si a n'est pas dans A on écrit $a \notin A$. Bien sûr ces symboles "$\in$" et

“$\notin$” sont mutuellement exclusifs. On appelle *sous-ensemble* de A un ensemble B ayant la propriété que tout élément dans B est également dans A. Ceci est noté de la façon suivante en langage logique :

$$b \in B \Rightarrow b \in A.$$

Lorsque B est sous-ensemble de A on écrit $B \subseteq A$. La notation $\subseteq$ n'exclut pas que B soit en fait égal à l'ensemble A. On notera ici un point important : deux ensembles A et B sont dits “égaux” si l'on a $B \subseteq A$ ainsi que $A \subseteq B$. Ceci signifie que lorsque l'on doit montrer que deux ensemble mathématiques sont égaux, il y a deux inclusions à vérifier. Inversement, un sous-ensemble non-vide $B \subseteq A$ sera dit *propre* si l'on a $B \neq A$, ce que l'on écrira comme $B \subset A$. C'est-à-dire qu'il existe $a \in A$ tel que $a \notin B$, ce qui s'écrit comme

$$\exists a \in A \mid a \notin B.$$

Lorsque B est sous-ensemble de A on a la notion naturelle de sous-ensemble *complémentaire* de B dans A :

$$B^c = \{a \in A \mid a \notin B\}.$$

Exercice 1.1. *Vérifiez de manière immédiate que $A^c = \varnothing$. Par ailleurs établissez que l'on a toujours $(B^c)^c = B$. De plus montrez que pour $A, B \subseteq E$ on a : $A = B \iff A^c = B^c$.*

Il y a deux opérations fondamentales sur les ensembles :

L'*union* de A et B : $A \cup B = \{a \mid a \in A \text{ ou } a \in B\}$.

L'*intersection* de A et B : $A \cap B = \{a \mid a \in A \text{ et } a \in B\}$.

Exercice 1.2. *Montrez que $A \cup B$ est le plus petit ensemble contenant A et B, alors que $A \cap B$ est le plus grand ensemble contenu dans A et B.*

De façon générale on introduit la notion d'union et d'intersection finie :

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup ... \cup A_n \quad \text{et} \quad \bigcap_{j=1}^{n} A_j = A_1 \cap A_2 \cap ... \cap A_n.$$

Et de même pour un ensemble d'indices I, qui peut être infini, on introduit $\bigcup\limits_{i \in I} A_i$ et $\bigcap\limits_{j \in I} A_j$.

EXERCICE 1.3. *Démontrez les lois de De Morgan :*

$$\bigcup_{i\in I} A_i = (\bigcap_{i\in I} A_i^c)^c \quad et \quad \bigcap_{i\in I} A_i = (\bigcup_{i\in I} A_i^c)^c,$$

en vérifiant les deux inclusions " $\subseteq$ " et " $\supseteq$ ".

Etant donné la notion d'ensemble, il nous sera intéressant de considérer des *applications* entre ensembles. Par application entre A et B nous entendons un façon d'associer à chaque $x \in A$ un élément $f(x) \in B$. Nous notons ceci mathématiquement par $f\colon A \to B$. L'élément $f(x)$ est appelé *image de* x par l'application f. Nous insistons ici pour que cette image soit *unique* pour chaque $x \in A$.

$f\colon A \to B$ est dite *injective* si : $x \neq y \Rightarrow f(x) \neq f(y)$. Par contraposition[1], ceci est équivalent à dire que : $f(x) = f(y) \Rightarrow x = y$. C'est-à-dire que des éléments différents de A sont envoyés sur des éléments différents de B.

$f\colon A \to B$ est dite *surjective* si : $\forall z \in B\ \exists x \in A \mid z = f(x)$. En termes plus usuels, pour tout $z \in B$ il existe un x dans A tel que z soit l'image de x par f.

Lorsque l'application $f\colon A \to B$ est à la fois injective et surjective, on dit que f est une application *bijective* et que les ensembles A et B sont en bijection. En particulier on dit qu'ils ont la même *cardinalité*. Par exemple les ensembles $\{a, b, c, ..., x, y, z\}$ et $\{1, 2, 3, ..., 24, 25, 26\}$ sont en bijection de manière évidente. En Théorie des ensembles, la notion de bijection est l'équivalence naturelle entre ensembles.

Le premier ensemble ayant une certaine utilité dans ce cours est l'ensemble des *entiers naturels* : $\mathbb{N} = \{1, 2, 3, 4, ...\}$.[2] Il s'agit d'un ensemble fort utile pour compter et il a la propriété d'être *infini*. Ceci est une conséquence du fait que l'ensemble des entiers naturels $\mathbb{N}$ ne possède pas d'élément maximal pour la relation d'ordre intuitive dont on peut le munir : soit $k \in \mathbb{N}$ un candidat éventuel au titre d'élément maximal, mais alors on sait que $k + 1$ est également dans $\mathbb{N}$ et ce nombre est plus grand que k, ce qui démontre notre assertion.

1. Opération logique naturelle que l'on formalise ici par : $A \Rightarrow B \iff \neg B \Rightarrow \neg A$.
2. Certains textes mathématiques incluent 0 dans les entiers naturels d'autres pas, ceci n'est qu'une convention de départ et n'est pas vraiment important.

Une propriété surprenante des ensembles infinis comme $\mathbb{N}$ est qu'ils peuvent être mis en bijection avec des sous-ensembles propres. Par exemple si l'on définit l'ensemble

$$2\mathbb{N} = \{2, 4, 6, ..., 2k, 2(k+1), ...\}$$

alors l'application : $\mathbb{N} \to 2\mathbb{N}$ donnée par $f(k) = 2k$ est une bijection entre ces deux ensembles, même si $2\mathbb{N}$ est strictement contenu dans $\mathbb{N}$. Ainsi les deux ensembles ont même cardinalité, même si $\mathbb{N}$ possède des éléments qui ne sont pas dans $2\mathbb{N}$!

Finalement étant donné deux applications $f\colon A \to B$ et $g\colon B \to C$ on peut considérer leur *composition* qui sera l'application $g \circ f\colon A \to C$ donnée par $g \circ f(x) = g(f(x))$. On vérifie aisément que la composition d'applications vérifie la propriété d'*associativité* :

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

si bien que les parenthèses sont inutiles dans la notation et elles seront omises en général.

2. Ensembles ordonnés et Principe d'induction

Une *relation d'ordre* notée "$\leq$" sur un ensemble S est une relation sur les éléments de S qui satisfait :

(1) $a \in S$ et $b \in S \Rightarrow a \leq b$ ou $b \leq a$.

(2) $a \leq b$ et $b \leq a \Rightarrow a = b$.

(3) $a \leq b$ et $b \leq c \Rightarrow a \leq c$.

Une relation d'ordre étant donnée, on peut se permettre une certaine flexibilité dans la notation : $a \leq b$ pourra être écrit comme $b \geq a$. De même si $a \leq b$ et que $a \neq b$ on écrira $a < b$ ou encore $b > a$.

Définition 2.1. *Un ensemble S muni d'une relation d'ordre $\leq$ est appelé un ensemble ordonné.*

Sur l'ensemble des nombres naturels $\mathbb{N}$ on met une relation d'ordre satisfaisant les axiomes suivants :

(1) Tout sous-ensemble non-vide S de $\mathbb{N}$ possède un *premier élément* : $\exists x \in S \mid y \in S \Rightarrow y \geq x$.

(2) Tout élément de $\mathbb{N}$, à l'exclusion du premier élément, possède un *prédécesseur immédiat* : $\forall x \in \mathbb{N},\ x \neq 1,\ \exists y \in \mathbb{N} \mid y < x$ et $z < x \Rightarrow z \leq y$.

(3) $\mathbb{N}$ n'a pas de dernier élément : $\forall x \in \mathbb{N}\ \exists y \in \mathbb{N} \mid x < y$.

L'ensemble $\mathbb{N}$ muni de sa relation d'ordre usuelle ($1 < 2 < 3 < ... < n < n+1 < ...$) satisfait clairement ces axiomes. De plus, on peut montrer dans un cours de Logique que si N et M sont des ensembles ordonnés satisfaisant les axiomes (1), (2) et (3), alors il existe une bijection entre N et M qui préserve leurs relations d'ordres respectives. Nous nous contenterons donc dans ce cours de prendre la relation d'ordre usuelle sur $\mathbb{N}$ comme fondement de notre construction des nombres.

EXERCICE 2.2. *Montrez que tout élément de* $\mathbb{N}$ *possède un* successeur immédiat *:*

$$\forall x \in \mathbb{N}\ \exists y \in \mathbb{N} \mid x < y \text{ et } x < z \Rightarrow y \leq z.$$

Un sous-ensemble $S \subset \mathbb{N}$ ayant un dernier élément est *fini*. Autrement, si S ne possède pas de dernier élément, on dit que S est un sous-ensemble *infini*. Ceci nous permet de définir un ensemble quelconque comme étant fini (respectivement infini) s'il peut être mis en bijection avec un sous-ensemble fini (respectivement infini) de $\mathbb{N}$.

Une propriété de la plus haute importance de l'ensemble ordonné $\mathbb{N}$ est qu'il permet d'établir le *Principe d'induction mathématique*. Nous en donnerons deux énoncés, un pour les ensembles dans $\mathbb{N}$ et un pour les propriétés indexées par $\mathbb{N}$.

THÉORÈME 1. **(Principe d'induction pour les ensembles)**
Soit $S \subseteq \mathbb{N}$ *un ensemble tel que*

(1) $1 \in S$

(2) $x \in S$ *implique que le successeur immédiat de* x *est dans* S.

Alors on a $S = \mathbb{N}$.

Preuve. Nous raisonnons par l'absurde[3]. Supposons que $S \neq \mathbb{N}$. Alors comme $S \subset \mathbb{N}$, on a S^c non-vide. Par l'axiome (1) sur l'ordre de $\mathbb{N}$, l'ensemble S^c possède alors un premier élément x. Comme $1 \in S$, on sait que $x \neq 1$. Par l'axiome (2) sur l'ordre de $\mathbb{N}$, on sait que x a un prédécesseur immédiat y. Comme x est le premier élément de S^c, forcément $y \in S$. Mais alors l'hypothèse (2) du théorème implique que le successeur immédiat de y est dans S. Mais ceci signifie que $x \in S$, une contradiction, donc notre hypothèse de départ était fausse et l'on a bien $S = \mathbb{N}$. $\square$

3. Nous supposons les étudiants familiers avec ce procédé de preuve où l'on suppose la conclusion fausse, on en dérive une contradiction mathématique, ce qui permet conclure que la conclusion devait être vraie.

Une seconde façon d'énoncer le Principe d'induction est de considérer un énoncé $\mathcal{P}$ qui dépend de $\mathbb{N}$. Pour montrer que l'énoncé $\mathcal{P}(n)$ est vrai pour tout $n \in \mathbb{N}$, on ne peut clairement pas vérifier explicitement $\mathcal{P}(1), \mathcal{P}(2), \mathcal{P}(3)$ et ainsi de suite... ceci prendrait une infinité de temps ! Le Principe d'induction est l'outil mathématique qui nous permet de résoudre le problème et il s'énonce de la manière suivante :

THÉORÈME 2. **(Principe d'induction pour les propositions)**
Etant donné un énoncé $\mathcal{P}(n)$ dépendant de $n \in \mathbb{N}$, si l'on a

(1) $\mathcal{P}(1)$ *vrai*

(2) $\mathcal{P}(k)$ *vrai* $\Rightarrow \mathcal{P}(k+1)$ *vrai, pour un k supposé quelconque.*

Alors on a $\mathcal{P}(n)$ vrai pour tout $n \in \mathbb{N}$.

Preuve. On définit tout simplement l'ensemble $S = \{n \in \mathbb{N} \mid \mathcal{P}(n) \text{ est vraie}\}$. Cet ensemble vérifie les hypothèses du Principe d'induction pour les ensembles, et on en conclut donc que $\mathcal{P}(n)$ est vraie pour tout $n \in \mathbb{N}$. □

L'hypothèse "$\mathcal{P}(1)$ vrai" s'appelle le *cas de base*. La première condition dit que l'énoncé est vrai pour le premier élément de l'ensemble ordonné $\mathbb{N}$. On appelle la seconde hypothèse "$\mathcal{P}(k)$ vrai $\Rightarrow$ $\mathcal{P}(k+1)$ vrai" le *pas d'induction.* Il est important de comprendre que cette condition doit s'appliquer à un k général, c'est-à-dire sur lequel on ne fait aucune hypothèse. Combinée à la seconde hypothèse, elle permet d'affirmer que $\mathcal{P}(2)$ est vrai. Alors à nouveau en invoquant la seconde hypothèse on a $\mathcal{P}(3)$ également vrai, et ainsi de suite, pour obtenir au final que pour tout $n \in \mathbb{N}$ on a $\mathcal{P}(n)$ vrai.

Sous cette forme, on voit très bien toute la puissance du Principe d'induction mathématique : il réduit une vérification d'un nombre infini de cas (la conclusion dans son énoncé) à la vérification d'un cas de base ($n = 1$) et une vérification du pas d'induction.

EXEMPLE 2.3. *Utilisons le Principe d'induction pour montrer la formule suivante exprimant de façon concise la valeur de la somme des n premiers entiers naturels :*

$$1 + 2 + 3 + ... + n = \frac{n(n+1)}{2}.$$

Bien sûr on peut directement démontrer le résultat en sommant $1+2+3+...(n-1)+n$ et $n + (n-1) + ... + 3 + 2 + 1$ terme à terme en remarquant qu'à chaque fois l'on obtient $n+1$ et que l'on a n telles sommes.

La preuve par induction demande de vérifier explicitement la validité de la formule dans le cas $n = 1$. *Or on a bien* $1 = 1(1+1)/2$. *Le pas d'induction nous demande de montrer que l'on a*

$$1 + 2 + 3 + ... + k = \frac{k(k+1)}{2},$$

si l'on suppose qu'il est vrai que

$$1 + 2 + 3 + ... + (k-1) = \frac{(k-1)k}{2}.$$

On peut écrire

$$\begin{aligned} 1 + 2 + 3 + ... + k &= (1 + 2 + 3 + ... + (k-1)) + k \\ &= \frac{(k-1)k}{2} + k \qquad \textit{(Par l'hypothèse du pas d'induction)} \\ &= \frac{k^2 + k}{2} \\ &= \frac{k(k+1)}{2}. \end{aligned}$$

Donc le pas d'induction a été démontré et on en conclut que la formule est vraie pour tout $n \in \mathbb{N}$.

Exercice 2.4. *Démontrez l'inégalité de Bernoulli : pour tout nombre* $x \geq -1$ *et tout entier* $n \in \mathbb{N}$, *on a*

$$(1 + x)^n \geq 1 + nx.$$

Exercice 2.5. *Si on définit le coefficient binômial*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

montrez par induction que ce nombre est toujours dans $\mathbb{N}$.

Exercice 2.6. *Démontrez par induction la formule suivante, pour tout* $x \neq 1$ *:*

$$1 + x + x^2 + x^3 + ... + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Etant donné un ensemble A, on appelle une *suite* d'éléments de A une application $f \colon \mathbb{N} \to A$. Il est plus pratique d'énumérer les éléments de la suite comme $a_1,\ a_2, a_3, ...$ où l'on a $a_k = f(k)$ pour tout $k \in \mathbb{N}$. On note alors notre suite par le symbole $\{a_n\}_{n \in \mathbb{N}}$, ou plus simplement $\{a_n\}$ s'il n'y a pas de risque de confusion dans les indices. Une *sous-suite* $\{a_{n_k}\}$ de $\{a_n\}$ est obtenue en choisissant un sous-ensemble *infini* $S \subseteq \mathbb{N}$, $S =$

$\{n_1, n_2, ..., n_k, n_{k+1}, ...\}$ ordonné dans $\mathbb{N}$ par $n_1 < n_2 < ... < n_k < ...$ et en considérant les éléments $\{a_{n_1}, a_{n_2}, ..., a_{n_k}, ...\}$ correspondants dans $\{a_n\}$. Nous reviendrons en détail sur la notion de suite lorsque nous aurons construit les nombres réels.

3. Les nombres rationnels

Nous rappelons ici brièvement la construction des nombres rationnels à partir de l'ensemble $\mathbb{N}$. On considère premièrement l'ensemble

$$\mathbb{Z} = \{..., -k, -(k-1), ..., -2, -1, 0, 1, 2, ..., k-1, k, ...\}$$

appelé l'ensemble des *entiers relatifs*. Sur $\mathbb{Z}$ on a l'opération d'*addition* dont l'élément neutre est 0 : $k + 0 = k$ $(\forall k \in \mathbb{Z})$. Chaque $k \in \mathbb{Z}$ possède un *inverse additif* : $k + (-k) = 0$. L'ensemble $\mathbb{Z}$ possède également une opération de *multiplication* dont l'élément neutre est $1 \in \mathbb{Z}$: $1 \cdot k = k$ $(\forall k \in \mathbb{Z})$. Ici nous constatons que les seuls éléments dans $\mathbb{Z}$ qui ont un *inverse multiplicatif* sont les nombres 1 et -1. Nous étendons donc $\mathbb{Z}$ à un ensemble où chaque élément de $\mathbb{Z}\backslash\{0\}$ possède un inverse multiplicatif : $k \cdot (k^{-1}) = 1$, où l'élément k^{-1} a été ajouté à $\mathbb{Z}$. On demande alors également que la somme de tels éléments soient dans l'ensemble, ainsi que leurs inverses additifs. Ceci résume la construction de l'ensemble

$$\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z} \text{ premiers entre eux}\}$$

appelé l'ensemble des *nombres rationnels*.

Les opérations d'addition et de multiplication sur $\mathbb{Q}$ satisfont les propriétés importantes que nous résumons ici :

(1) fermeture sous addition et multiplication : $a, b \in \mathbb{Q} \Rightarrow a + b \in \mathbb{Q}$, $ab \in \mathbb{Q}$.

(2) commutativité : $a + b = b + a$ et $ab = ba$.

(3) associativité : $(a + b) + c = a + (b + c)$ et $a(bc) = a(bc)$.

(4) existence d'éléments neutres : $0 + x = x$ et $1 \cdot x = x$.

(5) distributivité : $a(b + c) = ab + ac$.

(6) $a, b \in \mathbb{Q} \Rightarrow a + x = b$ possède une solution unique.

(7) $a, b \in \mathbb{Q} \Rightarrow ax = b$ possède une solution unique si $a \neq 0$.

(8) $\mathbb{Q}$ est un ensemble ordonné.

(9) $b \geq c \Rightarrow a + b \geq a + c$ pour tout $a \in \mathbb{Q}$.

(10) $b \geq c \Rightarrow ab \geq ac$ pour tout $a \geq 0$ dans $\mathbb{Q}$.

(11) $1 > 0$.

(12) $\mathbb{Q}$ est *archimédien* : si $a > 0$ et $b \in \mathbb{Q}$ quelconque, il existe $n \in \mathbb{N}$ tel que $na > b$.

On dit alors que $\mathbb{Q}$ est un *corps ordonné archimédien.*

Le lecteur ou la lectrice dont c'est une première mise en contact avec des mathématiques plutôt abstraites pourrait ici se demander quel est le sens d'écrire des évidences comme "$ab = ba$", "$0 + x = x$" ou, pire, "$1 > 0$" ? Ce que nous faisons ici est très courant en Mathématiques : abstraire certains énoncés pour en faire la base d'une théorie que nous sommes en train de développer. Si la méthode semble énoncer des "évidences" c'est parce que nous sommes extrêmement familiers avec ces propriétés des nombres. Mais l'histoire des Mathématiques abonde d'exemples [4] où une familiarité avec des notions plus ou moins bien explicitées ont mené même les mathématiciens les plus illustres à commettre des erreurs résultant en des énoncés faux. Ce n'est donc pas couper les cheveux en quatre que de s'attarder à pareil formalisme, mais bien plutôt une connaissance approfondie de la nature des objets mathématiques.
Si l'intuition occupe une place importante dans le développement des théories mathématiques, la présentation finale des idées devrait, idéalement, ne découler que de propriétées élémentaires sur lesquelles les mathématiciens se sont entendus au départ. Ensuite, chaque énoncé requiert une preuve formelle, même si l'on pense que le résultat semble *évident.* Par exemple $0 \cdot x = 0$ ($\forall x \in \mathbb{Q}$) peut sembler intuitivement clair - avec raison ! - mais on doit *démontrer* cet énoncé *au moins une fois* dans sa vie de mathématicien pour pouvoir le prendre pour acquis par la suite. Cet énoncé découle des propriétés énoncées ci-dessus. Par exemple :

$$\begin{aligned} 0 \cdot x &= (1-1) \cdot x \quad \text{(car -1 inverse additif de 1)} \\ &= 1 \cdot x - 1 \cdot x \quad \text{(par distributivié de la multiplication)} \\ &= x - x \quad \text{(car 1 élément neutre de la multiplication)} \\ &= 0 \quad \text{(car $-x$ est l'inverse additif de x)} \end{aligned}$$

4. Le livre de R. Godement cité dans la bibliographie donne des exemples saisissants !

Ou encore

$$\begin{aligned} 0 \cdot x &= (0+0) \cdot x \quad \text{(car 0 est élément neutre additif)} \\ &= 0 \cdot x + 0 \cdot x \quad \text{(par distributivié de la multiplication)} \\ 0 \cdot x - 0 \cdot x &= 0 \cdot x + 0 \cdot x - 0 \cdot x \quad \text{(en retranchant } 0 \cdot x \text{ de chaque côté)} \\ 0 &= 0 \cdot x \quad (\text{car } -0 \cdot x \text{ inverse additif de } 0 \cdot x) \end{aligned}$$

On aura remarqué au passage qu'il n'y a pas qu'une façon d'arriver à la solution de tels problèmes si bien que la créativité joue ici un certain rôle, même à un stade aussi peu avancé. Afin de pratiquer la manipulation de ces règles élémentaires de l'ensemble $\mathbb{Q}$ et pour développer un certain sens de la preuve mathématique, il est recommandable que les étudiants démontrent les énoncés suivants avec la même rigueur que ce qui a été fait ci-dessus pour "$0 \cdot x = 0$" :

(1) $\forall x \in \mathbb{Q}$ strictement positif $\exists n \in \mathbb{N}$ tel que $\frac{1}{n} < x < n$. En particulier, si un rationnel non-négatif α est plus petit que $\frac{1}{n}$ pour tout $n \in \mathbb{N}$, alors $\alpha = 0$.

(2) On définit $|x| = \max\{x, -x\}$. Montrez que $|x| \geq 0$, $|x| = |-x|$ et que

$$|x+y| \leq |x| + |y| \quad \text{(Inégalité du triangle)}$$

4. Les rationnels ne suffisent pas

Par rapport aux nombres naturels $\mathbb{N}$, qui ne permettent que de compter des quantiés entières, les nombres rationnels permettent de calculer avec beaucoup plus de doigté et d'ailleurs des cultures millénaires se sont contentés de ces nombres. Et pourtant on sait depuis l'Antiquité grecque que pour décrire le monde qui nous entoure nous avons besoin de plus. En effet considérons un triangle rectangle *et* isocèle, de côtés égaux de longueur 1. Un tel triangle possède une hypothénuse de longueur a, où le nombre a satisfait la relation de Pythagore :

$$a^2 = 1^2 + 1^2 = 2.$$

PROPOSITION 4.1. *Il n'existe pas de $a \in \mathbb{Q}$ tel que $a^2 = 2$.*

Preuve. Montrons en effet que $(\frac{m}{n})^2 \neq 2$ pour tous $m, n \in \mathbb{Z}$. Commençons par décomposer m et n en isolant les facteurs de 2 qu'ils contiennent éventuellement :

$$m = m_1 \cdot 2^r \quad \text{et} \quad n = n_1 \cdot 2^s$$

où $r, s \geq 0$ entiers et m_1 et n_1 sont des entiers impairs. Alors puisque m_1^2 et n_1^2 sont impairs on a forcément (ceci demande quelques instants de réflexion) :

$$m_1^2 \cdot 2^{2(r-s)} \neq 2 \cdot n_1^2.$$

Il s'en suit que

$$\left(\frac{m}{n}\right)^2 = \left(\frac{m_1}{n_1}\right)^2 \cdot 2^{2(r-s)} \neq 2.$$

□

Cette proposition nous dit que si l'on veut pouvoir mesurer des longueurs de triangles tels que celui décrit ci-dessus, il faut des nombres encore plus généraux que ceux contenus dans $\mathbb{Q}$. Nous allons maintenant présenter un défaut encore plus important des rationnels, mais pour ceci nous avons besoin de deux notions fort importantes au préalable, notions que nous rencontrerons à nouveau au cours des prochains chapîtres.

DÉFINITION 4.2. *Soit $A \subset \mathbb{Q}$. Un majorant pour A est un nombre $M \in \mathbb{Q}$ tel que pour tout $a \in A$ on ait $a \leq M$. De même un minorant pour A est un nombre $m \in \mathbb{Q}$ tel que pour tout $a \in A$ on ait $m \leq a$.*

On dit qu'un ensemble A dans $\mathbb{Q}$ est *borné inférieurement* (respectivement *borné supérieurement*) s'il possède un minorant (respectivement un majorant). Un ensemble est dit *borné* s'il possède à la fois un minorant et un majorant. Ainsi pour un ensemble borné $A \subset \mathbb{Q}$ il existe $M \in \mathbb{Q}$ positif tel que $-M \leq a \leq M$ ($\forall a \in A$). En fait, par la propriété archimédienne des rationnels, on peut même supposer que $M \in \mathbb{N}$.

Parmi tous les majorants d'un ensemble A, il se peut qu'il en existe un qui soit *minimal*. De même parmi tous les minorants de A il se peut qu'il y en ait un qui soit *maximal*.

DÉFINITION 4.3. *On appelle le plus petit majorant de l'ensemble A le suprémum de A, noté* $\sup A$. *On appelle le plus grand minorant de l'ensemble A l'infimum de A, noté* $\inf A$.

Nous devons bien insister sur le fait que le suprémum ou l'infimum d'un ensemble n'existent pas forcément. Par exemple l'ensemble $\mathbb{Z} \subset \mathbb{Q}$ ne possède ni suprémum, ni infimum, puisqu'il ne possède même pas de majorant ou de minorant ! De même on voit aisément que $\mathbb{N} \subset \mathbb{Q}$ ne possède pas suprémum, n'étant pas borné supérieurement, mais $\mathbb{N}$ possède un infimum et $\inf \mathbb{N} = 1$.

Défaut de $\mathbb{Q}$: Il existe des ensembles bornés $A \subset \mathbb{Q}$ qui n'ont pas de suprémum ou d'infimum dans $\mathbb{Q}$.

Par exemple considérons l'ensemble $A = \{x \in \mathbb{Q} \mid x^2 < 2\}$. Cet ensemble est clairement borné, puisque par exemple $M_0 = 4$ et $m_0 = -4$ sont respectivement des majorants et minorants de A. Montrons toutefois que A ne possède pas de suprémum dans $\mathbb{Q}$: nous établirons que si $M \in \mathbb{Q}$ est un majorant quelconque de A, alors on peut trouver un autre majorant $M' \in \mathbb{Q}$ tel que $M' < M$. Soit donc M tel que $M^2 > 2$ et $0 < M < 2$, qui sera clairement un majorant de A. On a alors $M^2 = 2 + h$ pour un certain h tel que $0 < h < 2$ par construction. Posons alors

$$M' = M - \frac{h}{4}$$

qui sera un nombre satisfaisant $M' < M$ et pour lequel

$$\begin{aligned} {M'}^2 = (M - \frac{h}{4})^2 &= M^2 - \frac{Mh}{2} + \frac{h^2}{16} \\ &> M^2 - \frac{Mh}{2} \quad (\text{car } \frac{h^2}{16} > 0) \\ &> M^2 - h \quad (\text{car } 0 < M < 2) \\ &= 2. \end{aligned}$$

Puisque ${M'}^2 > 2$, on a bien M' majorant de A, ce qui conclue notre preuve.

Il est à noter que le fait que cet exemple A ne possède pas de suprémum dans $\mathbb{Q}$ est lié de façon indirecte à la Proposition 4.1 : nous avons pu supposer que $M^2 > 2$ pour tout majorant dans $\mathbb{Q}$ parce que si $M^2 < 2$ alors M ne serait pas majorant (donnez les détails ici) et que l'on sait, par cette Proposition 4.1, que $M^2 = 2$ est impossible pour $M \in \mathbb{Q}$! Si nous avions un nombre α tel que $\alpha^2 = 2$ alors on aurait clairement construit $\sup A = \alpha$.

5. La construction des réels

Il y a fort à parier que pratiquement tous les étudiants de ce cours se sont fait enseigner il y a longtemps que les nombres réels forment une droite appelée... la droite réelle. On aura même utilisé cette droite rélle sous forme d'abscisse et d'ordonnée pour parler du graphe d'une fonction réelle. En réalité ces dessins - la droite réelle y compris ! - ne correspondent à rien de bien solide, ne serait-ce que parce que la résolution du trait d'un crayon sur une feuille de papier ou celle d'un écran d'ordinateur ne permet même pas de représenter adéquatement les nombres rationnels...La droite réelle utilisée depuis des années est tout au plus un outil d'intuition pour visualiser des nombres que l'on n'a même pas pris le temps de construire... C'est un peu le propre de la pédagogie moderne de ne pas vraiment définir les objets que l'on étudie, de peur de créer d'inutiles difficultés aux étudiants... alors qu'en réalité de telles omissions empêchent un esprit curieux de vraiment comprendre les objets qu'il est sensé étudier.

Un des buts de ce cours est de ne pas arrondir inutilement les coins sur les notions fondamentales et ainsi nous étudirons dans cette section l'approche de Dedekind pour la construction des nombres réels. Plutôt que d'être un rejet de l'intuition amenée par la visualisation des réels comme constituant la droite réelle, cette approche permet d'utiliser cette intuition de façon légitime. La matière présentée ici contient en elle une part d'intuition qui devra être mise à jour par les étudiants eux-mêmes dans leur étude. On ne lit pas des mathématiques comme on lit un roman, il faut être impliqué constamment de façon à ce que ces notions abstraites deviennent parlantes à chacun à travers des exemples que chacun prendra soin de construire.

DÉFINITION 5.1. *Un ensemble $A \subset \mathbb{Q}$ est appelé une coupure de Dedekind dans $\mathbb{Q}$ si*

(1) $A \neq \varnothing$ *et* $A \neq \mathbb{Q}$

(2) $\forall x \in A\ \exists a \in A$ *tel que* $a > x$

(3) $x \in A\ ,\ y < x \Rightarrow y \in A.$

Nous pouvons nous faire une idée intuitive de ce qu'est une coupure de Dedekind de la façon suivante. Comme l'ensemble A est non-vide par (1) il contient au moins un rationnel x. Par (3) il contient également tous les rationnels plus petits que x. Par (1) on sait également que $A \neq \mathbb{Q}$, donc il existe un $z \in \mathbb{Q}$ qui n'est pas dans A. Mais alors tout rationnel $z' > z$ ne sera pas dans A également : si $z' \in A$ alors en vertu de (3) on devrait avoir $z \in A$ puisque $z < z'$. Ceci signifie qu'une coupure de Dedekind A est toujours bornée supérieurement. Vient alors la propriété (2) qui dit que cet ensemble borné supérieurement ne possède pas d'élément maximal. Toutes ces propriétés mises ensemble permettent donc de s'imaginer une coupure de Dedekind dans $\mathbb{Q}$ comme une sorte de demi-intervalle de rationnels sur la droite réelle, borné à droite et n'ayant pas d'élément maximal.

Un point important concerne l'égalité entre deux coupures de Dedekind. Etant donné deux coupures de Dedekind dans $\mathbb{Q}$, A et B, comme ce sont des sous-ensembles de $\mathbb{Q}$, on doit interpréter “$A = B$” comme une équation d'ensembles ce qui signifie donc que l'on a à la fois $A \subseteq B$ et $B \subseteq A$.

Observons que tout rationnel $r \in \mathbb{Q}$ définit naturellement un coupure de Dedekind de la façon suivante : soit $A(r) = \{x \in \mathbb{Q} \mid x < r\}$. Clairement les conditions (1) et (3) sont satisfaites immédiatement. Pour ce qui est de (2), on remarque que si $x \in A$, alors le nombre $z = \frac{x+r}{2}$ satisfait $x < z < a$ et donne donc la propriété désirée. On appelle une telle coupure une *coupure de rationnels.* Notons de plus que pour $r \neq s$ des rationnels distincts, on a $A(r) \neq A(s)$. Finalement une lecture attentive de ce qui précède aura sûrement suggéré que pour cette construction $A(r)$, on a le fait remaquable suivant : $\sup A(r) = r$ (donnez les détails de la preuve de ceci pour vérifier que vous comprenez bien !).

Le défaut des rationnels rencontré à la section précédente peut alors être rephrasé en termes de coupures dans $\mathbb{Q}$:

Défaut de $\mathbb{Q}$: Il existe des coupures de Dedekind dans $\mathbb{Q}$ qui ne possèdent pas de suprémum dans $\mathbb{Q}$.

En effet, l'ensemble $A = \{x \in \mathbb{Q} \mid x < 0 \text{ ou } x^2 < 2\}$ est clairement une coupure (exercice) qui n'a pas de suprémum par le travail qui a été fait à la fin de la section 4. La construction des réels permet de se défaire de cette situation ennuyante.

DÉFINITION 5.2. *L'ensemble des nombres réels, noté $\mathbb{R}$, est l'ensemble de toutes les coupures de Dedekind dans $\mathbb{Q}$.*

En considérant la correspondance $r \leftrightarrow A(r)$, nous pouvons alors naturellement identifier $\mathbb{Q}$ comme un sous-ensemble de $\mathbb{R}$ en considérant les coupures de rationnels correspondant à chaque rationnel. Nous voulons ensuite munir l'ensemble $\mathbb{R}$ d'une relation d'ordre ainsi que d'opérations d'addition et de multiplication qui en feront un corps ordonné archimédien à l'instar de $\mathbb{Q}$.

La relation d'ordre que nous proposons sur l'ensemble des coupures de Dedekind dans $\mathbb{Q}$ est la suivante :

$$A \leq B \iff A \subseteq B.$$

L'égalité dans la relation d'ordre ($\leq$ et $\geq$) correspond alors à l'égalité des coupures en tant qu'ensembles. L'inégalité stricte $A < B$ sera employée lorsque l'on aura $A \subset B$.

Proposition 5.3. **(Principe de trichotomie dans $\mathbb{R}$)**
Etant donné deux coupures $A, B \in \mathbb{R}$, trois possibilités mutuellement exclusives existent : $A = B$, $A < B$ ou $A > B$.

Preuve. Par définition, on a directement que $A = B$ exclut $A < B$ et $A > B$. Montrons alors que $A < B$ et $A > B$ sont mutuellement exclusives. Supposons au contraire que les deux relations soient vraies simultanément. Comme $A < B$ il existe un rationnel $r \in B$ tel que $r \notin A$. De même puisque $B < A$, il y a un rationnel $s \in A$ qui n'est pas dans B. Mais par les propriétés d'une coupure, on sait que $s \in A$ et $r \notin A$ implique que $s < r$. Pour les mêmes raisons, $r \in B$ et $s \notin B$ impliquent que $r < s$. Mais r et s sont alors des rationnels pour lesquels à la fois $r < s$ et $r > s$, ce qui est absurde.
Finalement montrons qu'une des trois relations doit être vraie pour A et B quelconques. Supposons que $A \neq B$. On a alors que les ensembles A et B sont distincts. S'il existe $r \in A$ qui n'est pas dans B, alors on a $B \subset A$ par les propriétés des coupures et donc $B < A$. Sinon, il existe $s \in B$ qui n'est pas dans A et alors on doit avoir $A \subset B$ et ainsi $A < B$.

□

Par ailleurs on vérifie facilement à partir des définitions que si $A \leq B$ et $B \leq C$, alors $A \leq C$, ce qui achève de démontrer que l'on a bien une relation d'ordre sur $\mathbb{R}$. Notons que cette relation d'ordre coïncide avec la relation d'ordre naturelle sur $\mathbb{Q}$ au sens où l'on a pour $r, s \in \mathbb{Q}$ que $r \leq s$ si et seulement si $A(r) \leq A(s)$, où le symbole "$\leq$" désigne dans un cas la relation d'ordre sur $\mathbb{Q}$ et dans l'autres celle sur $\mathbb{R}$.

L'opération d'addition sur $\mathbb{R}$ se définit de la façon suivante : si $I, J \in \mathbb{R}$ sont des coupures dans $\mathbb{Q}$, alors on pose

$$I + J = \{x + y \mid x \in I \ , \ y \in J\}.$$

On vérifie facilement que $I + J$ est bien un réel (c'est-à-dire une coupure de Dedekind dans $\mathbb{Q}$) et que l'opération d'addition satisfait $I + J = J + I$ ainsi que $I + (J + K) = (I + J) + K$. L'élément neutre pour cette addition est donné par la coupure $O = \{r \in \mathbb{Q} \mid r < 0\}$, qui vérifie clairement $I + O = I$ pour tout $I \in \mathbb{R}$. Les réels tels que $I \leq O$ (respectivement $I < O$) sont dits *négatifs* (respectivement *strictement négatifs*), alors que les $J \geq O$ sont dits *positifs* (et les $J > O$ sont dits *strictement positifs*). Chaque $I \in \mathbb{R}$ possède alors un unique inverse additif simplement donné par la coupure

$$(-I) = \{-M \in \mathbb{Q} \,|M \text{ est un majorant de } I\}.$$

Ceci implique facilement que l'équation $I + X = J$ possède une unique solution X en fonction de I et J.

L'opération de multiplication $I \cdot J$ est un peu plus difficile à donner puisque le résultat dépend du caractère positif ou négatif à la fois de I et de J. Le cas de base est celui de la multiplication de réels I et J positifs : on définit la coupure

$$I \cdot J = O \cup \{x \cdot y \mid x \in I \ , \ y \in J \text{ positifs}\}.$$

Si l'on introduit préalablement la *valeur absolue* d'un réel I par

$$|I| = \begin{cases} I & \text{si } I \geq O \\ -I & \text{si } I < O \end{cases}$$

on peut alors définir la multiplication dans tous les cas par

$$I \cdot J = \begin{cases} |I| \cdot |J| & \text{(si } I \text{ et } J \text{ ont même signe)} \\ -(|I| \cdot |J|) & \text{(si } I \text{ et } J \text{ ont des signes opposés).} \end{cases}$$

Un détail technique doit être souligné ici. Nous avons défini une relation d'ordre ainsi que des opération d'addition et de multiplication sur les coupures de Dedekind, en particulier pour les coupures rationnelles. On devrait formellement vérifier que dans le cas des coupures rationnelles ceci correspond précisément à ce que l'on connaissait déjà dans le cas de $\mathbb{Q}$. Nous ne faisons pas les détails ici, mais c'est bien le cas (voir **??** Théorème 1.28). De même il est possible de montrer à partir de ce que nous avons présenté dans cette section que $\mathbb{R}$ satisfait les condition d'un corps ordonné. Finalement notons que $\mathbb{R}$ est archimédien : Si $I > O$ et $J \in \mathbb{R}$ quelconque, alors il existe $n \in \mathbb{N}$ tel que $n \cdot I > J$. En effet, J étant une coupure dans $\mathbb{Q}$ on sait qu'il existe un rationnel $q > J$. Comme $I > O$ on peut trouver un rationnel $r \in I$ tel que $r > 0$ et alors la propriété archimédienne dans $\mathbb{Q}$ nous donne n tel que $n \cdot r > q$. Il s'en suit que $n \cdot I > q > J$ comme voulu.

On pourrait ici se demander quelle est l'utilité d'avoir construit cet ensemble $\mathbb{R}$ qui contient $\mathbb{Q}$ et qui partage avec lui un ordre ainsi que les mêmes caractéristiques algébriques.

La différence essentielle pour ce cours est mise en évidence lorsque l'on considère les coupures de Dedekind dans $\mathbb{R}$. Nous rappelons que nous avons établi qu'il y a des coupures de rationnels qui ne possèdent pas de suprémum dans $\mathbb{Q}$. Si nous considérons maintenant une *coupure de réels*, c'est-à-dire un sous-ensemble $\mathcal{J} \subset \mathbb{R}$ tel que

(1) $\mathcal{J} \neq \varnothing$ et $\mathcal{J} \neq \mathbb{R}$

(2) $\forall J \in \mathcal{J}\ \exists K \in \mathcal{J}$ tel que $J < K$

(3) $\forall J \in \mathcal{J}\ ,\ L < J \Rightarrow L \in \mathcal{J}$

alors on a le résultat fondamental suivant :

Théorème 3. **(Théorème de complétude des réels)**
Toute coupure dans $\mathbb{R}$ admet un suprémum dans $\mathbb{R}$.

Avant de démontrer ceci, il est bon de remarquer que par construction une coupure de réels est un ensemble de réels dont les éléments sont des coupures de rationnels... Ce que le Théorème 3 nous dit en quelque sorte c'est que contrairement au passage des rationnels aux coupures de Dedekind dans $\mathbb{Q}$ où l'on gagnait quelque chose, c'est-à-dire les réels, le passage des réels aux coupures de Dedekind dand $\mathbb{R}$ n'élargit pas notre système de nombres.

Preuve. Soit $\mathcal{J}$ une coupure de réels et posons

$$U = \bigcup_{J \in \mathcal{J}} J.$$

On veut premièrement montrer que U définit une coupure dans $\mathbb{Q}$, auquel cas on aura $U \in \mathbb{R}$. Selon la Définition 5.1 il y a trois conditions à vérifier. Par construction, on a clairement $U \neq \varnothing$ et $U \neq \mathbb{Q}$, donc la première condition est satisfaite.
Si $x \in U$, alors la définition de U dit qu'il existe un $J_x \in \mathcal{J}$ tel que $x \in J_x$. Comme J_x est une coupure dans $\mathbb{Q}$, on a pour tout $y < x$ que $y \in J_x$ et donc également $y \in U$, ce qui vérifie la troisième condition.
De même, si $x \in U$ on a $x \in J_x$ et comme J_x est coupure dans $\mathbb{Q}$, on sait qu'il existe $z \in J_x$ tel que $x < z$. On a donc trouvé $z \in U$ tel que $x < z$ et la seconde condition est satisfaite.

Par définition de U et de la relation d'ordre sur $\mathbb{R}$ on a que $U \geq J$ pour tout $J \in \mathcal{J}$, ce qui signifie que U est un majorant de $\mathcal{J}$. Montrons finalement que U est en fait le plus petit majorant de $\mathcal{J}$: si V est un autre majorant de $\mathcal{J}$, alors $V \geq J$ pour tout $J \in \mathcal{J}$. Mais alors on a également que $V \supseteq U = \bigcup_{J \in \mathcal{J}} J$. Donc on a bien $U \leq V$ comme voulu. En résumé, nous avons construit $U \in \mathbb{R}$ tel que $U = \sup \mathcal{J}$. $\square$

Ce théorème se révèlera très important dans le cours, comme il sera vu aux chapitres suivants. Mais déjà il nous permet de nous débarrasser pour la suite du langage des coupures de Dedekind puisque l'on sait maintenant qu'à chaque coupure I dans $\mathbb{R}$ on peut associer de façon unique son suprémum $\sup I$ et qu'inversement tout réel x peut être vu comme le suprémum de la coupure $A(x) = \{r \in \mathbb{Q} \mid r < x\}$. L'usage veut alors que l'on adopte plutôt pour notation d'éléments de $\mathbb{R}$ des symboles comme x, y ou z plutôt que les coupures associées $A(x), A(y)$ ou $A(z)$. Ceci allègera d'ailleurs le texte !

Le Théorème 3 est tellement important que nous en montrons une variante qui est un peu plus générale :

Théorème 4. *Soit $S \subset \mathbb{R}$ un ensemble non-vide et borné supérieurement. Alors S possède un suprémum dans $\mathbb{R}$.*

Preuve. On veut se ramener au Théorème 3, donc on veut définir une coupure dans $\mathbb{R}$ à partir de S. Soit $\mathcal{U}$ défini par

$$x \in \mathcal{U} \iff \exists y \in S \mid y > x.$$

Intuitivement, $\mathcal{U}$ "remplit" S dans $\mathbb{R}$. Montrons que $\mathcal{U}$ est une coupure dans $\mathbb{R}$. On a immédiatement que $\mathcal{U} \neq \varnothing$ et $\mathcal{U} \neq \mathbb{R}$ car $S \neq \varnothing$ et S borné. De plus si $x \in \mathcal{U}$ et $z < x$, alors par construction de $\mathcal{U}$ il existe $y \in S$ tel que $y > x$. Mais alors $y > z$ donne $z \in \mathcal{U}$ à nouveau par définition de $\mathcal{U}$. Finalement, si $x \in \mathcal{U}$ on a $y \in S$ tel que $y > x$ et on sait alors construire un réel x' tel que $y > x' > x$, donc on a trouvé $x' \in \mathcal{U}$ satisfaisant $x' > x$. Ceci achève de montrer que $\mathcal{U}$ est une coupure dans $\mathbb{R}$.
Par le Théorème 3, on sait qu'il existe $U \in \mathbb{R}$ tel que $U = \sup \mathcal{U}$. On sait que $U \notin \mathcal{U}$ car $\mathcal{U}$ est une coupure et alors par construction de $\mathcal{U}$ on a que $y \leq U$ pour tout $y \in S$. Autrement dit, U est un majorant de S... et il sera le plus petit tel majorant puisque $U = \sup \mathcal{U}$ et que S contient des éléments plus grands que ceux de $\mathcal{U}$. $\square$

La dernière remarque que nous ferons dans cette section portant sur la construction des nombres réels est qu'il y a un résultat parfaitement analogue au Théorème 4 pour l'infimum des ensembles :

Tout sous-ensemble non-vide S de $\mathbb{R}$ qui est borné inférieurement possède un infimum $\inf S$ dans $\mathbb{R}$.

6. Le lien entre $\mathbb{Q}$ et $\mathbb{R}$

Il est intéressant de voir quel est le lien plus précis entre $\mathbb{Q}$ et l'ensemble $\mathbb{R}$ qui le contient. D'une part, il est facile de voir qu'entre deux rationnels r et s il existe toujours un réel... tout simplement le rationnel $\frac{r+s}{2}$! La proposition suivante est un peu plus surprenante à prime abord :

PROPOSITION 6.1. *$\mathbb{Q}$ est dense dans $\mathbb{R}$, c'est-à-dire que pour tous $x, y \in \mathbb{R}$ tels que $x < y$ il y a toujours un rationnel $r \in \mathbb{Q}$ tel que $x < r < y$.*

Preuve. Ceci découle de la définition des réels comme coupures dans $\mathbb{Q}$ et de ce que signifie alors $x < y$: tout simplement que leurs coupures associées I et J satisfont $I \subset J$ et constituent des ensembles distincts de $\mathbb{Q}$. Il existe donc un rationnel r qui est dans J sans être dans I. On a alors automatiquement $x < r$. La définition de coupure pour J et le fait que $r \in J$ impliquent alors également que $r < y$ tel que voulu. $\square$

Il est clair que le sous-ensemble $\mathbb{N} \subset \mathbb{R}$ n'est pas, quant à lui, dense dans $\mathbb{R}$ (exercice). On pourra donc être surpris lorsque l'on s'intéresse aux cardinalités des ensembles $\mathbb{N}$, $\mathbb{Q}$ et $\mathbb{R}$ d'apprendre que les deux premiers ont même cardinalité mais que la cardinalité de $\mathbb{R}$ est d'une nature complètement différente, comme le montrent les deux propositions suivantes.

PROPOSITION 6.2. *Les ensembles $\mathbb{N}$ et $\mathbb{Q}$ ont même cardinalité.*

Preuve. Pour mettre $\mathbb{N}$ et $\mathbb{Q}$ en bijection, on procède de la façon suivante. On ordonne les éléments de $\mathbb{Q}$ sous la forme d'un carré de longueur infinie en plaçant les rationnels comme suit :

$$\begin{array}{cccccc}
\frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots \\
\frac{2}{1} & \frac{2}{3} & \frac{2}{5} & \frac{2}{7} & \frac{2}{9} & \cdots \\
\frac{3}{1} & \frac{3}{2} & \frac{3}{4} & \frac{3}{5} & \frac{3}{7} & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{array}$$

C'est-à-dire que nous avons placé sur la $k^{\text{ème}}$ ligne les rationnels de numérateur égal à k, en prenant soin d'éliminer toutes les répétitions possibles (par exemple $\frac{2}{2}, \frac{2}{4}$ ou encore $\frac{3}{6}$). Ce carré infini sur la droite et vers le bas a la propriété de posséder des diagonales de longueur k, allant de la première entrée de la $k^{\text{ème}}$ ligne à la $k^{\text{ème}}$ entrée de la première colonne. En parcourant toutes ces diagonales l'une à la suite de l'autre, on établit une bijection entre $\mathbb{N}$ et les rationnels strictement positifs. Clairement la même chose peut alors être faite avec tout $\mathbb{Q}$. □

Proposition 6.3. *L'ensemble $\mathbb{R}$ n'est pas dénombrable.*

Preuve. Nous allons en fait montrer que l'ensemble $\{x \in \mathbb{R} \mid 0 < x < 1\}$ n'est pas dénombrable. Pour ce faire nous allons pour l'instant accepter une description très pratique des nombres réelles sur laquelle nous reviendrons plus tard : la description par développements décimaux. Supposons que notre ensemble est dénombrable. On peut alors énumérer ses éléments par

$$\begin{aligned} &\alpha_1 = 0, a_{11}\ a_{12}\ a_{13}\ ... \\ &\alpha_2 = 0, a_{21}\ a_{22}\ a_{23}\ ... \\ &\alpha_3 = 0, a_{31}\ a_{32}\ a_{33}\ ... \\ &... \end{aligned}$$

Pour chaque $i \in \mathbb{N}$ prenons alors un nombre $p_i = \begin{cases} 2 & \text{si } a_{ii} \neq 2 \\ 3 & \text{si } a_{ii} \neq 3. \end{cases}$
Il n'y a absolument rien de spécial à ces nombres 2 ou 3 choisis, outre le fait que nous venons alors de construire un nombre réel $x = 0, p_1\ p_2\ p_3\ ...$ dans $(0, 1)$ qui doit forcément être distinct de tous les α_i : comme la $i^{\text{ème}}$ décimale de chaque α_i est différente de la $i^{\text{ème}}$ décimale de x, on a automatiquement que $|\alpha_i - x| > 10^{-i} > 0$ et donc les réels sont distincts. Il s'en suit que $\mathbb{R}$ n'est pas dénombrable. □

Nous en déduisons qu'il existe au moins *deux* types d'infini mathématique bien distincts au niveau de la cardinalité des ensembles. Les méthodes de l'Analyse mathématique réussissent donc même à donner un sens à la notion d'infini, dcrit auparavant de façon floue par les philosophes. Il n'est pas faux de dire que le passage du Calcul différentiel intuitif tel que développé par Fermat, Newton et Leibnitz à l'Analyse mathématique telle que développée au XIX$^{\text{ème}}$ siècle correspond à une meilleure compréhension de phénomènes faisant intervenir des infinis mathématiques. Ceci sera fait dans les sections et chapitres suivants lorsque nous étudierons la convergence des suites et des séries, la continuité et dérivabilité des fonctions réelles et, si le temps le permet, l'intégration.

7. Le principe des intervalles emboîtés dans $\mathbb{R}$

Une grande partie des fondements de l'Analyse réelle repose sur le résultat fondamental, dont nous avons vu deux versions équivalentes, voulant qu'un ensemble borné dans $\mathbb{R}$ possède toujours un supremum. Pour construire la théorie à partir de là, il y a de nombreuses possibilités. On choisit ici de développer une bonne partie des notions de ce chapitre en utilisant une conséquence de ce résultat fondamental : le *principe des segments emboîtés*. Comme on le verra bientôt cet outil allègera certaines preuves ou, encore, en fera ressortir leur caractère géométriquement intuitif.

DÉFINITION 7.1. *Un intervalle $I \subseteq \mathbb{R}$ est un sous-ensemble de $\mathbb{R}$ ayant la propriété que si $x, y \in I$ satisfont $x < y$, alors pour tout z tel que $x < z < y$, on a $z \in I$.*

Par le Théorème 3 on sait que tout intervalle I qui est borné possède un infimum a et un supremum b. On appelle a et b les *extrémités* de I. Ceci est, bien sûr, motivé par l'image géométrique que l'on se fait de $\mathbb{R}$. On désignera I par divers symboles, selon que a ou b soient dans I :

$$
\begin{aligned}
&I = [a, b] \text{ si } a, b \in I && \text{(intervalle fermé)}\\
&I = (a, b) \text{ si } a, b \notin I && \text{(intervalle ouvert)}\\
&I = [a, b) \text{ si } a \in I, b \notin I && \text{(intervalle semi-ouvert)}\\
&I = (a, b] \text{ si } a \notin I, b \in I && \text{(intervalle semi-ouvert)}
\end{aligned}
$$

Etant donné un intervalle $I = [a, b]$, on définit sa *longueur* comme étant le nombre réel $l(I) = b - a$. Un ensemble d'intervalles $\{I_n\}$ est dit *décroissant* si pour chaque $n \in \mathbb{N}$ on a $I_{n+1} \subseteq I_n$. Etant donné un ensemble décroissant de segments $\{I_n\}$, on dit que l'ensemble est *emboîté* si, de plus, la suite $\{l(I_n)\}$ décroît vers 0, c'est-à-dire :

$$\forall \epsilon > 0 \; \exists N \in \mathbb{N} \text{ tel que } n > N \; \Rightarrow \; l(I_n) < \epsilon.$$

THÉORÈME 5. *(Principe des segments emboîtés) Soit $\{I_n\}$ un ensemble de segments fermés et emboîtés dans $\mathbb{R}$. Alors l'ensemble $\bigcap_{i=1}^{\infty} I_n$ consiste d'un unique point de $\mathbb{R}$.*

Preuve. Posons $I_n = [a_n, b_n]$ pour $n \in \mathbb{N}$. Comme l'ensemble $\{I_n\}$ est emboîté, on a

$$a_n \leq a_{n+1} < b_{n+1} \leq b_n.$$

En fait on a plus encore : pour tous $m, n \in \mathbb{N}$, on a $a_n < b_n$. En effet, si $n < m$, on sait que $a_n \leq a_m < b_m$, alors que si $m < n$ on a plutôt $a_n < b_n \leq b_m$, dans les deux cas on utilise la propriété d'emboîtement.
Ceci implique que chaque b_m est un majorant pour l'ensemble $\{a_n\}$. Donc $\{a_n\}$ possède un supremum : $u = \sup\{a_n\}$. Comme u est un majorant pour $\{a_n\}$, on a $a_n \leq u$ pour tout $n \in \mathbb{N}$. Et comme c'est également le plus petit majorant, on a que $u \leq b_n$ pour tout $n \in \mathbb{N}$. Ainsi on a construit $u \in \mathbb{R}$ tel que $a_n \leq u \leq b_n$ pour tout $n \in \mathbb{N}$, si bien que $u \in I_n$ pour tout $n \in \mathbb{N}$ puisque I_n est intervalle fermé. Donc

$$u \in \bigcap_{i=1}^{\infty} I_n.$$

Montrons maintenant qu'il s'agit de l'unique point dans cette intersection. C'est ici que l'on utilise le fait que la longueur des intervalles I_n décroit vers 0. Soit $x \neq u$, si bien que $|x - u| > 0$. Comme la suite $\{b_n - a_n\}$ décroît vers 0, on en déduit qu'il existe $N \in \mathbb{N}$ tel que si $n > N$ alors $|x - u| > b_n - a_n$. En se souvenant que $u \in I_n$, on en conclut que $x \notin I_n$ pour $n > N$ et donc, à fortiori que $x \notin \bigcap_{i=1}^{\infty} I_n$, ce qui complète la preuve. □

Il est important de bien saisir que la preuve utilise de façon cruciale toutes les hypothèses. D'une part le résultat serait faux si l'on considérait la notion d'intervalles fermés et emboîtés dans $\mathbb{Q}$ car rien n'oblige le point d'intersection trouvé dans la preuve pour $\mathbb{R}$ à être un nombre rationnel ! On pourra aisément construire des exemples de ceci, par approximation de réels quelconques par des rationnels. En second lieu, il est important que le lecteur ou la lectrice se convainque que les hypothèses I_n fermés et I_n emboîtés sont également essentielles. On cherchera, par exemple, à construire un ensemble d'intervalles ouverts et emboîtés pour lequel l'intersection de ses éléments est vide, ou encore un exemple d'ensemble d'intervalles fermés dont la longueur ne décroît pas vers 0 pour lequel l'intersection est également vide... et ainsi de suite. Plus le lecteur ou la lectrice fera attention à ces exemples, meilleure sera la compréhension de ce résultat fondamental.

8. Suites convergentes et suites de Cauchy dans $\mathbb{R}$

Dans l'hypothèse même d'un ensemble de segments emboîtés, on a déjà introduit implicitement la notion de suite convergeant vers 0.

DÉFINITION 8.1. *Etant donné une suite de réels* $\{a_n\}$ *et un élément* $l \in \mathbb{R}$*, on dit que* $\{a_n\}$ *converge vers* l *si*

$$\forall \epsilon > 0 \ \exists N \in \mathbb{N} \text{ tel que } n > N \ \Rightarrow |a_n - l| < \epsilon.$$

On dit alors que la suite $\{a_n\}$ a pour limite l et on écrit

$$\lim_{n\to\infty} a_n = l \text{ ou encore } \{a_n\} \to l.$$

On remarque la propriété suivante de la notion de limite : si elle existe, alors elle est *unique* : en effet, si l_1 et l_2 sont deux limites d'une même suite $\{a_n\}$, on montre que $l_1 = l_2$. Pour ce faire, il suffit de démontrer que quelque soit le réel positif $\epsilon > 0$ choisi, on a $|l_1 - l_2| < \epsilon$.[5] En effet, soient

$$\begin{aligned} N_1 \in \mathbb{N} \text{ tel que } n > N_1 \ &\Rightarrow |a_n - l_1| < \frac{\epsilon}{2} \\ N_2 \in \mathbb{N} \text{ tel que } n > N_2 \ &\Rightarrow |a_n - l_2| < \frac{\epsilon}{2}. \end{aligned}$$

Alors on a

$$\begin{aligned} |l_1 - l_2| &= |l_1 - a_n + a_n - l_2| \\ &\leq |l_i - a_n| + |a_n - l_2| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Géométriquement, la condition $\{a_n\} \to l$ signifie que si l'on regarde un voisinage I_ϵ de rayon $\epsilon > 0$ et centré en l, avec $\epsilon > 0$ quelconque mais fixé, alors à partir d'un certain $N \in \mathbb{N}$, la condition $n > N$ implique que $a_n \in I_\epsilon$. Il est important de saisir que pour chaque tel $\epsilon > 0$ on pourra trouver un tel $N = N_\epsilon$.

EXEMPLE 8.2. *La suite* $\{\frac{1}{n} \mid n \in \mathbb{N}\}$ *converge vers* 0*. Si ceci semble intuitivement clair, le but premier d'un tel exemple est de montrer que l'on peut et doit établir ceci de façon formelle, uniquement à partir des propriétés déjà connues de* $\mathbb{R}$*, sans appel à l'intuition dans la forme finale de la preuve par des imprécisions telles "il est évident que" ou "on voit que" . Il est important de bien saisir ces idées dans leur détail le plus complet.*

Soit donc $\epsilon > 0$ *quelconque que l'on fixe. Comme* ϵ *est un nombre réel, la propriété archimédienne de* $\mathbb{R}$ *implique qu'il existe* $N \in \mathbb{N}$ *tel que* $N > \frac{1}{\epsilon}$*. Ceci peut se réécrire*

5. Ceci est un truc souvent utilisé en Analyse, expliquez pourquoi ceci est suffisant.

comme $\frac{1}{N} < \epsilon$*. Mais alors pour tout* $n > N$*, on a*

$$\frac{1}{n} < \frac{1}{N} < \epsilon,$$

ce qui est exactement la condition requise pour avoir la convergence de $\{\frac{1}{n}\}$ *vers* 0*.*

EXEMPLE 8.3. *Si on a un nombre réel* $0 < \alpha < 1$*, alors la suite* $\{\alpha^n\}$ *converge vers* 0*. Ici, on verra que la preuve est plus élaborée. Posons* $\beta = \frac{1}{\alpha}$*, si bien que* $\beta > 1$*. Alors* $\beta = 1 + h$ $(h > 0)$*. Par l'inégalité de Bernoulli, on a*

$$\beta^n = (1+h)^n > 1 + nh,$$

si bien qu'en particulier $\beta > nh$*. Soit alors* $\epsilon > 0$ *quelconque fixé. Par la propriété archimédienne de* $\mathbb{R}$*, il existe un* $N \in \mathbb{N}$ *tel que* $N > \frac{1}{\epsilon h}$ *Il s'en suit que* $Nh > \frac{1}{\epsilon}$ *et si* $n > N$*, alors ce qui précède donne*

$$\beta^n > nh > Nh > \frac{1}{\epsilon}.$$

En revenant à notre suite $\{\alpha^n\}$*, on a donc obtenu*

$$\alpha^n = \frac{1}{\beta^n} < \epsilon,$$

tel que désiré.

On aura noté les efforts non-négligeables déployés pour obtenir ces résultats que le lecteur ou la lectrice hâtifs auraient imaginé *évidents*. Il est clair que dans des exemples beaucoup plus complexes, cette approche très pédestre, malgré toute sa rigueur, n'est pas efficace. Si l'on veut néanmoins préserver l'aspect rigoureux de cette approche, il faudra développer, au cours du chapitre, des outils théoriques permettant de manipuler algébriquement et efficacement les limites de suites.

EXERCICE 8.4. *Calculez les limites des suites* $\{\frac{3n+7}{n}\}$ *et* $\{\frac{3n}{2n+1}\}$*.*

EXERCICE 8.5. *Montrez que* $\{a_n\} \to a \iff \{a_n - a\} \to 0$*.*

Une notion qui semble plus subtile que la notion de convergence d'une suite est celle de suite de Cauchy :

DÉFINITION 8.6. *Une suite* $\{a_n\}$ *est dite suite de Cauchy si la condition suivante est satisfaite :*

$$\forall \epsilon > 0 \ \exists N \in \mathbb{N} \mid \ \ n, m > N \ \Rightarrow |a_n - a_m| < \epsilon.$$

D'une manière intuitive, être suite de Cauchy signifie que les termes a_n et a_m se rapprochent de plus en plus si n et m sont "grands". On n'affirme cependant pas que ces valeurs de plus en plus rapprochées convergent vers un réel. La convergence d'une suite $\{a_n\} \to l$ implique clairement que la suite est également de Cauchy, en utilisant l'inégalité du triangle :

$$|a_n - a_m| \leq |a_n - l| + |a_m - l| < \epsilon/2 + \epsilon/2 = \epsilon,$$

pour $n, m > N = N_\epsilon$ donné par l'hypothèse de convergence de $\{a_n\}$ vers l.
Par contre il n'y a, à priori, aucune raison pour laquelle une suite de Cauchy quelconque dans $\mathbb{R}$ devrait converger dans $\mathbb{R}$. En fait, si l'on remplace $\mathbb{R}$ par $\mathbb{Q}$ on sait déjà depuis le début de ce chapitre qu'il existe des suites de rationnels qui sont de Cauchy mais qui ne convergent pas dans $\mathbb{Q}$. Le lecteur ou la lectrice réfléchira à quelques exemples. Avant de revenir au lien précis entre les suites de Cauchy et les suites convergentes dans $\mathbb{R}$, donnons une caractérisation plus géométrique de la notion de suite de Cauchy :

PROPOSITION 8.7. *Une suite $\{a_n\}$ est suite de Cauchy si et seulement si*

$$\forall \epsilon > 0 \ \exists I_\epsilon \text{ intervalle fermé avec } l(I_\epsilon) = \epsilon \text{ et } \exists N \in \mathbb{N} \text{ tels que } n > N \ \Rightarrow a_n \in I_\epsilon.$$

Preuve. Dans une direction, supposons $\{a_n\}$ de Cauchy : pour $n, m > N$ on a $|a_n - a_m| < \epsilon/2$, en particulier, pour $n > N$ on a $|a_n - a_{N+1}| < \epsilon/2$. Posons alors

$$I_\epsilon = [a_{N+1} - \epsilon/2, a_{N+1} + \epsilon/2]$$

et on obtient donc bien que $a_n \in I_\epsilon$ pour tout $n > N$.
Inversement, pour $\epsilon > 0$ quelconque fixé, supposons que $n > N \ \Rightarrow a_n \in I_\epsilon/2$ pour un certain $N \in \mathbb{N}$. Alors pour $n, m > N$, on a $a_n, a_m \in I_\epsilon/2$ et donc en particulier on trouve

$$|a_n - a_m| \leq l(I_\epsilon/2) = \epsilon/2 < \epsilon,$$

et $\{a_n\}$ est bien de Cauchy. □

EXERCICE 8.8. *Montrez qu'une suite de Cauchy dans $\mathbb{R}$ est toujours bornée.*

EXERCICE 8.9. *Soient a_0 et a_1 des réels distincts et définissons la suite $\{a_n\}$ par récurrence en posant $a_k = \frac{a_{k-1}+a_{k-2}}{2}$ pour tout $k \geq 2$. Montrez que $\{a_n\}$ est une suite de Cauchy.*

Il s'agit d'une autre manifestation de la *complétude* de $\mathbb{R}$ que l'on puisse démontrer le résultat suivant :

THÉORÈME 6. *Toute suite de Cauchy dans $\mathbb{R}$ converge dans $\mathbb{R}$.*

Preuve. On utilise la caractérisation des suites de Cauchy en termes d'intervalles fermés :

$$\forall \epsilon > 0 \ \exists I_\epsilon \text{ intervalle fermé avec } \ l(I_\epsilon) = \epsilon \text{ et } \exists N \in \mathbb{N} \text{ tels que } n > N \ \Rightarrow a_n \in I_\epsilon.$$

Soient alors

$$I_1 \ , \ l(I_1) < 1, \text{ avec } n > N_1 \ \Rightarrow a_n \in I_1,$$
$$I_2' \ , \ l(I_2') < \frac{1}{2}, \text{avec } n > N_2' \ \Rightarrow a_n \in I_2'.$$

L'ensemble $I_2 = I_1 \cap I_2'$ satisfait $I_2 \subseteq I_1$ et $l(I_2) \leq \frac{1}{2}$. De plus, si $n > N = \max\{N_1, N_2'\}$, on a $a_n \in I_2$. En continuant ainsi ce procédé, on construit un ensemble d'intervalles fermés $\{I_k\}$ et une suite de naturels $\{N_k\}$ tels que

$$I_{k+1} \subseteq I_k \ , \ l(I_k) \leq \frac{1}{k} \text{ et } \forall n > N_k \text{ on a } a_n \in I_k.$$

Comme $\{I_k\}$ est ensemble de segments fermés et emboîtés dans $\mathbb{R}$, le principe des segments emboîtés donne

$$\bigcap_{i=1}^{\infty} I_n = l.$$

Démontrons que ce nombre l est en fait la limite de la suite $\{a_n\}$. Soit $\epsilon > 0$, alors on peut trouver, par la construction qui précède, I_k tel que $l(I_k) < \epsilon$ et N_k tel que pour tout $n > N_k$ on ait $a_n \in I_k$, puisque $\{a_n\}$ est suite de Cauchy. Mais par construction on a également $l \in I_k$, si bien que l'on a pour $n > N_k$ que

$$|a_n - l| \leq l(I_k) < \epsilon.$$

Ceci est exactement la condition pour que $\{a_n\}$ converge vers l. □

Donnons maintenant quelques propriétés des suites convergentes dans $\mathbb{R}$:

Proposition 8.10. *Toute suite convergente est bornée :* $\exists M \in \mathbb{R}$ *t.q.* $|a_n| < M \ \forall n \in \mathbb{N}$.

Preuve. Supposons que $\{a_n\} \to l$. Alors on a, par exemple, un $N \in \mathbb{N}$ tel que $n > N$ implique $|a_n| < 1$. Donc $|a_n| < |l| + 1 \ \forall n > \mathbb{N}$. On pose alors

$$M = \max\{|a_1|, |a_2|, ..., |a_N|, |l|\} + 1,$$

et l'on aura bien $|a_n| < M$ $(\forall n \in \mathbb{N})$ tel que désiré. □

Proposition 8.11. *Si* $\{a_n\}$ *converge vers* l *et que* $\{a_{n_k}\}$ *est une sous-suite de* $\{a_n\}$, *alors* $\{a_{n_k}\} \to l$ *également.*

Preuve. Facile et laissée en exercice. □

PROPOSITION 8.12. *Si on a $\{a_n\} \to a$ et $\{b_n\} \to b$ ainsi que $\alpha, \beta \in \mathbb{R}$, alors*

$$\{\alpha a_n + \beta b_n\} \to \alpha a + \beta b.$$

Preuve. Il suffit de montrer que $\{\alpha a_n\} \to \alpha a$ et que $\{a_n + b_n\} \to a + b$. Pour la première affirmation, on distingue deux cas. Si $\alpha = 0$, c'est trivial. Sinon, $\alpha \neq 0$ et fixons $\epsilon > 0$ arbitraire. Comme $\{a_n\} \to a$, on a $N \in \mathbb{N}$ tel que pour $n > N$ on ait

$$|a_n - a| < \frac{\epsilon}{|\alpha|}.$$

Alors, pour la suite $\{\alpha a_n\}$ et $n > N$, on a

$$|\alpha a_n - \alpha a| = |\alpha||a_n - a| < |\alpha|\frac{\epsilon}{|\alpha|} = \epsilon,$$

tel que voulu.
Par ailleurs, à nouveau pour $\epsilon > 0$ donné, la convergence de $\{a_n\}$ vers a et de $\{b_n\}$ vers b donne

$$\begin{aligned} &\exists N_1 \in \mathbb{N} \text{ tel que } \; n > N_1 \; \Rightarrow |a_n - a| < \frac{\epsilon}{2}, \\ &\exists N_2 \in \mathbb{N} \text{ tel que } \; n > N_2 \; \Rightarrow |b_n - b| < \frac{\epsilon}{2}. \end{aligned}$$

Ainsi pour $n > \max\{N_1, N_2\}$, on a

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

comme voulu. □

PROPOSITION 8.13. *Si $\{a_n\} \to a$ et $\{b_n\} \to b$, alors $\{a_n b_n\} \to ab$*

Preuve. Preuve laissée en exercice. □

PROPOSITION 8.14. *Si $\{a_n\} \to a$ et $\{b_n\}$ est une suite qui n'est jamais nulle et satisfait $\{b_n\} \to b \neq 0$, alors la suite quotient $\{a_n/b_n\}$ converge vers a/b.*

Preuve. Comme $\{b_n\}$ est toujours non-nulle et que sa limite est également non-nulle, il existe $\alpha > 0$ tel que $|b_n| > \alpha$ pour tout $n \in \mathbb{N}$ et donc $|b| \geq \alpha$ (ces deux affirmations devraient demander au moins quelques instants de réflexion...). En outre, $\{a_n\}$ converge, donc elle est bornée en vertu de la Proposition (8.10) : $|a_n| < M$ $(\forall n \in \mathbb{N})$ pour un certain

$M \in \mathbb{R}$.
Soit alors $\epsilon > 0$ donné. Par convergence des suites en question, on a $N \in \mathbb{N}$ tel que

$$|a_n - a| < \frac{\epsilon\alpha}{2} \text{ ainsi que } |b_n - b| < \frac{\epsilon\alpha^2}{2M}.$$

Alors pour $n > N$, on a également

$$\begin{aligned}|\frac{a_n}{b_n} - \frac{a}{b}| &= \frac{|a_n b - b_n a|}{|b_n b|} = \frac{|a_n b - a_n b_n + a_n b_n - b_n a|}{|b_n b|} \\ &\leq \frac{|a_n||b - b_n|}{|b_n b|} + \frac{|a_n - a|}{|b|} < \frac{M(\epsilon\alpha^2/2M)}{\alpha^2} + \frac{\epsilon\alpha/2}{\alpha} = \epsilon,\end{aligned}$$

où l'on a utilisé toutes les inégalités obtenues au cours de la preuve. Ceci donne bien le résultat escompté. □

Il est désirable d'avoir des critères de convergence que l'on peut appliquer facilement dans de nombreux cas. Nous en avons déjà vu un, le critère de Cauchy : pour montrer qu'un suite converge dans $\mathbb{R}$, il suffit de vérifier que c'est une suite de Cauchy. Lorsque l'on a plus d'information sur la suite, on peut parfois conclure directement que la suite converge, comme dans le cas suivant :

Théorème 7. *Toute suite croissante et bornée de $\mathbb{R}$ converge dans $\mathbb{R}$.*

Preuve 1. Comme $\{a_n\}$ est croissante et bornée, on a

$$a_{n+1} \geq a_n \ (\forall n \in \mathbb{N}),$$

$$a_n < M \ (\forall n \in \mathbb{N}).$$

Choisissons $\epsilon > 0$ ainsi que $\alpha < a_1$. Par la propriété archimédienne de $\mathbb{R}$, il existe alors $l \in \mathbb{N}$ tel que $\alpha + l\epsilon > M$. On peut donc construire un $k \in \mathbb{N}$ tel que $\alpha + k\epsilon < a_N$ pour un certain $N \in \mathbb{N}$ mais tel que $\alpha + (k+1)\epsilon \geq a_n$ pour tout $n \in \mathbb{N}$ (simplement enlever juste assez d'entiers à $l \in \mathbb{N}$). Comme $\{a_n\}$ est croissante, on a alors

$$\alpha + k\epsilon < a_N \leq a_{N+1} \leq a_{N+2} \leq ...$$

Mais alors pour les entiers $n, m > N$, on a ainsi

$$|a_n - a_m| < \alpha + (k+1)\epsilon - (\alpha + k\epsilon) = \epsilon.$$

Ceci signifie que $\{a_n\}$ est suite de Cauchy et donc on sait qu'elle converge. □

Preuve 2. Une preuve plus immédiate peut être donnée si l'on utilise la caractérisation du supremum d'un ensemble. Fixons $\epsilon > 0$ arbirairement. Comme $\{\alpha_n\}$ est bornée, elle possède un supremum $\alpha \in \mathbb{R}$. En tant que supremum, α a la propriété qu'il existe $N \in \mathbb{N}$ tel que

$$\alpha - \epsilon < a_n \leq \alpha = \sup\{a_n\}.$$

Mais alors le fait que $\{a_n\}$ soit croissante implique

$$\alpha - \epsilon < a_N \leq a_n \leq \alpha \ (\forall n > N),$$

ce qui dit exactement que $\{a_n\}$ converge vers α. □

En combinant ce théorème ainsi que la Proposition 8.10, on peut énoncer le résultat de manière légèrement différente : Supposons que $\{a_n\}$ est une suite croissante. Alors $\{a_n\}$ converge si et seulement si $\{a_n\}$ est suite bornée. Bien sûr il existe des suites bornées qui ne convergent pas et par ce que nous venons de faire, nous savons alors qu'elles ne peuvent pas être croissantes.

EXERCICE 8.15. *Montrez que la suite $\{a_n\}$ donnée de façon récursive par $s_1 = \sqrt{2}$ et $s_k = \sqrt{2 + \sqrt{s_{k-1}}}$ $(k \in \mathbb{N})$ converge.*

Le résultat suivant est particulièrement utile lorsque l'on veut réduire la vérification de la convergence d'une suite à celle de deux autres, plus faciles à traiter et convergeant vers la même limite :

THÉORÈME 8.16. **(Théorème des gendarmes)**
Soit $\{a_n\}$ une suite pour laquelle on peut trouver deux suites $\{x_n\}$ et $\{y_n\}$ convergeant toutes deux vers une même limite l et satisfaisant

$$x_n \leq a_n \leq y_n \ (\forall n \in \mathbb{N}).$$

Alors on a que $\{a_n\} \to l$ également.

Preuve. Ces deux inégalités pour la suite $\{a_n\}$ en fonction des suites $\{x_n\}$ et $\{y_n\}$ donnent :

$$a_n - l \leq y_n - l \qquad \text{et} \qquad l - a_n \leq l - x_n.$$

On en déduit alors

$$|a_n - l| \leq \max\{l - x_n, y_n - l\}$$

et la convergence de $\{x_n\}$ et $\{y_n\}$ vers l donne bien la convergence de $\{a_n\}$ vers l également. $\square$

Ce résultat a pour application un critère qui peut souvent se révéler utile pour démontrer la convergence d'une suite vers 0 ou, encore, pour démontrer que la suite diverge :

PROPOSITION 8.17. *Soit $\{a_n\}$ une suite telle que $a_n \neq 0$ $(\forall n \in N)$ et que l'on ait*

$$\lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| = \alpha.$$

Alors si $\alpha > 1$ la suite $\{a_n\}$ diverge et si $\alpha < 1$ la suite $\{a_n\}$ converge vers 0.

Preuve. On donne rapidement l'idée pour le cas $\alpha < 1$. Pour N assez grand, on a par la condition de limite que pour $\alpha < r < 1$,

$$\begin{aligned}
&|a_{N+1}| < r|a_N| \\
&|a_{N+2}| < r|a_{N+1}| < r^2|a_N| \\
&|a_{N+3}| < r|a_{N+2}| < r^2|a_{N+1}| < r^3|a_N| \\
&\quad\vdots \\
&|a_{N+k}| < r|a_{N+k-1}| < \ldots < r^k|a_N|.
\end{aligned}$$

Ce qui donne donc

$$-r^k|a_N| < a_{N+k}| < r^k|a_N|.$$

Comme $|a_N|$ est ici un terme constant, on a que les membres de gauche et de droite convergent vers 0, puisque $r < 1$ par construction. Le théorème des gendarmes donne alors que $\{a_n\} \to 0$ comme voulu. Pour le cas $\alpha > 1$, laissé en exercice, il s'agit de démontrer que la suite satisfait alors

$$|a_{N+k}| > r^k|a_N| \quad \text{pour } \alpha > r > 1 \text{ et } N \text{ assez grand.}$$

$\square$

EXERCICE 8.18. *Montrez que la suite $\{a^n\}$ converge pour tout $|a| < 1$ et diverge pour tout $|a| \geq 1$.*

EXEMPLE 8.19. **(La construction du nombre e)**
Considérons la suite $\{(1 + \frac{1}{n})^n\}$. Nous allons montrer que cette suite est croissante et que

3 en est un majorant. Il découlera alors du Théorème 7 que cette suite converge.
Par le binôme de Newton, on peut écrire

$$(1+\frac{1}{n})^n = 1+1+\frac{1}{2!}\frac{n-1}{n}+\frac{1}{3!}\frac{n-1}{n}\frac{n-2}{n}+... \text{ (somme de } n+1 \text{ termes)}$$
$$(1+\frac{1}{n+1})^{n+1} = 1+1+\frac{1}{2!}\frac{n}{n+1}+\frac{1}{3!}\frac{n}{n+1}\frac{n-1}{n+1}+... \text{ (somme de } n+2 \text{ termes)}$$

Comme on a $\frac{k}{n} < \frac{k+1}{n+1}$ *pour tout* $k \in \mathbb{N}$, *la croissance de la suite est claire en comparant terme à terme les deux équations ci-dessus.*

A nouveau par le binôme de Newton, on a

$$\begin{aligned}(1+\frac{1}{n})^n &< 1+1+\frac{1}{2!}+\frac{1}{3!}+...+\frac{1}{n!}\\ &< 1+1+\frac{1}{2}+\frac{1}{2^2}+...+\frac{1}{2^{n-1}}\\ &= 1+\frac{1-(\frac{1}{2})^n}{1-\frac{1}{2}}\\ &= 3-\frac{1}{2^{n-1}} < 3.\end{aligned}$$

Ceci nous donne donc la convergence de la suite $\{(1+\frac{1}{n})^n\}$ *en vertu du Théorème 7. On* définit *alors le nombre réel*

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

Donnons maintenant une estimation de cette limite en donnant deux inégalités pour ce nombre e. *Premièrement fixons* $n \in \mathbb{N}$. *Alors pour tout* $m > n$ *on obtient*

$$(1+\frac{1}{m})^m > 1+1+\frac{1}{2!}\frac{m-1}{m}+\frac{1}{3!}\frac{(m-1)(m-2)}{m^2}+...+\frac{1}{n!}\frac{(m-1)...(m-n+1)}{m^{n-1}}.$$

Il est important de réaliser que s'il est vrai que le membre de droite dépend de m, *il s'agit néanmoins d'une somme finie de* n *termes. On peut donc prendre la limite de chaque côté lorsque* m *tend vers l'infini pour obtenir l'inégalité :*

$$e \geq 1+1+\frac{1}{2!}+\frac{1}{3!}+...+\frac{1}{n!}.$$

Par ailleurs, on a également l'inégalité suivante par le binôme de Newton, pour $m > n$ *quelconques :*

$$(1+\frac{1}{m})^m < 1+1+\frac{1}{2!}+...+\frac{1}{n!}+\frac{1}{2^n}+\frac{1}{2^{n+1}}+...+\frac{1}{2^m},$$

où l'on remarque que les $m-n$ derniers termes représentent une somme d'une progression géométrique de raison $\frac{1}{2}$. On a donc

$$e \leq 1 + 1 + \frac{1}{2!} + ... + \frac{1}{n!} + \frac{1}{2^{n-1}} \ (\forall n \in \mathbb{N}).$$

Ces deux inégalités donnent donc déjà l'approximation plutôt précise du nombre e :

$$1 + 1 + \frac{1}{2!} + ... + \frac{1}{n!} \leq e \leq 1 + 1 + \frac{1}{2!} + ... + \frac{1}{n!} + \frac{1}{2^{n-1}}.$$

Voilà un exemple du savoir-faire des mathématiciens des siècles passés pour évaluer un nombre sans avoir accès aux outils de l'informatique...

9. Le Théorème de Bolzano-Weierstrass

On a vu plus tôt que $\{a_n\}$ bornée et croissante implique $\{a_n\}$ convergente. Si on laisse tomber l'hypothèse de croissance de la suite $\{a_n\}$, il est facile de trouver des exemples de suites qui convergent et d'autres qui ne convergent pas :

$$\{a_n\} = \{(-1)^n \ \frac{1}{n}\} \text{ converge : } |a_n| = \frac{1}{n} \text{ qui converge vers } 0.$$
$$\{b_n\} = \{(-1)^n\} \text{ diverge : } b_{2k} = +1 \ , \ b_{2k+1} = -1 \ \forall k \in \mathbb{N}.$$

Le théorème suivant est donc très utile car il nous permet tout de même d'isoler des sous-suites de $\{a_n\}$ au comportement favorable.

Théorème 8. **(Théorème de Bolzano-Weierstrass)**
Si $\{a_n\}$ est une suite bornée dans $\mathbb{R}$, alors $\{a_n\}$ possède une sous-suite convergente.

Preuve. La preuve présentée ici utilise le principe des segments emboîtés. L'hypothèse faisant en sorte que $\{a_n\}$ est bornée implique l'existence d'un intervalle fermé $I_0 = [a, b]$ tel que $\{a_n\} \subseteq I_0$. On divise I_0 en deux sous-segments égaux,

$$I_1' = [a, \frac{b-a}{2}] \text{ et } \ I_1'' = [\frac{b-a}{2}, b].$$

Soit alors I_1 un de ces deux sous-segments contenant une infinité des a_n. On divise alors I_1 en deux sous-segments égaux I_2' et I_2'', en choisissant ensuite I_2 contenant une infinité des a_n ; on répète le processus ainsi de suite. On a alors construit successivement une suite décroissante d'intervalles fermés et bornés

$$I_0 \supset I_1 \supset I_2 \supset ... \supset I_n \supset ...,$$

avec $l(I_k) = \frac{1}{2} \cdot l(I_{k-1})$ et chaque I_k contient une infinité de termes de la suite $\{a_n\}$ par construction. La condition sur les longueurs des segments nous place alors dans les hypothèses du théorème des segments emboîtés, qui donne alors

$$x = \bigcap_{n=1}^{\infty} I_n.$$

Il est maintenant facile de construire une sous-suite de $\{a_n\}$ convergeant vers x. En effet, soit $a_{n_1} \in I_1$. Comme I_2 contient une infinité d'éléments de $\{a_n\}$, on peut trouver $n_2 > n_1$ tel que $a_{n_2} \in I_2$. Et ainsi de suite, on construit donc

$$n_1 < n_2 < ... < n_k < ...$$

avec $a_{n_k} \in I_k$ pour chaque $k \in \mathbb{N}$. Par construction, on a

$$|a_{n_k} - x| \leq \frac{1}{2^k},$$

ce qui donne bien la convergence de la sous-suite $\{a_{n_k}\}$. □

On peut donner une formulation différente du résultat de Bolzano et Weierstrass, plus géométrique, pour laquelle on a besoin de la notion suivante :

DÉFINITION 9.1. *Soit $E \subseteq \mathbb{R}$ un ensemble. On dit que $x \in \mathbb{R}$ est un point d'accumulation de E si pour tout intervalle ouvert J_x contenant le point x, on a $J_x \cap E$ ensemble infini.*

Remarquons en premier lieu qu'un point d'accumulation x d'un ensemble E n'est pas forcément dans E. Par exemple, si $E = (0, 1)$, le point 0 est point d'accumulation de E mais n'en est pas un élément. Ceci dit, même si x n'est pas forcément dans E, on peut trouver une suite de points *dans* E convergeant vers x : en effet soient dans cet exemple

$$J_x(n) = (x - \frac{1}{n}, x + \frac{1}{n}).$$

Comme x est point d'accumulation de E, on peut trouver, pour chaque n, un élément $x_n \in E \cap J_n(x)$. Alors la suite $\{x_n\} \subseteq E$ ainsi construite converge vers x puisque l'on a

$$|x_n - x| < \frac{1}{n} \ (\forall n \in \mathbb{N}).$$

THÉORÈME 9. **(Version ensembliste de Bolzano-Weierstrass)**
Tout ensemble infini et borné E dans $\mathbb{R}$ possède au moins un point d'accumulation.

Preuve. Comme E est borné, on peut trouver un intervalle I_0 tel que $E \subseteq I_0$. Comme E est infini on peut, comme dans la preuve précédente, construire une suite d'intervalles fermés et emboîtés

$$I_0 \supset I_1 \supset I_2 \supset \ldots \supset I_n \supset \ldots$$

dont chacun contient une infinité de points de E. Soit alors

$$x = \bigcap_{n=1}^{\infty} I_n.$$

On veut montrer que x est point d'accumulation de E : soit J_x quelconque contenant x, $J_x = (\alpha, \beta)$ où $\alpha < x < \beta$. On a alors un $n \in \mathbb{N}$ tel que $n > N$ implique $l(I_n) < \min\{x - \alpha, \beta - x\}$. Ainsi $E \cap I_n \subseteq E \cap J_x$ contient une infinité de points de E et x est point d'accumulation de E. □

Remarque 9.2. *On peut également définir la notion de point d'accumulation d'une suite, une notion moins forte que la notion de limite d'une suite, car une suite n'a qu'au plus une limite alors qu'elle peut avoir de nombreux points d'accumulation selon la définition suivante : x est point d'accumulation de $\{a_n\}$ si pour tout J_x intervalle ouvert autour de x, une infinité de termes de $\{a_n\}$ sont dans J_x. Une troisième version du théorème de Bolzano-Weierstrass s'énonce alors de la façon suivante :*

Toute suite bornée dans $\mathbb{R}$ possède au moins un point d'accumulation.

Exemple 9.3. *On sait que l'ensemble $\mathbb{Q}$ est dénombrable, donc on peut voir $\mathbb{Q}$ comme une suite $\{r_n\}$ de rationnels. On peut se demander quels sont les points d'accumulation de $\mathbb{Q}$ dans $\mathbb{R}$? La réponse est... tout $\mathbb{R}$!*

En effet, on se souvient que l'on a montré (Proposition 6.1) que $\mathbb{Q}$ est dense dans $\mathbb{R}$: entre deux réels $x < y$ il y a toujours un rationnel r. Soit alors $x \neq y$ et prenons la suite d'intervalles fermés et emboîtés

$$I_n = \{[x, \frac{y - (2^n - 1)x}{2^n}]\}.$$

On a par construction que x est dans l'intersection de tous les I_n et dans chaque I_k on sait construire un $r_{n_k} \in \mathbb{Q}$. Comme

$$|x - r_{n_k}| < l(I_k) \ (\forall k \in \mathbb{N}),$$

on a bien construit une suite de rationnels convergeant vers x, si bien que x est point d'accumulation de $\mathbb{Q}$.

10. Quelques commentaires pratiques

Une grande partie de ce qui a été vu jusqu'ici dans ce premier chapitre est de nature théorique. Si ceci est parfaitement normal pour construire rigoureusement les fondements de l'Analyse réelle, il n'en demeure pas moins que le lecteur ou la lectrice voulant bien comprendre la matière doit être capable d'appliquer ces notions abstraites dans des problèmes plus concrets tels le calcul de supréma, la détermination de la convergence ou divergence d'une suite donnée et calcul explicite de limites. Cette section se veut donc un résumé plus pratique autour de ces questions.

Etant donné $A \subset \mathbb{R}$, comment calculer $\sup A$?

En premier lieu il faut se poser la question à savoir si l'ensemble possède ou non un supremum. Pour cela l'ensemble doit être non-vide et borné. Pour montrer qu'un ensemble est borné, il suffit de trouver un nombre réel M plus grand que tout élément dans A. Si l'on doit, au contraire, montrer que l'ensemble n'est pas borné, on devra prendre un réel quelconque N (pas un nombre précis tel 10, π, 1 000 000 !) et montrer qu'il existe toujours des éléments de A plus grands que N. Ceci repose souvent sur la propriété archimédienne des réels.

Une fois que l'on sait que $\sup A$ existe, de façon pratique il y a quelques façons de procéder pour trouver le supremum. D'une part, on peut montrer qu'un réel α est égal à $\sup A$ en montrant que :
(1) Si $M > \alpha$ est un majorant pour A, il existe un autre majorant M' tel que $\alpha < M' < M$.
(2) Si $m < \alpha$, alors il existe $x \in A$ tel que $m < x \leq \alpha$.

La première condition vérifie que tout nombre supérieur à α ne peut être supremum de A, car il existe alors un plus petit majorant. La seconde condition vérifie que tout nombre inférieur à α ne peut être supremum de A car ce n'est pas un majorant. Ces deux conditions pratiques donnent bien que le nombre α est le plus petit majorant possible pour l'ensemble A.

D'autre part, on peut montrer qu'un nombre α est égal à $\sup A$ en vérifiant les deux conditions suivantes :
(1) Le nombre α est majorant de A.

(2) Si M est un majorant quelconque de A, alors $M \geq \alpha$.

La partie (2) est souvent démontrée par contraposition : si un nombre L satisfait $L < \alpha$, alors L n'est pas majorant de A. Ceci fait souvent intervenir la propriété archimédienne des nombres réels.

EXEMPLE 10.1. *Soit* $A = \{\frac{1}{1+1/n} \mid n \in \mathbb{N}\}$. *Intuitivement, on peut penser que* $\sup A = 1$ *et c'est ce que l'on va essayer de montrer. Comme*

$$0 < \frac{1}{1+1/n} < 1 \ (\forall n \in \mathbb{N}),$$

on sait que A *est borné supérieurement, donc* $\sup A$ *existe. Soit* $\alpha < 1$. *Peut-on trouver* $x \in A$ *tel que* $x > \alpha$ *? Si oui, on doit avoir :*

$$\frac{1}{1+1/n} > \alpha \iff 1+1/n > \frac{1}{\alpha} \iff \frac{1}{n} < \frac{1}{\alpha} - 1 \iff n > \frac{1}{\frac{1}{\alpha}-1}.$$

Puisque $\alpha \neq 1$, *en vertu de la propriété archimédienne des nombres réels, ceci est possible. Donc* α *n'est pas un majorant de* A *et l'on a ainsi* $\sup A \geq 1$ *pour avoir un majorant. D'un autre côté, on a que le nombre* 1 *est un majorant pour* A *puisque pour tout* $n \in \mathbb{N}$,

$$1 + \frac{1}{n} > 1 \Rightarrow \frac{1}{1+\frac{1}{n}} < 1.$$

Et donc $\sup A \leq 1$ *pour que ce majorant soit le plus petit. Les deux inégalités nous donnent bien* $\sup A = 1$. *Calculez pareillement* $\inf A$ *en exercice.*

Convergence ou divergence d'une suite

Il y a de nombreuses façons de démontrer la convergence d'une suite :

(1) Utiliser l'agèbre des limites et des résultats bien connus (par exemple $\{\frac{1}{n}\} \to 0$). Il faut tout de même faire attention de justifier et ne pas utiliser des formules vides de sens comme *"il est évident que..."*, ou *"on voit que..."*

(2) Utiliser le théorème des gendarmes (Théorème 8.16) avec des suite plus simples que celle à étudier et convergeant vers la limite conjecturée.

(3) Utiliser le critère de Cauchy (ceci est rarement plus utile).

(4) Utiliser les théorèmes de convergence du cours (mais s'assurer que toutes les hypothèses sont satisfaites !).

Pour montrer qu'une suite diverge, il y a également de nombreuses approches :

(1) Voir si l'on peut montrer que la suite n'est pas bornée.
(2) Trouver deux points d'accumulation distincts. C'est-à-dire deux sous-suites convergeant vers des valeurs distinctes.
(3) Utiliser le critère de Cauchy par contraposition : $\exists \epsilon_0 > 0$ tel que $\forall N \in \mathbb{N}$ il existe $n, m > N$ avec $|a_n - a_m| \geq \epsilon_0$.

EXEMPLE 10.2. *La suite $\{a_n = \{(-1)^n\}$ diverge. Une façon de voir cela est que la suite possède deux points d'accumulation, $+1$ et -1, limites respectivement des sous-suites $\{a_{2n}\}$ et $\{a_{2n+1}rb$. On peut également utiliser le critère de Cauchy, puisque l'on remarque que*

$$|a_{k+1} - a_k| = 2 \ (\forall k \in \mathbb{N}).$$

Alors pour $\epsilon_0 = 1$, on a que pour tout $N \in \mathbb{N}$, il existe k et $k+1$ tels que

$$|a_{k+1} - a_k| \geq \epsilon_0,$$

et donc $\{a_n\}$ n'est pas de Cauchy.

Pour montrer que $\{a_n\} \to l$, on peut souvent ne pas utiliser directement la définition mais plutôt montrer que l'on a

$$|a_n - l| \leq C \cdot |b_n| \ (\forall n > N),$$

pour N assez grand fixé, C une constante et $\{b_n\}$ une suite dont on sait déjà qu'elle tend vers 0. En effet, pour un nombre $\epsilon/C > 0$ quelconque, on peut trouver N_0 tel que pour $n > N_0$ on ait $|b_n| < \epsilon/C$. Alors si $n > \max\{N, N_0\}$, on aura bien

$$|a_n - l| < \epsilon$$

tel que voulu.

EXEMPLE 10.3. *Soit la suite $\{a_n\} = \{\frac{1}{n^2}\}$. Alors on sait que*

$$|a_n| \leq \frac{1}{n} \ (\forall n \in \mathbb{N}).$$

Comme la suite $\{\frac{1}{n}\}$ tend clairement vers 0 (propriété archimédienne de $\mathbb{R}$), on a directement $\{a_n\} \to 0$.

Exemple 10.4. *Considérons la suite*

$$\{a_n\} = \{\frac{4}{n\sqrt{1+1/n^2}}\}.$$

On obtient immédiatement

$$|\frac{4}{n\sqrt{1+1/n^2}}| = |\frac{4}{\sqrt{1+1/n^2}}| \cdot \frac{1}{n} \leq 4 \cdot \frac{1}{n},$$

si bien que la suite $\{a_n\}$ *tend vers* 0. *Notez que l'on évite ainsi les longues manipulations nécessaires si l'on veut procéder à partir de la définition de convergence.*

Limites de suites données par récurrence

La première étape ici est d'avoir une idée intuitive des limites possibles. A supposer qu'une limite existe, disons l, en remplaçant les valeurs de la suite dans l'équation de récurrence, il est souvent possible d'avoir de l'information. On donne ici un exemple plutôt que de développer des notions générales.

Exemple 10.5. *On cherche la limite, si elle existe, de la suite suivante donnée par récurrence :*

$$s_{n+1} = \frac{2}{1+s_n} \qquad s_0 = 0.$$

Supposons que l *est l'éventuelle limite, alors on doit avoir* $\{s_{n+1}\} \to l$ *et* $\{s_n\} \to l$. *L'algèbre des limites donne donc*

$$l = \frac{2}{1+l},$$

si bien que les deux seules possibilités à ce stade sont $l = -2$ *et* $l = 1$. *Si la limite existe, ce sera forcément une de ces deux valeurs. Comme l'expression par récurrence pour la suite dit immédiatement que la suite est positive, on peut exclure* $l = -2$ *! Essayons alors de montrer que la suite* $\{s_n\}$ *converge. Comme*

$$s_{n+1} - 1 = \frac{2}{1+s_n} - 1 = \frac{1-s_n}{1+s_n},$$

on obtient

$$|s_{n+1} - 1| = \frac{|s_n - 1|}{|1+s_n|}.$$

Cette équation dit que la suite $\{|s_n - 1|\}$ *est décroissante. En outre, on sait qu'elle est bornée inférieurement (par le nombre* 0 *!), si bien que la suite doit converger. Ceci implique immédiatement que* $\{s_n\}$ *converge, et sa limite est donc* 1

11. Le théorème de recouvrement de Borel

On dit de façon tout à fait générale qu'un ensemble E est *recouvert* par une collection d'ensembles $\mathcal{J}$ si pour chaque $x \in E$ il existe un certain $J \in \mathcal{J}$ contenant x. Etant donné $E \subseteq \mathbb{R}$ recouvert par une collection $\mathcal{J}$, on peut se demander si l'on peut remplacer la condition

$$E \subseteq \bigcup_{J \in \mathcal{J}} J \qquad \text{par la condition} \qquad E \subseteq \bigcup_{i=1}^{n} J_i$$

où $\{J_1, J_2, ..., J_n\}$ est un sous-ensemble *fini* de $\mathcal{J}$. C'est-à-dire : étant donné un recouvrement $\mathcal{J}$ de E, peut-on trouver un *sous-recouvrement fini* ? Cette section se propose d'étudier cette question dans un cas particulier.

Exemple 11.1. *L'ensemble* $(0,1)$ *est recouvert par la collection*

$$\mathcal{J} = \{(\frac{1}{n}, 1 - \frac{1}{n}) \mid n \in \mathbb{N}\},$$

comme le lecteur ou la lectrice s'en convaincra aisément. Par contre, tout sous-ensemble fini de $\mathcal{J}$ *ne pourra pas couvrir entièrement* $(0,1)$*. En effet, soient* $J_1, J_2, ..., J_N$ *quelconques choisis dans* $\mathcal{J}$*. Il existe alors par construction un* $m \in \mathbb{N}$ *tel que* $J_m = (\frac{1}{m}, 1 - \frac{1}{m})$ *soit un élément de* $\mathcal{J}$ *qui ne soit pas dans* $J_1 \cup J_2 \cup ... J_N$*. Comme* $J_m \subseteq (0,1)$*, on sait conséquemment qu'il existe des éléments de* $(0,1)$ *qui ne sont pas dans* $J_1 \cup J_2 \cup ... J_N$*, tel qu'annoncé.*

Intuitivement, à partir de la construction présentée ci-dessus, on peut penser que ce qui nous permet de construire un recouvrement de $(0,1)$ n'admettant pas un sous-recouvrement fini de cet ensemble est le fait que l'élément 1 est point d'accumulation de $(0,1)$ sans être dans cet ensemble. Ceci nous met alors sur la bonne voie pour énoncer et démontrer le résultat suivant :

Théorème 10. **(Théorème de Borel)**
Etant donné un intervalle fermé et borné $I = [a, b]$ *et un recouvrement quelconque par des intervalles ouverts,* $\mathcal{J}$*, il existe un sous-recouvrement fini de* I*.*

Preuve. La preuve est tout de même subtile et demandera une lecture attentive. On définit le sous-ensemble suivant de I :

$$E = \{x \in I \mid [a, x] \text{ est recouvrable par une nombre fini d'intervalles dans } \mathcal{J}\}.$$

On a alors $E \neq \varnothing$ puisque $a \in E$ par l'hypothèse que $\mathcal{J}$ est recouvrement de I. De plus, on remarque que si $x \in \mathbb{E}$ et que l'on a $a < y < x$, alors à fortiori $y \in E$. Aussi E étant

borné par construction, posons $\alpha = \sup E$.
Si $\alpha = b$ on n'a plus rien à démontrer car alors $E = [a, b]$ (ceci n'est pas immédiat alors justifiez l'affirmation complètement). Autrement, on suppose que $\alpha < b$ et on cherchera à obtenir une contradiction pour conclure la preuve. Soit alors $J \in \mathcal{J}$ tel que $\alpha \in J$. Prenons $y \in J$ avec $y < \alpha$. Alors ce y est construit tel que $[a, y]$ soit recouvrable par un sous-ensemble fini $J_1, J_2, ..., J_N$ de $\mathcal{J}$. Comme J est un intervalle ouvert et que $\alpha \in J$, on peut trouver un élément $z \in J$ tel que $z > \alpha$. Mais alors tout élément entre y et z est dans l'intervalle J, si bien que $[a, z]$ est couvert par $J_1, J_2, ..., J_N$ <u>et</u> J. Ceci signifie que $z \in E$ et on a donc contredit le fait que $\alpha = \sup E$. $\square$

En fait, le théorème de Borel n'est pas le résultat le plus général possible dans cette direction. Nous avons choisi de présenter un cas particulier pour faire mieux ressortir l'idée de la preuve. De façon plus générale, le théorème peut être énoncé pour tout ensemble fermé[6] et borné K dans $\mathbb{R}$. La preuve est similaire : on utilise le fait que $a = \inf K$ est dans K et au lieu de considérer $[a, x]$ pour la définition de E comme ci-dessus, on doit utiliser $[a, x] \cap K$, et la preuve peut être adaptée à ce cas plus général. D'autre part, le résultat est vrai pour des recouvrements ouverts quelconques, pas seulement par des intervalles. On obtient alors la forme la plus générale, connue sous le nom de *Théorème de Heine-Borel* :

De tout recouvrement ouvert d'un ensemble fermé et borné dans $\mathbb{R}$, on peut extraire un sous-recouvrement fini.

12. Ensembles ouverts et fermés dans $\mathbb{R}$

Dans certains manuels d'introduction à l'Analyse réelle on passe un certain temps à définir le langage de la Topologie générale dans le contexte très particulier de $\mathbb{R}$. Il semble pourtant discutable d'introduire si tôt dans le parcours académique des étudiants des notions telles les espaces métriques et topologiques, qui seront amplement revues par la suite, surtout que ceci est souvent fait au détriment de notions fines de l'Analyse dans $\mathbb{R}$ (convergence de séries, continuité uniforme, etc). Dans cette section de nature plutôt facultative, nous nous penchons un peu sur les notions d'ensemble *ouvert* et d'ensemble *fermé* dans $\mathbb{R}$ pour discuter de certains aspects topologiques sans que cela ne requière un arsenal de vocabulaire qui nous semblerait déplacé.

6. Voir la section suivante pour cela.

DÉFINITION 12.1. *Un ensemble $\mathcal{U} \subseteq \mathbb{R}$ est dit ouvert dans $\mathbb{R}$ si pour chaque $x \in \mathcal{U}$ il existe un intervalle ouvert J_x contenant x et tel que $J_x \subseteq \mathcal{U}$.*

Cette notion d'ensemble ouvert généralise clairement la notion d'intervalle ouvert, dans le sens où l'ensemble $\mathcal{U}$ ressemble *localement* à un intervalle ouvert par la condition de la définition. Par exemple l'ensemble $(0,1) \cup (2,3)$ est un ensemble ouvert qui n'est pas lui-même un intervalle. Par contre l'ensemble $(0,1] \cup (2,3)$ ne constitue pas un ensemble ouvert puisqu'autour du point $x = 1$ ont ne peut trouver aucun intervalle ouvert contenu dans notre ensemble.

PROPOSITION 12.2. *Toute réunion d'ensembles ouverts dans $\mathbb{R}$ est un ensemble ouvert. Toute intersection <u>finie</u> d'ensembles ouverts dans $\mathbb{R}$ est en ensemble ouvert.*

Preuve. Soit $x \in \bigcap_{i=1}^{n} \mathcal{U}_i$ où chacun des $\mathcal{U}_i$ est ouvert. Comme x est dans $\mathcal{U}_i$ ouvert, on peut trouver un J_x^i contenant x tel que $J_x^i \subseteq \mathcal{U}_i$. Mais alors il est facile de voir [7] que $\bigcap_{i=1}^{n} J_x^i$ est également un intervalle ouvert contenant x, qui est contenu dans $\bigcap_{i=1}^{n} \mathcal{U}_i$ par construction.
Pour la seconde affirmation, soit $x \in \bigcup_{i=1}^{\infty} \mathcal{U}_i$. On sait alors que $x/in\mathcal{U}_k$ pour un certain $k \in \mathbb{N}$. Comme $\mathcal{U}_k$ est ouvert, on a $J_x \subseteq \mathcal{U}_k$ intervalle ouvert autour de x. Mais alors on a trivialement que $J_x \subseteq \bigcup_{i=1}^{\infty} \mathcal{U}_i$ également, ce qui montre que l'ensemble est ouvert. □

EXERCICE 12.3. *Donnez un exemple d'intersection infinie d'ouverts qui n'est pas elle-même un ensemble ouvert. Trouvez dans la preuve de la proposition précédente où est-ce que l'on a utilisé de façon cruciale l'hypothèse d'intersection finie.*

Le théorème suivant donne une caractérisation complète des ouverts dans $\mathbb{R}$. Comme cette matière n'est pas centrale au cours, nous n'en donnons par la preuve ici.

THÉORÈME 12.4. *Si $\mathcal{U} \subseteq \mathbb{R}$ est un ensemble quelconque, alors $\mathcal{U}$ est réunion d'un nombre dénombrable d'intervalles ouverts disjoints.*

L'autre type de sous-ensemble de $\mathbb{R}$ qui nous intéresse dans cette section est celui d'ensemble fermé :

DÉFINITION 12.5. *On dit que $\mathcal{F} \subseteq \mathbb{R}$ est fermé si son complément $\mathcal{F}^c = \mathbb{R} \backslash \mathcal{F}$ est un ensemble ouvert.*

7. Fournissez les détails au besoin.

Ainsi par exemple $[0,1]$ est un ensemble fermé puisque l'on vérifie aisément que son complément $(-\infty,0)\cup(1,+\infty)$ est ouvert dans $\mathbb{R}$. Par contre $[0,1)$ n'est pas fermé puisque son complément $(-\infty,0)\cup[1,+\infty)$ n'est pas ouvert.

Proposition 12.6. *Toute intersection d'ensembles fermés dans $\mathbb{R}$ est un ensemble fermé. Toute réunion finie d'ensembles fermés dans $\mathbb{R}$ est une ensemble fermé.*

Preuve. Ceci est une conséquence directe de la Proposition 12.2 en utilisant les lois de De Morgan pour les ensembles. $\square$

Exercice 12.7. *Donnez un exemple d'une famille d'ensembles fermés $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ telle que $\bigcup_{n=1}^{\infty}\mathcal{F}_n$ ne soit pas un ensemble fermé.*

Exercice 12.8. *Montrez que $\mathbb{Q}\subset\mathbb{R}$ n'est ni ouvert ni fermé dans $\mathbb{R}$.*

On peut facilement caractériser les ensembles fermés dans $\mathbb{R}$ grâce à la notion de point d'accumulation :

Théorème 12.9. *$\mathcal{F}\subseteq\mathbb{R}$ est fermé $\iff$ $\mathcal{F}$ contient tous ses points d'accumulation.*

Preuve. Dans un sens, supposons $\mathcal{F}$ fermé, si bien que l'on sait que $\mathcal{F}^c$ est ouvert. Si $x\in\mathcal{F}^c$, on a $J_x\subseteq\mathcal{F}^c$ et donc $J_x\cap\mathcal{F}=\varnothing$. On a donc trouvé un voisinage ouvert de x qui ne contient pas de points de $\mathcal{F}$ et Il s'en suit que x n'est pas un point d'accumulation de $\mathcal{F}$.
Dans l'autre direction, soit $\mathcal{F}$ contenant tous ses points d'accumulation. Prenons $x\in\mathcal{F}^c$. Comme x n'est pas point d'accumulation de $\mathcal{F}$, il existe un intervalle J_x autour de x tel que $J_x\cap\mathcal{F}$ consiste d'un nombre fini de point $\{x_1,x_2,...,x_N\}$. Si l'on pose

$$d=\min\{|x-x_1|,|x-x_2|,...,|x-x_N|\},$$

alors on a $(x-d,x+d)$ ouvert qui n'intersecte pas $\mathcal{F}$, donc $(x-d,x+d)\subseteq\mathcal{F}^c$. Ceci signifie que $\mathcal{F}^c$ est ouvert et donc que $\mathcal{F}$ est fermé. $\square$

On peut par ailleurs se demander dans quelle mesure les propriétés d'être ouvert et fermé peuvent coexister, et la réponse est surprenante par sa simplicité :

Proposition 12.10. *Les seuls ensembles à la fois ouvert et fermé dans $\mathbb{R}$ sont $\varnothing$ et $\mathbb{R}$.*

Preuve. Il est clair que $\varnothing$ et $\mathbb{R}$ satisfont les deux conditions simultanément. Soit $\mathcal{U}\subset\mathbb{R}$ un ouvert non-vide distinct de $\mathbb{R}$. Par le Théorème 12.4, $\mathcal{U}$ est une réunion dénombrable d'intervalles ouverts comme comme $\mathcal{U}$ n'est pas tout $\mathbb{R}$, un de ces intervalles possède un point d'accumulation qui n'est pas dans $\mathcal{U}$ et donc $\mathcal{U}$ ne peut pas être fermé. $\square$

Nous concluons cette section en donnant des extensions à des ensembles fermés et bornés de divers résultats rencontrés plus tôt.

PROPOSITION 12.11. *Soit $\mathcal{F}$ ensemble fermé dans $\mathbb{R}$ et $\{a_n\}$ une suite dans $\mathcal{F}$ qui converge vers $a \in \mathbb{R}$. Alors $a \in \mathcal{F}$.*

Preuve. On peut supposer $a_n \neq a$ $(\forall n \in \mathbb{N})$ sinon c'est fini. Alors tout intervalle ouvert J_a autour de a contient une infinité de a_n car $\{a_n\} \to a$. Donc a est point d'accumulation de $\mathcal{F}$ et ainsi $a \in \mathcal{F}$ par le Théorème 12.9. $\square$

THÉORÈME 12.12. *Si $\{\mathcal{F}_n\}$ est une suite décroissante d'ensembles fermés et bornés dans $\mathbb{R}$, alors $\mathcal{F} = \bigcap_{n=1}^{\infty} \mathcal{F}_n$ est non-vide.*

Preuve. La preuve de ce résultat qui généralise le Principe des segments emboîtés est laissée en exercice. $\square$

CHAPITRE 2

Séries infinies

En principe, ce chapitre n'est pas plus ardu que le précédent puisque l'étude de la convergence ou divergence des séries infinies n'est rien d'autre que l'étude de ce que l'on appellera la suite des sommes partielles de la série. On verra cependant au cours de ce chapitre que les séries infinies présentent aux mathématiciens des suites de réels dont la limite peut parfois être bien difficile à calculer. La difficulté vient du fait qu'il peut être très ardu, et parfois contre-intuitif, d'évaluer la somme obtenue par addition d'un nombre infini de termes. Ainsi, par exemple, la suite $\{a_n\} = \{\frac{1}{n}\}$ converge vers 0 mais l'on verra au cours du chapitre que l'expression

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

n'a aucun sens mathématique. Ou encore que l'expression

$$\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$$

peut être réarrangée pour représenter un nombre réel quelconque ! Ceci et bien d'autres subtilités rendent l'étude des séries un des chapitres difficiles du cours, mais il s'agit d'une passage obligatoire pour comprendre la notion de séries de fonctions tant utilisée dans de nombreux domaines des mathématiques pures et appliquées..

1. Sommes partielles de séries infinies

Etant donné une suite $\{a_n\}$ dans $\mathbb{R}$, on aimerait dans ce chapitre donner un sens mathématique à des expressions de la forme

$$a_1 + a_2 + a_3 + ... + a_n + a_{n+1} + ...$$

Comme il s'agit là d'un somme d'une infinité de termes, bien qu'on puisse l'écrire on ne sait pas si cette formule correspond à un nombre réel. Par exemple, en considérant la suite

$\{a_n\} = 1$, on a

$$\begin{aligned} a_1 + a_2 + a_3 + ... + a_k + ... &= 1 + 1 + 1 + ... + 1 + ... \\ &> k \qquad (\forall k \in \mathbb{N}), \end{aligned}$$

et donc il n'existe pas de réel $\alpha \in \mathbb{R}$ correspondant à cette somme infinie par la propriété archimédienne des réels.

Un exemple encore plus frappant est donné par $\{a_n\} = \{(-1)^n\}$, pour laquelle on écrit formellement

$$a_1 + a_2 + a_3 + ... + a_n + ... = -1 + 1 - 1 + ...$$

En regroupant les termes par paires en commençant par les deux premiers, cette expression semble alors exprimer la somme infinie

$$0 + 0 + 0 + 0 + ... = 0$$

Mais si l'on laisse plutôt le premier terme et que l'on commence ensuite à somme les termes par paires, on semble plutôt exprimer la somme infinie

$$-1 + 0 + 0 + 0 + ... = -1$$

On constate donc que l'on a un problème de taille pour donner un sens *réel* à l'expression $a_1 + a_2 + a_3 + ...$ dans cet exemple ! Le but premier sera donc de résoudre ce genre de problème.

DÉFINITION 1.1. *On appelle série infinie associée à une suite $\{a_n\}$ l'expression*

$$a_1 + a_2 + a_3 + ... + a_n + ... = \sum_{n=1}^{\infty} a_n.$$

Pour étudier une série infinie $\sum_{n=1}^{\infty} a_n$ il convient de définir sa *suite des sommes partielles* $\{s_n\}$ définie par

$$s_1 = a_1 \ , \ s_2 = a_1 + a_2 \ , \ , ... \ , \ s_n = a_1 + a_2 + ... + a_n,$$

dont le $n^{\text{ième}}$ terme est tout simplement la somme des n premiers termes de la suite $\{a_n\}$.

EXEMPLE 1.2. *Considérons la série $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + ... + \frac{1}{2^n} + ...$ La suite des sommes partielles associée à cette série infinie est calculable comme somme d'une progression géométrique*

(faites le calcul explicite)

$$s_n = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + ... + \frac{1}{2^n} = \frac{2^n - 1}{2^n}.$$

La suite des sommes partielles $\{s_n\}$ peut être étudiée avec les outils développés au cours du chapitre précédent. Cette suite peut converger ou diverger et c'est là le point de départ de l'étude de la convergence des séries :

DÉFINITION 1.3. *On dit que la série* $\sum\limits_{n=1}^{\infty} a_n$ *converge si sa suite de sommes partielles* $\{s_n\}$ *est une suite convergente dans* $\mathbb{R}$. *La valeur de la série est alors définie par*

$$\sum_{n=1}^{\infty} a_n = s = \lim_{n\to\infty} s_n.$$

Par ailleurs, on dit que la série diverge si $\{s_n\}$ *ne converge pas.*

EXEMPLE 1.4. *Pour la série infinie* $\sum\limits_{n=1}^{\infty} (-1)^n$, *la suite des sommes partielles donne*

$$s_{2k} = 0 \qquad \textit{et} \qquad s_{2k+1} = -1,$$

si bien que la suite $\{s_n\}$ *ne converge pas et donc la série diverge.*

EXEMPLE 1.5. *La série* $\sum\limits_{n=1}^{\infty} n$ *diverge diverge car sa suite de sommes partielles satisfait*

$$s_n > n \ (\forall n \in \mathbb{N}),$$

et est donc clairement divergente.

EXEMPLE 1.6. *La série* $\sum\limits_{n=1}^{\infty} \frac{1}{2^n}$ *déjà introduite au cours de l'exemple 1.2 a pour suite de sommes partielles*

$$\{s_n\} = \{\frac{2^n - 1}{2^n}\}.$$

Montrons que $\lim\limits_{n\to\infty} s_n = 1$ *:*

$$\begin{aligned}
\lim_{n\to\infty} \frac{2^n - 1}{2^n} &= \lim_{n\to\infty} \frac{2^n(1 - 1/2^n)}{2^n} \\
&= \lim_{n\to\infty} 1 - \frac{1}{2^n} \\
&= 1 - \lim_{n\to\infty} \frac{1}{2^n}
\end{aligned}$$

Comme on a $0 < \frac{1}{2^n} < \frac{1}{n}$ *pour tout* $n \in \mathbb{N}$*, on obtient bien le résultat. On peut alors écrire en toute rigueur*

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

Proposition 1.7. *Si la série* $\sum\limits_{n=1}^{\infty} a_n$ *converge, alors en particulier la suite* $\{a_n\}$ *converge vers* 0

Preuve. On démontre la contraposée de la proposition. Supposons au contraire que $\{a_n\} \nrightarrow 0$: il existe alors $\alpha > 0$ tel que pour tout $N \in \mathbb{N}$ il existe $n > N$ avec $|a_N| \geq \alpha$. Alors comme on a, par la définition de la suite des sommes partielles,

$$a_n = s_n - s_{n-1},$$

on obtient

$$|s_n - s_{n-1}| \geq \alpha$$

et donc la suite $\{s_n\}$ n'est pas suite de Cauchy et ne peut, par conséquent, converger. La série diverge donc tel que voulu. □

La dernière proposition est un premier outil pour montrer qu'une série donnée ne converge pas : on vérifie que la suite $\{a_n\}$ apparaissant dans la définition de la série ne converge pas vers 0 et on invoque la proposition précédente pour conclure que la série diverge.

Mise en garde. La proposition ne dit pas que si la suite $\{a_n\}$ tend vers 0 alors la série converge ! Ce dernier énoncé constitue la réciproque et rien dans preuve suggère qu'une telle réciproque est vraie. On demande aux lecteurs et aux lectrices de porter une attention particulière à ceci. Remarquons que si on arrivait à démontrer que $\{a_n\} \to 0$ implique que sa série converge, alors on aurait *complètement réduit* l'étude de la convergence des séries à la convergence des suites vers 0 (réfléchissez quelques instants sur cette dernière affirmation). Ce n'est malheureusement pas le cas - comme on le verra dans le prochain théorème - et c'est ce qui rend la théorie des séries infinies à la fois complexe et intéressante.

Nous énonçons le prochain résultat comme un théorème important même s'il ne s'agit en fait que de la construction d'un exemple. Ceci est fait dans le but de bien faire comprendre au lecteur ou à la lectrice que la réciproque de la Proposition (1.7) est fausse en général.

Il y a fort à parier que malgré toutes ces précautions, il y aura toujours des étudiants qui affirmeront un jour, en examen par exemple, qu'une série donnée **doit** converger car son terme général converge vers 0. La psychologie de l'apprentissage est parfois obscure...

THÉORÈME 11. *Il existe des séries infinies* $\sum_{n=1}^{\infty} a_n$ *qui ne convergent pas, bien que* $\{a_n\}$ *converge vers* 0.

Preuve. On construit tout simplement un exemple d'une série $\sum_{n=1}^{\infty} a_n$ qui diverge mais pour laquelle $\{a_n\} \to 0$. Considérons la **série harmonique** :

$$\sum_{n=1}^{\infty} \frac{1}{n}.$$

Regardons sa suite de sommes partielles

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

On montre que cette suite réelle diverge. Premièrement on montre par induction que

$$s_{2^k} > \frac{k}{2} \ (\forall k \in \mathbb{N}).$$

Le cas $k = 1$ est clair puisque $s_2 = 1 + 1/2 > 1/2$. Supposons le résultat vrai pour l'étape k. On a alors

$$s_{2^{k+1}} = s_{2^k} + [\frac{1}{2^k + 1} + \frac{1}{2^k + 2} + ... + \frac{1}{2^{k+1}}].$$

Comme le membre de gauche est une somme de 2^{k+1} termes et que s_{2^k} est somme de 2^k termes, on en déduit que l'expression entre crochets dans le membre droit de l'équation contient également 2^k termes dans la somme. En combinant ceci à l'hypothèse d'induction, on a alors l'inégalité suivante

$$s_{2^{k+1}} > \frac{n}{2} + \frac{1}{2^{k+1}} \cdot 2^k = \frac{n+1}{2},$$

ce qui complète l'induction.

On a donc, en particulier, que la suite $\{s_n\}$ n'est pas bornée et donc elle ne peut pas converger dans $\mathbb{R}$ (on rappelle qu'une suite qui converge doit, en particulier, être borné). La série harmonique est dont une série qui diverge dont le terme général tend néanmoins vers 0. $\square$

On se souvient de la propriété remarquable des suites dans $\mathbb{R}$: une suite converge dans $\mathbb{R}$ si et seulement si elle est suite de Cauchy. En appliquant ceci à la suite des sommes partielles d'une série infinie, on obtient ainsi une condition nécessaire et suffisante pour la convergence des séries :

PROPOSITION 1.8. *Une série* $\sum_{n=1}^{\infty} a_n$ *converge si et seulement si elle satisfait*

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ t.q.\ \forall n > 0, \forall p > 0\ \textit{on ait}\ |a_{n+1} + a_{n+2} + ... + a_{n+p}| < \epsilon.$$

2. Séries à termes positifs

Dans cette section on s'intéresse à la convergence des séries $\sum_{n=1}^{\infty} a_n$ pour lesquelles la condition $a_n \geq 0$ est vérifiée pour tout $n \in \mathbb{N}$. Cette condition impose certaines restrictions au comportement de la série et nous permettra d'obtenir des critères de convergence variés. Ce qui rend une série à termes positifs intéressante est le fait qu'une telle série possède alors forcément une suite de sommes partielles $\{s_n\}$ qui est *croissante*. Si l'on se souvient que l'on a montré au Chapitre 1 (Théorème 7) que toute suite croissante et bornée converge, on obtient alors un critère de convergence pour les séries positives :

PROPOSITION 2.1. *Soit* $\sum_{n=1}^{\infty} a_n$ *une série à termes positifs. Alors la série converge si et seulement sa suite des sommes partielles* $\{s_n\}$ *est bornée.*

Preuve. La suite des sommes partielles est croissante par l'hypothèse que la série est à termes positifs. La preuve découle alors imédiatement de la Proposition 8.10 et du Théorème 7. □

EXEMPLE 2.2. *Considérons la série* $\sum_{n=1}^{\infty} \frac{1}{n!}$. *On se souvient que lors de notre étude du nombre* $e = \lim_{n\to\infty}(1 + 1/n)^n$ *dans l'exemple 8.19 au cours du Chapitre 1, on avait établi l'inégalité*

$$e \geq \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + ... + \frac{1}{n!}\ (\forall n \in \mathbb{N}).$$

On en déduit donc pour notre série infinie à termes positifs que la suite de ses sommes partielles est bornée par le nombre réel e *et conséquemment*

$$\sum_{n=1}^{\infty} \frac{1}{n!}\ \textit{est une série convergente.}$$

En reprenant les détails de l'exemple 8.19 quelle est la valeur de cette série ?

EXEMPLE 2.3. *(Les séries géométriques)*
On appelle série géométrique de terme constant $a > 0$ et de raison $r > 0$ la série exprimée comme

$$\sum_{n=0}^{\infty} ar^n.$$

On montre qu'une telle série converge si et seulement si $r < 1$. En effet, d'une part on remarque que pour $r \geq 1$, la suite $\{ar^n\}$ ne converge pas vers 0 car on a $ar^n \geq a$. Ceci démontre une des deux implications. Réciproquement, si l'on a $r < 1$, alors on peut écrire

$$s_n = a + ar + ar^2 + ... + ar^{n-1}$$

$$rs_n = ar + ar^2 + ar^3 + ... + ar^{n-1} + ar^n$$

En soustrayant la seconde ligne de la première on a

$$(1-r)s_n = a - ar^n \iff s_n = a\frac{1-r^n}{1-r}$$

Comme on a $r < 1$, on sait que $\{r^n\} \to 0$ et donc

$$\lim_{n\to\infty} s_n = \frac{a}{1-r}$$

Et ainsi la série $\sum\limits_{n=1}^{\infty} ar^n$ converge vers $\frac{a}{1-r}$.

On se souvient qu'une des façons de démontrer la convergence d'une suite était de la comparer à des suites que l'on connait mieux, en utilisant, par exemple, le théorème des gendarmes. On peut appliquer le même genre d'idées dans le contexte des séries infinies, afin de comparer une série donnée à d'autres que l'on connait mieux. Ceci se résume sous la forme du test de comparaison suivant :

THÉORÈME 12. **(Test de comparaison)**

(1) *Soit $\{a_n\}$ une suite positive telle que la série $\sum\limits_{n=1}^{\infty} a_n$ converge et supposons que $\{b_n\}$ est une autre suite positive pour laquelle on ait un $M > 0$ tel que*

$$b_n \leq Ma_n \ (\forall n \in \mathbb{N}).$$

Alors la série $\sum\limits_{n=1}^{\infty} b_n$ converge également.

(2) *Soit $\{a_n\}$ une suite positive telle que la série $\sum\limits_{n=1}^{\infty} a_n$ diverge et supposons que b_n est une autre suite positive pour laquelle on ait un $M > 0$ tel que*

$$b_n \geq Ma_n \ (\forall n \in \mathbb{N}).$$

Alors la série $\sum_{n=1}^{\infty} b_n$ *diverge également.*

Preuve.

(1) Soient

$$\{s_n\} \text{ suite des sommes partielles de } \{a_n\}$$
$$\{t_n\} \text{ suite des sommes partielles de } \{b_n\}.$$

Comme il s'agit de suites positives, on sait que $\{s_n\}$ et $\{t_n\}$ sont des suites croissantes et l'on a, par l'hypothèse de convergence de la série de $\{a_n\}$ que $\lim_{n\to\infty} s_n = s$. On a alors les inégalités

$$t_n \leq M s_n \leq M s,$$

en utilisant le fait que $b_n \leq M a_n$. Mais alors $\{t_n\}$ est suite croissante et bornée, si bien qu'elle converge et ainsi la série $\sum_{n=1}^{\infty} b_n$ converge.

(2) Si $\sum_{n=1}^{\infty} a_n$ diverge, on sait alors que $\{s_n\}$ est suite non-bornée (on utilise ici crucialement le fait que $\{a_n\}$ soit positive) et alors l'inégalité

$$t_n \geq M s_n$$

implique que $\{t_n\}$ n'est pas bornée également et donc elle ne peut converger, c'est-à-dire que $\sum_{n=1}^{\infty} b_n$ diverge.

□

On devrait remarquer ici que les hypothèses impliquant les inégalités dans l'énoncé du dernier théorème peuvent être plus faibles. En effet, on peut remplacer la condition "pour tout $n \in \mathbb{N}$" par une condition telle "pour tout $n > N$ pour une certaine constante $n \in \mathbb{N}$". En effet, si la série

$$\sum_{n=N+1}^{\infty} b_n$$

converge, alors on aura également

$$\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{N} b_n + \sum_{n=N+1}^{\infty} b_n$$

convergente. Même chose pour la divergence.

Le dernier théorème pose tout naturellement la question de l'éventuelle existence d'une série convergente à termes positifs $\{a_n\}$ que l'on pourrait utiliser par comparaison pour *toutes* les séries convergentes :

$$\sum_{n=1}^{\infty} b_n \text{ converge } \iff \exists M > \text{tel que } b_n \leq M a_n \ (\forall n \in \mathbb{N}).$$

La théorie serait alors remarquablement simple, du moins en principe, car on n'aurait qu'à comparer une série à étudier à cette série universelle pour étudier sa convergence... Malheureusement, on montre maintenant que ce n'est point le cas.

Définition 2.4. *On dit qu'une série* $\sum_{n=1}^{\infty} b_n$ *converge plus lentement que la série* $\sum_{n=1}^{\infty} a_n$ *si les deux séries convergent et que les suites* $\{a_n\}$ *et* $\{b_n\}$ *satisfont*

$$\lim \frac{b_n}{a_n} = +\infty,$$

c'est-à-dire que pour tout $M > 0$ *il existe* $N \in \mathbb{N}$ *tel que*

$$n > N \Rightarrow \frac{b_n}{a_n} > M.$$

Proposition 2.5. *Etant donné n'importe quelle série convergente à termes positifs, il existe une série convergeant plus lentement.*

Preuve. La preuve n'est pas sans subtilités. Soit $\sum_{n=1}^{\infty} a_n$ convergente avec valeur de la série égale à s. Alors on a

$$\forall \epsilon > 0 \ \exists n \in \mathbb{N} \text{ tel que } |a_1 + a_2 + ... + a_n - s| < \epsilon.$$

En particulier, pour chaque $k \in \mathbb{N}$, on peut trouver $n_k \in \mathbb{N}$ tel que

$$|a_1 + a_2 + ... + a_{n_k} - s| < \frac{1}{k2^k} \tag{1}$$

Comme $\{a_n\}$ est une suite positive (et donc la suite des sommes partielles est croissante), ceci permet de construire

$$n_1 < n_2 < ... < n_k < ...$$

satisfaisant l'équation (1). Posons alors

$$b_n = \begin{cases} a_n & \text{si } n \leq n_k \\ k a_n & \text{si } n_k + 1 \leq n \leq n_{k+1}. \end{cases}$$

On a alors

$$\begin{aligned}|a_{n_k+1}+a_{n_k+2}+...+a_{n_{k+1}}| &= |a_1+...+a_{n_k}+a_{n_k+1}+...+a_{n_{k+1}}-s-(a_1+...+a_{n_k}-s)|\\ &< \frac{1}{(k+1)2^{k+1}}+\frac{1}{k2^k}<\frac{1}{k2^k}+\frac{1}{k2^k}=\frac{1}{k2^{k-1}}.\end{aligned}$$

Alors par la construction de $\{b_n\}$, on obtient

$$|b_{n_k+1}+b_{n_k+2}+...+b_{n_{k+1}}|<\frac{1}{2^{k-1}}.$$

On en déduit alors facilement que la suite des sommes partielles de $\sum_{n=1}^{\infty} b_n$ est suite de Cauchy et donc la série converge tel que requis. Par ailleurs, par la définition de $\{b_n\}$, on a

$$\lim_{n\to\infty}\frac{b_n}{a_n}\geq\frac{ka_n}{a_n}=k\ (\forall k\in\mathbb{N}),$$

et donc la série de $\{b_n\}$ converge plus lentement que la série de $\{a_n\}$. □

3. Tests de convergence pour les séries

Nous développons ici des critères de nature plus pratique pour pouvoir évaluer si une série donnée de manière explicite converge ou diverge. Il est à noter qu'il est bon d'avoir plusieurs tests de convergence, puisque selon le type d'exemple traité certains tests seront plus aisément appliqués que d'autres. Il est important de bien se souvenir des hypothèses nécessaires pour l'application de chacun de ces tests de convergence.

Théorème 13. *(Test de la racine de Cauchy)*

(1) *Soit $\sum_{n=1}^{\infty} a_n$ une série à termes positifs pour laquelle il existe $N\in\mathbb{N}$ et $\alpha<1$ tels que*

$$n>N\Rightarrow\sqrt[n]{a_n}<\alpha.$$

Alors la série converge.

(2) *Si au contraire $\sqrt[n]{a}_n\geq 1$ pour un nombre infini de valeurs de n, alors la série diverge.*

Preuve. La seconde assertion est claire puisque l'on a alors $a_n\geq 1$ pour une infinité de termes et donc la suite $\{a_n\}$ ne peut converger vers 0, une condition nécessaire pour la

convergence de sa série.
Quant à la première affirmation, on a pour $n > N$ que la suite $\{a_n\}$ satisfait

$$a_n < \alpha^n.$$

Comme la série géométrique tronquée $\sum_{n=N+1}^{\infty} \alpha^n$ converge pour $\alpha < 1$, on obtient par le Théorème 12 que la série $\sum_{n=1}^{\infty} a_n$ converge également (pourquoi les N premiers termes n'importent pas ?). □

L'analogue de la Proposition 8.17 pour les séries donne un autre test utile pour examiner la convergence de certaines séries :

THÉORÈME 14. *(Test du quotient de d'Alembert)*

(1) *Si* $\sum_{n=1}^{\infty} a_n$ *est une série à termes positifs pour laquelle il existe un* $N \in \mathbb{N}$ *et* $\alpha < 1$ *tels que*

$$n > N \Rightarrow \frac{a_{n+1}}{a_n} < \alpha,$$

alors la série converge.

(2) *Si au contraire, à partir de on a* $\beta > 1$ *tel que*

$$n > N \Rightarrow \frac{a_{n+1}}{a_n} > \beta,$$

alors la série diverge.

Preuve. L'idée de base est très simple : on veut comparer le comportement de notre série à une série géométrique. Dans le premier cas ($\alpha < 1$), exactement comme dans la preuve de la Proposition (8.17), on obtient pour notre série à termes positifs

$$a_1 + a_2 + \ldots + a_N + a_{N+1} + \ldots + a_n + \ldots \leq a_1 + a_2 + \ldots + a_N + a_N\alpha + a_N\alpha^2 + \ldots + a_N\alpha^{n-N} + \ldots$$

Notre série $\sum_{n=1}^{\infty} a_n$ converge car elle est majorée par une série qui converge, constituée du terme constant $a_1 + a_2 + \ldots + a_N$ et de la série géométrique tronquée de raison $\alpha < 1$.

Le second cas est encore plus facile puisque l'hypothèse sur la suite $\{a_n\}$ nous permet d'obtenir, pour N assez grand afin que l'on ait

$$\frac{a_{n+1}}{a_n} > \beta > 1 \ (\forall n > N),$$

l'inégalité suivante :

$$a_n \geq \beta^{n-N} a_N.$$

En particulier on voit alors que la suite $\{a_n\}$ ne peut converger vers 0, ce qui exclut la possibilité de convergence de sa série tel que voulu. $\square$

EXEMPLE 3.1. *Considérons la série* $\sum_{n=1}^{\infty} \frac{n^4}{2^n}$. *On applique le test du quotient pour étudier la convergence de cette série :*

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)^4/2^{n+1}}{n^4/2^n} = \frac{1}{2}\Big(\frac{n+1}{n}\Big)^4.$$

Comme on a $\lim_{n\to\infty} \frac{n+1}{n} = 1$, *alors* $\lim_{n\to\infty} \left(\frac{n+1}{n}\right)^4 = 1$. *Il en découle que pour* $\alpha = \frac{2}{3}$ *il existe un* $N \in \mathbb{N}$ *tel que pour* $n > N$ *quelconque, on trouve pour tout* $n > N$,

$$\frac{1}{2} < \frac{(n+1)^4/2^{n+1}}{n^4/2^n} < \frac{2}{3},$$

ce qui donne la convergence de notre série.

Digression. On introduit ici la notion de limite supérieure et inférieure.

Si $\{a_n\}$ est une suite de réels, posons

$$\mathcal{A}(a_n) = \{x \in \mathbb{R} \mid x \text{ est point d'accumulation de } \{a_n\}\}.$$

Une autre façon de voir cet ensemble $\mathcal{A}(a_n)$ est de dire qu'il s'agit de l'ensemble des limites des sous-suites de $\{a_n\}$. Par exemple, si on a que $\{a_n\} \to a$, alors $\mathcal{A}(a_n) = \{a\}$ car toute sous-suite de $\{a_n\}$ doit converger vers a, donc c'est l'unique point d'accumulation de $\{a_n\}$. Par contre, pour la suite $\{b_n\} = \{(-1)^n\}$, on voit facilement que $\mathcal{A}(a_n) = \{-1, +1\}$. Le lecteur ou la lectrice pourra s'amuser à construire des suites dont l'ensemble des points d'accumulation $\mathcal{A}(a_n)$ est infini.

DÉFINITION 3.2. *On appelle limite supérieure, respectivement limite inférieure, de la suite* $\{a_n\}$ *les nombres*

$$\limsup_{n\to\infty} a_n = \sup \mathcal{A}(a_n)$$
$$\liminf_{n\to\infty} a_n = \inf \mathcal{A}(a_n)$$

PROPOSITION 3.3. *On a les caractérisations suivantes pour les notions de limite supérieure et inférieure :*

(1) *Le nombre* $\limsup_{n\to\infty} a_n$ *est l'unique réel tel que*
   (a) $\limsup_{n\to\infty} a_n \in \mathcal{A}(a_n)$
   (b) *Si* $x > \limsup_{n\to\infty} a_n$ $\exists N \in \mathbb{N}$ *tel que* $n > N \Rightarrow a_n < x$.
(2) *Le nombre* $\liminf_{n\to\infty} a_n$ *est l'unique réel tel que*
   (a) $\liminf_{n\to\infty} a_n \in \mathcal{A}(a_n)$
   (b) *Si* $x < \liminf_{n\to\infty} a_n$ $\exists N \in \mathbb{N}$ *tel que* $n > N \Rightarrow a_n > x$.

Preuve. Voir le livre de Rudin, pour la culture générale plus que pour le contenu de ce cours. □

Ayant maintenant introduit ces notions de limite supérieure et inférieure lors de la disgression, on peut alors donner des versions plus générales des tests de la racine de Cauchy et du quotient de d'Alembert :

Test de la racine : Soit $\alpha = \limsup_{n\to\infty} \sqrt[n]{a_n}$. Il y a alors deux cas de figure pour conclure :

$$\text{Si } \alpha < 1 \text{ , alors } \sum_{n=1}^{\infty} a_n \text{ converge.}$$
$$\text{Si } \alpha > 1 \text{ , alors } \sum_{n=1}^{\infty} a_n \text{ diverge.}$$

Test du quotient : Soient $\alpha = \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}$ et $\beta = \liminf_{n\to\infty} \frac{a_{n+1}}{a_n}$. Il y a lors deux cas de figure pour conclure :

$$\text{Si } \alpha < 1 \text{ , alors } \sum_{n=1}^{\infty} a_n \text{ converge.}$$
$$\text{Si } |\frac{a_{n+1}}{a_n}| \geq 1, \text{ pour tout } n > N_0, \text{alors } \sum_{n=1}^{\infty} a_n \text{ diverge.}$$
$$\text{Si } \beta \leq 1 \leq \alpha, \text{ on ne peut rien conclure.}$$

Les preuves de ces versions généralisées sont essentiellement les mêmes que celles données préalablement, alors le lecteur ou la lectrice pourra suppléer les détails en exercice. Une question qu'il est naturel de se poser est la force relative de ces deux tests. Il arrive souvent que le test du quotient de d'Alembert soit plus facile à appliquer que le test de la racine de Cauchy, ce qui est un facteur important lorsque l'on doit choisir lequel utiliser. Cependant, ce dernier test *domine* le test du quotient dans la mesure où :

(1) Si le test du quotient donne la convergence, alors le test de la racine donnera également la convergence.

(2) Si le test de la racine ne permet pas de conclure, alors le test du quotient ne permettra pas de conclure également.

Ceci est établi en démontrant les inégalités suivantes dont on trouvera la preuve dans le livre de Rudin :

$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n\to\infty} \sqrt[n]{a_n}.$$

$$\limsup_{n\to\infty} \sqrt[n]{a_n} \leq \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

Dans la théorie des séries infinies, il y a de nombreux autres tests que l'on n'a pas forcément le loisir d'étudier dans un cours d'introduction à l'Analyse réelle. Certains sont particulièrement subtils et ont été développés parce que ceux que l'on a introduit ci-haut ne fonctionnaient pas, même dans certains exemples simples à énoncer. Ainsi, par exemple, pour la famille de séries

$$\sum_{n\to\infty} \frac{1}{n^{\xi}}$$

dépendant d'un paramètre réel ξ, tant le test de la racine que le test du quotient ne donnent pas de résultat concluant (faites l'exercice de calcul de limite dans les deux cas). On passera donc le reste de cette section à développer d'autres tests complémentaires, qui seront essentiellement des variations de ceux déjà rencontrés au cours du chapitre.

Théorème 3.4. *Soient* $\sum_{n=1}^{\infty} a_n$ *et* $\sum_{n=1}^{\infty} b_n$ *deux séries à termes strictement positifs. Supposons qu'il existe* $N \in \mathbb{N}$ *tel que*

$$n > N \Rightarrow \frac{a_{n+1}}{a_n} \leq \frac{b_{n+1}}{b_n}.$$

(1) *Si* $\sum_{n=1}^{\infty} a_n$ *diverge, alors* $\sum_{n=1}^{\infty} b_n$ *diverge.*

(2) *Si* $\sum_{n=1}^{\infty} b_n$ *converge, alors* $\sum_{n=1}^{\infty} a_n$ *converge.*

Preuve. Les deux énoncés sont en fait équivalents d'un point de vue logique, alors on ne montre que (2). On a pour $n > N$:

$$\frac{a_n}{a_N} = \frac{a_{N+1}}{a_N}\frac{a_{N+2}}{a_{N+1}}\frac{a_{N+3}}{a_{N+2}}\cdots\frac{a_n}{a_{n-1}} = \prod_{k=N}^{n-1}\frac{a_{k+1}}{a_k}$$
$$\leq \prod_{k=N}^{n-1}\frac{b_{k+1}}{b_k} = \frac{b_n}{b_N}.$$

Et donc pour tout $n > N$ on a

$$a_n \leq \frac{a_N}{b_N} b_n$$

et par le théorème de comparaison on obtient bien la convergence de la série $\sum_{n=1}^{\infty} a_n$. □

Exemple 3.5. *Soit $\{a_n\}$ une suite telle que pour $n > N$ on ait*

$$\frac{a_{n+1}}{a_n} \leq \left(\frac{n}{n+1}\right)^\alpha \text{ pour } \alpha > 1.$$

Remarquons que si la limite du quotient existe, alors elle satisfait

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n} \leq 1$$

mais il est possible que l'on ne puisse conclure par le test de d'Alembert (cas où la limite serait exactement 1). On a alors

$$\frac{a_{n+1}}{a_n} \leq \left(\frac{1/n+1}{1/n}\right)^\alpha$$

et par le théorème précédent on pourra montrer que $\sum_{n=1}^{\infty} a_n$ converge si l'on sait que $\sum_{n=1}^{\infty}\frac{1}{n^\alpha}$ converge pour $\alpha > 1$, ce que nous démontrerons très bientôt.

Théorème 3.6. *Soit $\{a_n\}$ une suite décroissante et positive. Alors*

$$\sum_{n=1}^{\infty} a_n \text{ converge} \iff \sum_{n=1}^{\infty} 2^n a_{2^n} \text{ converge.}$$

Preuve. Considérons la suite des sommes partielles $\{s_n\}$ de la série. Alors

$$s_{2^{n+1}-1} = a_1 + (a_2 + a_3) + (a_4 + \dots + a_7) + \dots + (a_{2^n} + a_{2^n+1} + \dots + a_{2^{n+1}-1}).$$

Comme la suite $\{a_n\}$ est décroissante, on a les inégalités

$$2^k a_{2^{k+1}} \leq a_{2^k} + a_{2^k+1} + \dots + a_{2^{k+1}-1} \leq 2^k a_{2^k} \text{ pour } k \in \mathbb{N}.$$

Mais alors en combinant ceci et l'expression ci-dessus pour $s_{2^{n+1}-1}$, on obtient pour la suite des sommes partielles les deux inégalités

$$s_{2^{n+1}-1} \geq a_1 + \frac{1}{2}\sum_{k=2}^{n+1} 2^k a_{2^k} \qquad\qquad s_{2^{n+1}-1} \leq a_1 + \sum_{k=1}^{n} 2^k a_{2^k}.$$

Le théorème de comparaison et ces deux inégalités impliquent directement le résultat. □

EXEMPLE 3.7. *Etudions le comportement de la série dépendant d'un paramètre $\alpha \in \mathbb{R}$ fixe :*

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}.$$

On sait déjà traiter un cas : $\alpha = 1$ donne la série harmonique dont on a établi la divergence. Le cas $\alpha < 1$ est alors facile puisque l'on peut comparer à la série harmonique :

$$n^\alpha < n \ (\forall n \in \mathbb{N}) \Rightarrow \frac{1}{n^\alpha} > \frac{1}{n} \ (\forall n \in \mathbb{N})$$

et on obtient la divergence de la série pour tout $\alpha < 1$.
Dans le cas où $\alpha > 1$, on applique le théorème précédent : considérons la suite

$$\{b_n\} = \{2^n \frac{1}{(2^n)^\alpha}\}.$$

On a alors

$$b_n = \frac{2^n}{(2^n)^\alpha} = \big(\frac{1}{2^{\alpha-1}}\big)^n.$$

Comme $\alpha > 1$, on a que $1 < 2^{\alpha-1}$ si bien que $\{b_n\}$ est suite géométrique de raison $\frac{1}{2^{\alpha-1}} < 1$ et pour laquelle $\sum\limits_{n=1}^{\infty} b_n$ converge. Le théorème précédent implique ainsi que $\sum\limits_{n=1}^{\infty} \frac{1}{n^\alpha}$ converge si $\alpha > 1$.

4. Séries à termes quelconques

Jusqu'à présent nous avons toujours imposé certains comportements à la suite $\{a_n\}$ dont on étudiait la convergence de la série associée (suite positive, décroissante, etc). Dans cette section, nous considèrons le problème de convergence sans aucune hypothèse de départ sur la suite $\{a_n\}$. On se souvient que deux résultats fondamentaux ont déjà été établis dans ce cadre tout-à-fait général. Tout d'abord une condition nécessaire pour avoir la convergence de la série :

$$\sum_{n=1}^{\infty} a_n \text{ converge } \Rightarrow \{a_n\} \to 0.$$

Ensuite une caractérisation de Cauchy de la convergence des séries :

$$\sum_{n=1}^{\infty} a_n \text{ converge } \iff \forall \epsilon > 0\, \exists N \in \mathbb{N} \text{ t.q. } \forall n > 0, \forall p > 0 \text{ on ait } |a_{n+1}+a_{n+2}+...+a_{n+p}| < \epsilon.$$

Pour étudier une série quelconque $\sum_{n=1}^{\infty} a_n$ il sera bien utile d'utiliser les deux séries auxiliaires suivantes

$$\sum_{n=1}^{\infty} p_n \qquad \text{et} \qquad \sum_{n=1}^{\infty} q_n,$$

où $\{p_n\}$ est sous-suite de termes *positifs* de $\{a_n\}$ alors que $\{q_n\}$ est sous-suite de termes *négatifs* de $\{a_n\}$. Posons alors pour la suite :

$$\{s_n\} \text{ : suite des sommes partielles de } \sum_{n=1}^{\infty} p_n$$

$$\{t_n\} \text{ : suite des sommes partielles de } \sum_{n=1}^{\infty} q_n$$

Nous distinguons alors 4 cas dans notre étude en fonction du comportement de ces suites de sommes partielles :

(I) $\{s_n\}$ et $\{t_n\}$ suites bornées.
(II) $\{s_n\}$ bornée et $\{t_n\}$ non-bornée.
(III) $\{s_n\}$ non-bornée et $\{t_n\}$ bornée.
(IV) $\{s_n\}$ et $\{t_n\}$ non-bornées.

Le concept que l'on va définir à l'instant est réellement fondamental, tant pour ce qui va suivre dans la section que pour le développement de la théorie générale des séries.

DÉFINITION 4.1. *On dit qu'une séries* $\sum_{n=1}^{\infty} a_n$ *converge absolument si la série* $\sum_{n=1}^{\infty} |a_n|$ *converge.*

On pourrait penser qu'il n'y a pas vraiment de lien entre les convergences des séries $\sum_{n=1}^{\infty} a_n$ et $\sum_{n=1}^{\infty} |a_n|$, mais en réalité le théorème suivant nous permettra souvent d'étudier en détail la convergence d'une série à partir de la série de valeurs absolues associée.

THÉORÈME 15. *Si une série* $\sum_{n=1}^{\infty} a_n$ *converge absolument, alors en particulier la série est convergente.*

Preuve. On utilise la caractérisation de Cauchy des séries convergentes ainsi que l'inégalité du triangle. Si la série $\sum_{n=1}^{\infty} |a_n|$ converge, alors pour $\epsilon > 0$ donné on a $N \in \mathbb{N}$ tel que pour tout $n > N$ et tout $p > 0$, on ait

$$||a_{n+1}| + |a_{n+2}| + ... + |a_{n+p}|| < \epsilon.$$

Mais alors comme

$$|a_{n+1} + a_{n+2} + ... + a_{n+p}| \leq |a_{n+1}| + |a_{n+2}| + ... + |a_{n+p}| < \epsilon,$$

on a bien que la série $\sum_{n=1}^{\infty} a_n$ converge par la caractérisation de Cauchy des séries convergentes. □

Le grand avantage à passer d'une suite donnée à la suite des valeurs absolues pour l'étude de la convergence des séries est que l'on a alors à étudier une série à termes positifs et l'on a développé plusieurs outils au cours de ce chapitre pour étudier de telles séries. Si la série des valeurs absolues converge, on peut conclure que la série elle-même converge, en vertu du dernier théorème. Le seul hic est le suivant : si la série des valeurs absolues ne converge pas, alors on ne peut conclure sans plus d'information... On verra en effet des exemples de séries convergentes qui ne sont pas *absolument* convergentes. Mais avant ceci, montrons le résultat suivant :

THÉORÈME 4.2. **(Théorème de Leibnitz)**
Soit $\{a_n\}$ une suite décroissante, positive et telle que $\{a_n\} \to 0$. Alors la série

$$\sum_{n=1}^{\infty} (-1)^{n+1} a_n$$

est une série convergente.

Preuve. La preuve est très jolie et ingénieuse. Montrons que la suite des sommes partielles de la série satisfait :

(1) $\{s_{2k}\}$ est croissante.
(2) $\{s_{2k+1}\}$ est croissante.
(3) $s_{2k} \leq s_{2l+1}$ pour tous $k, l \in \mathbb{N}$.

Pour (1), on remarque immédiatement que

$$s_{2n+2} = s_{2n} + a_{2n+1} - a_{2n+2} \geq s_{2n} \quad \text{car } \{a_n\} \text{ décroissante et positive,}$$

donc la suite $\{s_{2k}\}$ est bien croissante. Pour (2), on a de la même manière

$$s_{2n+3} = s_{2n+1} - a_{2n+2} + a_{2n+3} \leq s_{2n+1} \quad \text{car } \{a_n\} \text{ décroissante et positive,}$$

donc la suite $\{s_{2k+1}\}$ est croissante. Ces deux propriétés sont cruciales pour établir la troisième. Premièrement remarquons le cas particulier de (3) :

$$s_{2k} = s_{2k-1} - a_{2k} \leq s_{2k-1} \quad (\text{car } a_{2k} \geq 0).$$

Pour $2k$ et $2l+1$ donnés, soit alors $n \in \mathbb{N}$ tel que $2n \geq 2k$ et $2n-1 \geq 2l+1$. Alors par (1) et (2) on obtient

$$s_{2k} \leq s_{2n} \leq s_{2n-1} \leq s_{2l+1},$$

démontrant (3). Les propriétés (1) et (3) donnent alors que $\{s_n\}$ est croissante, bornée supérieurement si bien que l'on peut poser

$$\alpha = \sup\{s_{2n}\} = \lim_{n\to\infty}\{s_{2n}\}.$$

De même (2) et (3) donnent que $\{s_{2n+1}\}$ est décroissante, bornée inférieurement, si bien qu'on peut poser

$$\beta = \inf\{s_{2n+1}\} = \lim_{n\to\infty}\{s_{2n+1}\}.$$

On déduit facilement de (3) que $\alpha \leq \beta$ et comme $s_{2n+1} - s_{2n} = a_{2n+1}$ et $\lim\limits_{n\to\infty}\{a_n\} = 0$, on a en fait $\alpha = \beta$. Ceci démontre donc que $\lim\{s_n\}$ existe et donc $\sum\limits_{n=1}^{\infty}(-1)^{n+1}a_n$ converge. $\square$

Exemple 4.3. *La série* $\sum\limits_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$ *converge mais elle ne converge pas absolument. En effet, on peut utiliser le théorème de Leibnitz pour obtenir directement la convergence puisque* $\{\frac{1}{n}\}$ *est clairement suite décroissante et positive. Mais la série des valeurs absolues associée est nulle autre que la série harmonique dont on sait qu'elle diverge !*

En ayant en tête les 4 cas distingués ci-dessus pour le comportement des suites de sommes partielles pour les séries de termes positifs ($\{s_n\}$) et de termes négatifs ($\{t_n\}$), on montre ci-dessous que le seul cas pouvant mener à une convergence absolue de la série est le **Cas (I)** :

Proposition 4.4. *La série* $\sum\limits_{n=1}^{\infty} a_n$ *converge absolument si et seulement si* $\{s_n\}$ *et* $\{t_n\}$ *sont bornées.*

Preuve. Supposons premièrement que l'on est dans le Cas (I) : il existe $M, M' > 0$ tels que pour tout $n \in \mathbb{N}$ on ait

$$p_1 + p_2 + ... + p_n < M$$

$$-(q_1 + q_2 + ... + q_n) < M'$$

Alors on a également pour tout $n \in \mathbb{N}$

$$|a+1| + |a_2| + ... + |a_n| < M + M'.$$

Ainsi la suite des sommes partielles de la série $\sum_{n=1}^{\infty} |a_n|$ est croissante et bornée, donc elle converge et notre série converge absolument. La suite de la preuve est de montrer que dans les autres cas (II,III et IV) la série ne peut converger absolument. Pour ce qui est des Cas (II) et Cas (III), la série diverge absolument puisque l'on a les inégalités

$$\sum_{n=1}^{N} |a_n| \geq -\sum_{n=1}^{N} q_n \qquad \text{et} \qquad \sum_{n=1}^{N} |a_n| \geq \sum_{n=1}^{N} p_n,$$

et une de ces deux inégalités donnera (selon que l'on soit dans le Cas (II) ou le Cas (III)) que la suite des sommes partielles de $\sum_{n=1}^{\infty} |a_n|$ n'est pas bornée, si bien qu'elle ne peut converger. Le Cas (IV) sera traité séparément en élargissant notre étude. $\square$

Illustrons une des subtilités de la convergence des séries par un exemple qui, à prime abord, semble engendrer une contradiction mathématique... Considérons la série harmonique alternée

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n}.$$

On sait par le théorème de Leibnitz que cette série converge vers un nombre réel $x \in \mathbb{R}$. Effectuons les manipulations algébriques suivantes :

$$\begin{aligned} x &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + ... \\ &= 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \frac{1}{7} - \frac{1}{14} - \frac{1}{16} + ... \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{(1 terme positif suivi de 2 termes négatifs)} \\ &= (1 - \frac{1}{2}) - \frac{1}{4} + (\frac{1}{3} - \frac{1}{6}) - \frac{1}{8} + (\frac{1}{5} - \frac{1}{10}) - \frac{1}{12} + (\frac{1}{7} - \frac{1}{14}) - \frac{1}{16} + ... \\ &= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \frac{1}{14} - \frac{1}{16} + ... \\ &= \frac{1}{2} \cdot (1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + ...) \\ &= \frac{1}{2} \cdot x \end{aligned}$$

Ces manipulation algébriques mènent donc à la conclusion que $x = \frac{1}{2}x$ et donc que $x = 0$! Or dans la preuve du théorème de Leibnitz on avait que $x \geq s_2 = \frac{1}{2}$... D'où vient donc cette contradiction mathématique ? Si l'on regarde de plus près les manipulations effectuées ci-haut, on se rend compte que l'on a remanié une infinité de termes dans notre suite $\{(-1)^{n+1}\frac{1}{n}\}$ et avons supposé que ceci n'affecterait pas le calcul de la série. En fait, comme on le verra ci-dessous, ceci est une erreur...

DÉFINITION 4.5. *Soit $\phi \colon \mathbb{N} \to \mathbb{N}$ une application bijective. On dit que la suite $\{b_n\} = \{a_{\phi(n)}\}$ est un réarrangement de la suite $\{a_n\}$.*

Même si ceci peut sembler contre-intuitif au lecteur ou à la lectrice, si $\{b_n\}$ est un réarrangement de $\{a_n\}$, il n'y a absolument aucune raison pour que

$$\lim_{n\to\infty} b_1 + b_2 + ... + b_n \qquad \text{et} \qquad \lim_{n\to\infty} a_1 + a_2 + ... + a_n$$

soient toujours égales. On peut en fait avoir pour une permutation $\phi \colon \mathbb{N} \to \mathbb{N}$ que

$$\sum_{n=1}^{\infty} a_{\phi(n)} \neq \sum_{n=1}^{\infty} a_n.$$

Avant de montrer ceci, établissons que dans le cas de séries absolument convergentes, ceci ne peut se produire.

THÉORÈME 16. *Si $\sum\limits_{n=1}^{\infty} a_n$ converge absolument, alors tout réarrangement de la suite $\{a_n\}$ mène à une série qui converge absolument et aura la même valeur pour sa série.*

Preuve. Supposons que $\sum_{n=1}^{\infty} |a_n| = M$. Alors pour tout réarrangement $\{a_{\phi(n)}\}$ on a pour tout $n \in \mathbb{N}$ que

$$\sum_{k=1}^{n} |a_{\phi(k)}| \leq M,$$

si bien que la suite des sommes partielles de la série réarrangée converge et donc la série $\sum_{n=1}^{\infty} |a_{\phi(n)}|$ converge.

Soit $s = \sum_{n=1}^{\infty} a_n$ et prenons $\epsilon > 0$. Il existe alors $N \in \mathbb{N}$ tel que

$$|s - \sum_{k=1}^{N} a_k| < \frac{\epsilon}{2} \qquad \text{et} \qquad \sum_{k=N+1}^{\infty} |a_k| < \frac{\epsilon}{2},$$

par convergence de la série de $\{a_n\}$ et l'inégalité du triangle généralisée. Soit N' tel que $a_1, a_2, ..., a_N$ soient parmi $a_{\phi(1)}, a_{\phi(2)}, ..., a_{\phi(N')}$. Alors pour $n > N'$ on a

$$\begin{aligned} |s - \sum_{k=1}^{n} a_{\phi(k)}| &< |s - \sum_{k=1}^{N} a_k| + \sum_{k=N+1}^{\infty} |a_k| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Donc on a bien

$$\sum_{n=1}^{\infty} a_{\phi(n)} = s.$$

□

Nous rappelons que nous devons encore exclure le Cas (IV) ($\{s_n\}$ et $\{t_n\}$ non-bornées) dans la preuve de la Proposition 4.4. Pour ceci nous utilisons le fait qu'une série convergeant absolument verra tous ses réarrangements converger vers la *même* valeur. La preuve sera complète une fois que nous aurons démontré le résultat suivant qui est à la fois d'intérêt indépendant et plutôt spectaculaire !

THÉORÈME 17. **(Théorème de réarrangement de Riemann)**
Soit $\sum_{n=1}^{\infty} a_n$ *une série convergente mais pas absolument convergente. Alors pour tout réel* $\alpha \in \mathbb{R}$ *il existe un réarrangement* $\{a_{\phi(n)}\}$ *de* $\{a_n\}$ *tel que*

$$\sum_{n=1}^{\infty} a_{\phi(n)} = \alpha.$$

Preuve. La preuve est loin d'être immédiate, aussi le lecteur ou la lectrice aura intérêt d'en saisir avant tout la saveur géométrique avant de s'attaquer à retenir les détails. Soient $\sum_{n=1}^{\infty} p_n$ et $\sum_{n=1}^{\infty} q_n$ les sous-séries positive et négative respectivement. Comme la série d'origine ne converge pas absolument, on sait qu'au moins une de ces sous-séries diverge. En fait, les deux doivent diverger : il s'agit dans chaque cas de séries à termes de signe unique et donc elles divergent si et seulement si les suite de sommes partielles $\{s_n\}$ et $\{t_n\}$ sont non-bornées. Or si seulement une de ces deux suites de sommes partielles était non-bornée, on ne pourrait avoir que la suite des sommes partielles de $\sum_{n=1}^{\infty} a_n$ soit bornée et donc cette dernière série divergerait !

Supposons sans perte de généralité que $\alpha > 0$. Voici maintenant l'idée centrale de la preuve. Soit N_1 le plus petit entier tel que

$$\sum_{n=1}^{N_1} p_n > \alpha.$$

On sait qu'un tel N_1 existe car la série $\sum_{n=1}^{\infty} p_n$ diverge. On a alors par construction

$$\sum_{n=1}^{N_1-1} p_n \leq \alpha.$$

Posons $S_1 = \sum_{n=1}^{N_1} p_n$. Alors on trouve $S_1 - \alpha \leq p_{N_1}$. Soit M_1 le plus petit entier tel que

$$T_1 = S_1 + \sum_{n=1}^{M_1} q_n < \alpha.$$

Comme M_1 est le plus petit tel entier, on a

$$\alpha + q_{M_1} \leq S_1 + \sum_{n=1}^{M_1-1} q_n + q_{M_1} = T_1.$$

Et donc

$$\alpha - T_1 \leq -q_{M_1}.$$

On continue ce processus indéfiniment, obtenant des sommes

$$S_k = T_{k-1} + \sum_{n=N_{k-1}+1}^{N_k} p_n \qquad \text{et} \qquad T_k = S_k + \sum_{n=M_{k-1}+1}^{M_k} q_n,$$

respectivement supérieures et inférieures au nombre α, chaque fois choisissant N_k et M_k les plus petits possibles avec la propriété requise.

$\square$

CHAPITRE 3

Fonctions continues

1. Définition et quelques équivalences

Au cours du Chapitre 2 nous avons étudié en détail les suites de réels et vu qu'une suite est définie comme une fonction $f\colon \mathbb{N} \to \mathbb{R}$ avec $n \mapsto f(n) = a_n$. La convergence d'une suite $\lim\limits_{n\to\infty} a_n = \alpha$ a alors été rigoureusement interprétée comme

$$n \to +\infty \Rightarrow \{f(n)\} \to \alpha.$$

Il s'agit là d'un très bon point de départ pour comprendre la continuité de fonctions d'une variable réelle. Si $f\colon \mathbb{R} \to \mathbb{R}$ désigne maintenant une telle fonction, par analogie avec la notion de convergence des suites, on fera la définition suivante :

DÉFINITION 1.1. *On dit que $f\colon \mathbb{R} \to \mathbb{R}$ est continue au point $a \in \mathbb{R}$ si pour toute suite $\{x_n\} \subset \mathbb{R}$, on a*

$$\{x_n\} \to a \Rightarrow \{f(x_n)\} \to f(a).$$

Avec les conventions définies au chapitre sur les suites, on peut également alors écrire rigoureusement la condition de continuité comme

$$\lim_{x_n\to a} f(x_n) = f(a), \text{ pour toute suite } \{x_n\}.$$

De ce point de vue, la continuité des fonctions d'une variable réelle apparait comme une généralisation de la notion de limite d'une suite, ce qui es un avantage pour développer l'intuition. Un inconviénient de cette définition, par contre, est qu'elle demande de considérer *toutes* les suites convergeant ver a pour pouvoir démontrer la continuité de f en a.

Il est important de bien comprendre que la continuité est une notion *ponctuelle* : on parle de continuité *au point* $a \in \mathbb{R}$. Lorsque la condition de continuité est réalisée pour tout point $a \in A$ d'un ensemble $A \subseteq \mathbb{R}$, on peut alors dire que la fonction est continue *sur* A. On pourra, par exemple, parler de continuité sur un intervalle ouvert (a, b) ou encore sur tout $\mathbb{R}$.

Une façon intuitive de comprendre la continuité des fonctions sur $\mathbb{R}$ est obtenue à partir de la notion de graphe de $f\colon \mathbb{R} \to \mathbb{R}$. La continuité s'exprime alors (assez vaguement on en conviendra) en disant que l'on peut tracer le graphe de f sans lever le crayon. Il est important de réfléchir sur le fait que la définition proposée de la continuité n'est rien d'autre qu'une formalisation de cette intuition où l'on ne se fie plus à des notions imagées telles la *droite réelle* mais bien à des concepts analytiques rigoureux.

Si la définition de continuité a une apparence simple, on se convainc rapidement du contraire en la réecrivant de manière complètement explicite avec ce que l'on a fait dans le contexte de la section sur les suites réelles :

$$\{\forall \delta > 0 \; \exists N \in \mathbb{N} \; t.q. \; n > \mathbb{N} \Rightarrow |x_n - a| < \delta\}$$
$$\Rightarrow \{\forall \epsilon > 0 \; \exists N' \in \mathbb{N} \; t.q. \; n > N' \Rightarrow |f(x_n) - f(a)| < \epsilon\}.$$

Cette expression nous mène à une autre approche pour exprimer la continuité des fonctions réelles :

Théorème 1.2. *Une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est continue en $a \in \mathbb{R}$ si et seulement si*

$$\forall \epsilon > 0 \; \exists \delta > 0 \; t.q. \; |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Preuve. Dans une direction, soit $\epsilon > 0$ quelconque choisi et prenons un $\delta > 0$ tel que $|x - a| < \delta$ implique $|f(x) - f(a)| < \epsilon$. Pou une suite quelconque $\{x_n\}$ convergeant vers a, on a qu'il existe $N \in \mathbb{N}$ tel que $n > N$ implique $|x_n - a| < \delta$ (notez que N dépend de $\delta > 0$ ici). Mais alors on a que $|f(x_n) - f(a)| < \epsilon$ dès que $n > N$, ce qui donne bien la continuité de f en a.
Inversement, supposons que la fonction f est continue et choisissons $\epsilon > 0$ quelconque. Par la continuité de f on peut choisir $\delta > 0$ tel que pour $n > \max\{N, N'\}$ on ait

$$|x_n - a| < \delta \Rightarrow |f(x_n) - f(a)| < \epsilon.$$

Alors pour tout x tel que $|x-a| < \delta$, on a en particulier pour la suite constante $\{x_n\} = \{x\}$ que $|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon$ tel que désiré. □

Remarquons que dans la majorité des livres d'introduction à l'Analyse, la caractérisation "$\epsilon - \delta$" de la continuité est utilisée comme définition. L'étudiant ou l'étudiante aura grand intérêt à s'assurer qu'il ou elle comprend bien les deux points de vue puisque chacun a ses avantages. Par exemple, lorsque l'on doit se prononcer de façon pratique sur la continuité d'une fonction f donnée, l'approche à travers les suites sera favorable si l'on essaie de montrer que la fonction n'est pas continue (car alors il suffit d'exhiber une *seule* suite $\{x_n\}$ pour laquelle la condition n'est pas satisfaite : $\{x_n\} \to a$ mais $\{f(x_n)\} \nrightarrow f(a)$). L'utilisation des "$\epsilon - \delta$" permettra plutôt d'établir la continuité, jusqu'à ce que l'on ait développé d'autres outils plus pratiques.

EXEMPLE 1.3. *Considérons la fonction donnée par*

$$f(x) = \begin{cases} x^{1/k} & si\ x \geq 0 \\ 0 & sinon. \end{cases}$$

Suite Ex

Donnons maintenant une autre caractérisation de la continuité, plus globale et qui pourrait mener à l'étude de la continuité dans le cadre beaucoup plus général des espaces topologiques (voir Analyse II ou Topologie Générale). Cette approche repose sur la notion d'ensembles ouverts et de leur comportement par rapport à une fonction.

DÉFINITION 1.4. *Un sous-ensemble $\mathcal{U} \subseteq \mathbb{R}$ est dit ouvert si pour tout $x \in \mathcal{U}$ on peut trouver un intervalle ouvert I_x inclus dans $\mathcal{U}$.*

On considère par défaut que l'ensemble vide $\varnothing \subset \mathbb{R}$ est ouvert. Cette définition peut être comprise intuitivement de la façon suivante : un ouvert $\mathcal{U}$ de $\mathbb{R}$ est un ensemble qui a la forme d'un intervalle ouvert autour de chacun de ses éléments.

La notion complémentaire à celle d'ensemble ouvert dans $\mathbb{R}$ est celle d'ensemble *fermé* : $\mathcal{F} \subseteq \mathbb{R}$ est dit fermé si son complémentaire $\mathcal{F}^c = \mathbb{R} - \mathcal{F}$ est un ensemble ouvert.

THÉORÈME 1.5. *Une fonction $f \colon \mathbb{R} \to \mathbb{R}$ est continue sur $\mathbb{R}$ si et seulement si pour tout ouvert $\mathcal{U} \subseteq \mathbb{R}$, l'ensemble pré-image $f^{-1}(\mathcal{U}) = \{x \in \mathbb{R} \mid f(x) \in U\}$ est ouvert dans $\mathbb{R}$.*

Preuve. Soit $f \colon \mathbb{R} \to \mathbb{R}$ continue et $\mathcal{U} \subseteq \mathbb{R}$ un ouvert quelconque. Prenons $x \in f^{-1}(\mathcal{U})$. Alors $f(x) \in \mathcal{U}$ et comme $\mathcal{U}$ est ouvert, il existe un $\epsilon > 0$ tel que l'intervalle $J = (f(x)-\epsilon, f(x)+\epsilon)$ soit inclus dans $\mathcal{U}$. Par continuité de f on peut alors trouver $\delta > 0$ tel que

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Donc on a $f(y) \in J \subseteq \mathcal{U}$ dès que $y \in (x - \delta, x + \delta)$. C'est-à-dire que l'on a $(x - \delta, x + \delta) \subseteq f^{-1}(\mathcal{U})$ et donc l'ensemble $f^{-1}(\mathcal{U})$ est bien ouvert.
De façon réciproque, soit f non-continue en $a \in \mathbb{R}$. Alors il existe un $\epsilon_0 > 0$ tel que pour

tout $\delta > 0$ on puisse trouver $x_0 \in \mathbb{R}$ satisfaisant $|x_0 - a| < \delta$ et $|f(x_0 - f(a)| \geq \epsilon_0$. Donc on a $f(x_0) \notin J_0 = (f(a) - \epsilon_0, f(a) + \epsilon_0)$. Mais alors nous avons construit J_0 ouvert tel que $f^{-1}(J_0)$ ne soit pas ouvert, puisque tout intervalle autour de $a \in f^{-1}(J_0)$ contient, par construction, des points n'étant pas dans $f^{-1}(J_0)$. □

C'est donc dire que l'on caractérise les fonctions continues sur $\mathbb{R}$ comme celles pour lesquelles la pré-image de tout ouvert de $\mathbb{R}$ est également un ouvert de $\mathbb{R}$. Il est à noter que l'on parle ici de continuité sur tout $\mathbb{R}$. On pourrait facilement adapter le résultat à un sous-ensemble ouvert de $\mathbb{R}$, mais dans tous les cas, on ne parle pas ici de continuité en un point seulement.

Exercice 1.6. *Trouvez les points de continuité et de discontinuité des fonctions suivantes :*

(1) *Soit $\{r_n\}$ l'ensemble des rationnels et posons*

$$f(x) = \begin{cases} 0 & si\ x \in \mathbb{Q} \\ 1 & sinon. \end{cases}$$

(2) *Soit $\mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z}, p > 0, (p, q) = 1\}$ et posons*

$$f(x) = \begin{cases} continuer! \end{cases}.$$

2. Ensembles de points de discontinuité

Soit $f \colon \mathbb{R} \to \mathbb{R}$ et I un intervalle ouvert. Si l'ensemble $\{f(x) \mid x \in I\} \subset \mathbb{R}$ est borné, on pose

$$M(f, I) = \sup\{f(x) \mid x \in I\}\ ,\ m(f, I) = \inf\{f(x) \mid x \in I\}.$$

Définition 2.1. *On appelle amplitude de f sur I le nombre réel*

$$w(f, I) = M(f, I) - m(f, I).$$

Il est important de noter que cette notion dépend de la fonction f mais également de l'intervalle I choisi. Egalement, on montrera aisément que si l'on a $I_1 \subseteq I_2$, alors $w(f, I_1) \leq w(f, I_2)$.

Définition 2.2. *Le saut de f en un point $a \in \mathbb{R}$ est le nombre*

$$w(f, a) = \inf_{I_a} w(f, I_a),$$

où l'infimum est pris sur tous les intervalles I_a contenant le point a.

Quelle est l'idée derrière cette construction ? Pour chaque intervalle I_a, on obtient une amplitude de f sur I_a et en prenant des intervalles de plus en plus petits autour du point a, si la fonction f est continue, on s'attend à ce que cette amplitude diminue vers 0. A l'inverse, si la fonction n'est pas continue en a, on s'attend à ce que ce saut soit non-nul. Ceci est formalisé dans la proposition suivante :

PROPOSITION 2.3. *Une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est continue en $a \in \mathbb{R}$ si et seulement si $w(f,a) = 0$.*

Preuve. Supposons que $w(f,a) = 0$ et choisissons un $\epsilon > 0$ arbitraire. Il y a alors un intervalle J_a contenant a tel que $w(f,I_a) < \epsilon$ par la définition de $w(f,a)$ comme infimum des $w(f,I_a)$ sur tous les I_a contenant a. Pour tout élément $x \in J_a$, on a alors

$$|f(x) - f(a)| \leq w(f,J_a) < \epsilon.$$

La condition "$\epsilon - \delta$" pour la continuité implique alors que f est continue.

Inversement, supposons f continue en a et fixons $\epsilon > 0$. On a alors un $\delta > 0$ tel que $|x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon/2$. Posons $J = (a-\delta, a+\delta)$. Alors si $x, y \in J$,

$$|f(x)| - |f(y)| \leq |f(x) - f(y)| \leq |f(x) - f(a)| + |f(y) - f(a)|.$$

Il s'en suit que

$$M(f,J) - m(f,J) \leq |f(x) - f(a)| + |f(y) - f(a)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Comme $\epsilon > 0$ est arbitraire, on obtient bien $w(f,a) = 0$. □

On peut se demander s'il est possible de construire des fonctions réelles dont l'ensemble des points de continuité serait déterminé comme bon nous semble. On a déjà rencontré des exemples de fonctions continues sur les nombres irrationnels et discontinues sur l'ensemble des rationnels. Ce type de comportement était difficile à imaginer à partir d'une approche intuitive de la continuité, où notre esprit a naturellement tendance à séparer les points de discontinuité. On peut en fait rendre l'ensemble des points de continuité d'une fonction très restraint, comme le montre l'exercice suivant :

Exercice 2.4. *Démontrez que la fonction suivante définie sur tout $\mathbb{R}$ n'est continue qu'en un seul point (dont vous trouverez au passage l'identité !)*

$$f(x) = ...(continuer).$$

On démontre maintenant un résultat de nature plutôt théorique sur l'ensemble des points de discontinuité d'une fonction réelle, qui impose certaines restrictions aux types d'ensembles que l'on peut espérer voir comme ensemble de discontinuité d'une fonction.

Théorème 2.5. *Soit $f\colon \mathbb{R} \to \mathbb{R}$ et $\mathcal{D}$ l'ensemble des points de discontinuité de f. Alors $\mathcal{D}$ est réunion dénombrable d'ensembles fermés.*

Preuve. Cette preuve fait appel à la notion du saut $w(f, x)$ d'une fonction f en un point $x \in \mathbb{R}$. On se souvient que le saut est nul précisément lorsque la fonction est continue en ce point. Soit $\alpha > 0$ un nombre. Si l'on suppose que $w(f, x) < \alpha$, alors on a par construction que $w(f, J_x) < \alpha$ pour un certain intervalle J_x contenant x. Alors on a aussi $w(f, y) < \alpha$ pour tout $y \in J_x$, puisque J_x peut également servir d'intervalle ouvert autour de y. Il s'en suit que l'ensemble $\{y \in \mathbb{R} \mid w(f, y) < \alpha\}$ est ouvert. Donc son complément

$$D_\alpha = \{y \in \mathbb{R} \mid w(f, y) \geq \alpha\}$$

est un ensemble fermé. Mais alors on peut maintenant exprimer l'ensemble $\mathcal{D}$ des points de discontinuité de f comme

$$\mathcal{D} = \bigcup_{n=1}^{\infty} D_{1/n}$$

ce qui termine la preuve. □

Corollaire 2.6. *Il n'existe pas de fonction $f\colon \mathbb{R} \to \mathbb{R}$ qui ne soit continue que sur $\mathbb{Q}$ (ou sur tout sous-ensemble dénombrable et dense dans $\mathbb{R}$).*

Preuve. En effet, l'ensemble $\mathbb{R} - \mathbb{Q}$ n'est pas réunion dénombrable d'ensembles fermés, la preuve établissant que $\mathbb{R}$ n'est pas dénombrable en utilisant le principe des segments emboîtés pouvant être ici adaptée). □

3. Opérations sur les fonctions continues

Au cours de cette section, nous établissons que la continuité est préservée sous les opérations algébriques usuelles sur les fonctions. Les résultats ne sont pas très difficiles en soi et ils permettent de démontrer rapidement que de nombreuses fonctions sont continues en s'y référant. On remarquera que les preuves suivent toutes essentiellement la même logique et exploitent la définition séquentielle de la continuité.

PROPOSITION 3.1. *Si $f, g : \mathbb{R} \to \mathbb{R}$ sont deux fonctions continues et $\alpha \in \mathbb{R}$, alors $\alpha f + g : \mathbb{R} \to \mathbb{R}$ est continue.*

Preuve. Soit $\{x_n\} \to x$ quelconque. Par l'hypothèse de continuité en x on a $\{f(x_n)\} \to f(x)$ et $\{g(x_n)\} \to g(x)$. Alors la suite $\{(\alpha f + g)(x_n)\} = \{\alpha f(x_n) + g(x_n)\}$ converge vers $\alpha f(x) + g(x) = (\alpha f + g)(x)$ par la linéarité de la limite de suites, et donc la fonction $\alpha f + g$ est continue en x. $\square$

Il s'en suit que l'ensemble des fonctions continues sur $\mathbb{R}$ est un espace vectoriel sur $\mathbb{R}$. On peut aisément montrer (voir un cours d'Algèbre Linéaire) que cet espace vectoriel est de dimension infinie.

PROPOSITION 3.2. *Soient $f, g : \mathbb{R} \to \mathbb{R}$ continues. Alors $fg : \mathbb{R} \to \mathbb{R}$ est continue. Si en outre $g(x) \neq 0$, alors $f/g : \mathbb{R} \to \mathbb{R}$ est continue.*

Preuve. Pour $\{x_n\} \to x$, on a par continuité de f et g que $\{f(x_n)\} \to f(x)$ et $\{g(x_n)\} \to g(x)$ et donc $\{(fg)(x_n)\}$ converge vers $(fg)(x)$ par la formule de limite d'un produit de suites convergentes. De la même manière on établit le cas du quotient de deux fonctions dont le dénominateur $g(x)$ n'est pas nul. $\square$

Les deux dernières propositions donnent donc que l'ensemble des fonctions continues sur $\mathbb{R}$ est une algèbre sur $\mathbb{R}$ avec les opérations d'addition et de produit de fonctions.

PROPOSITION 3.3. *Soient $f, g : \mathbb{R} \to \mathbb{R}$ continues. Alors $f \circ g : \mathbb{R} \to \mathbb{R}$ est continue.*

Preuve. Soit $\{x_n\} \to x$ quelconque. On veut établir que $\{(f \circ g)(x_n)\} \to \{(f \circ g)(x)\}$. Par continuité de g en x on a $\{g(x_n)\} \to g(x)$, alors que par continuité de f en $g(x)$ on a $\{f(g(x_n))\} \to f(g(x))$, ce qui donne bien la continuité de $f \circ g$. $\square$

EXERCICE 3.4. *Etablissez qu'un polynôme $p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n$ définit une fonction réelle continue sur $\mathbb{R}$.*

EXERCICE 3.5. *Démontrez que la fonction $f(x) = \sqrt{x^8 + 3x^2}$ est continue sur $\mathbb{R}$.*

4. Propriétés fondamentales de la continuité des fonctions

Nous étudions dans cette sections des propriétés importantes des fonctions continues, allant souvent bien au-delà des conséquences immédiates des définitions. Il s'agit donc d'une section centrale du chapitre.

PROPOSITION 4.1. *Soit $f\colon \mathbb{R} \to \mathbb{R}$ une fonction continue pour laquelle $f \equiv 0$ sur un ensemble $E \subseteq \mathbb{R}$ dense dans $\mathbb{R}$. Alors f est identiquement nulle sur tout $\mathbb{R}$.*

Preuve. Soit $x \in \mathbb{R}$. Comme E est dense dans $\mathbb{R}$, on peut trouver une suite $\{x_n\} \subset E$ convergeant vers x. Alors par l'hypothèse sur f on a $f(x_n) = 0$ $(\forall n \in \mathbb{N})$. Par continuité de f au point x, il en découle que

$$f(x) = \lim\{f(x_n)\} = 0$$

tel que voulu. □

COROLLAIRE 4.2. *Si $f, g : \mathbb{R} \to \mathbb{R}$ sont des fonctions continues telles que*

$$f(p/q) = g(p/q) \ (\forall p/q \in \mathbb{Q}),$$

alors on a $f \equiv g$ sur $\mathbb{R}$.

Preuve. Considérer la fonction $h = f - g$ et appliquer la proposition précédente. □

PROPOSITION 4.3. *Soit $f\colon \mathbb{R} \to \mathbb{R}$ continue et partout strictement positive : $f(x) > 0$ $(\forall x \in \mathbb{R})$. Prenons I un intervalle fermé et borné quelconque. Alors il existe $\alpha > 0$ tel que pour tout $x \in I$ on ait $f(x) > \alpha$.*

Preuve. Commençons par rappeler que si une suite strictement positive $\{a_n\}$ converge vers une valeur $a > 0$, alors il existe $\alpha > 0$ tel que $a_n > \alpha$ $(\forall n \in \mathbb{N})$. Nous utilisons cette propriété des suites et procédons par contraposition pour la preuve de la proposition. Supposons donc, au contraire de la conclusion, que pour chaque $n \in \mathbb{N}$ on ait $x_n \in I$ tel que $f(x_n) \leq 1/n$. Alors $\{x_n\}$ est une suite dans I borné et possède donc une sous-suite convergente en vertu du Théorème de Bolzano-Weierstrass : $\{x_{n_k}\} \to x$. Comme I est également fermé, on a que $x \in I$. La continuité de f donne alors que $\{x_{n_k}\} \to x \Rightarrow \{f(x_{n_j})\} \to f(x)$. Mais alors ceci et la condition $0 < f(x_{n_k}) < 1/n_k$ impliquent que $f(x) = 0$ ce qui contredit le fait que f est supposée strictement positive. □

EXERCICE 4.4. *Donnez deux preuves différentes de cette dernière propoposition : (1) en utilisant un argument de type $\epsilon - \delta$ (2) en utilisant le théorème de recouvrement de Borel.*

EXERCICE 4.5. *Montrez, exemples à l'appui, que la conclusion de la proposition est fausse si l'on omet une des deux hypothèses sur l'ensemble I.*

THÉORÈME 18. *Soit $f\colon \mathbb{R} \to \mathbb{R}$ continue et I un intervalle fermé et borné. Alors la fonction f est bornée sur I : $\exists M \in \mathbb{R}$ t.q. $|f(x)| \leq M$ $(\forall x \in I)$.*

Preuve. Supposons que f n'est pas bornée sur I. Alors, en particulier, pour chaque $n \in \mathbb{N}$ il existe $x_n \in I$ tel que $|f(x_n)| > n$. Ceci définit une suite dans I, intervalle fermé et borné, donc également une sous-suite convergente dans I, $\{x_{n_k}\} \to x$. Mais alors, par construction, on a $\{f(x_{n_k})\} \not\to f(x)$ puisque $|f(x_{n_k})| > n_k$, donc f ne serait pas continue en x, une contradiction. Il s'en suit que f doit être bornée sur I. □

Ce résultat est fondamental à plusieurs égards en Analyse. En particulier, il implique que toute fonction continue sur un intervalle fermé et borné y possède un supremum et un infimum bien définis :

$$\sup_{x\in I}\{f(x)\}\ ,\ \inf_{x\in I}\{f(x)\}.$$

Lorsque la fonction f atteint son supremum ou son infimum sur I, on parle alors de *maximum* ou de *minimum* de f sur I.

THÉORÈME 4.6. *Toute fonction continue $f\colon \mathbb{R} \to \mathbb{R}$ atteint son supremum et son infimum sur tout intervalle fermé et borné I.*

Preuve. Par le Théorème 18, on a f bornée sur I, soit donc

$$M = \sup_{x\in I}\{f(x)\}.$$

Par la définition du supremum, pour chaque $n \in \mathbb{N}$, on a un $x_n \in I$ tel que

$$M \geq f(x_n) \geq M - 1/n.$$

Ceci définit donc une suite $\{x_n\}$ dans I telle que $\{f(x_n)\} \to M$. Comme I est fermé et borné, on a également une sous-suite $\{x_{n_k}\} \to x_0 \in I$. Mais alors la continuité de f donne que $\{f(x_{n_k})\} \to f(x_0)$ et ainsi on a bien $M = f(x_0)$ avec $x_0 \in I$. La preuve pour l'infimum $m = \sup\limits_{x\in I}\{f(x)\}$ est identique. □

S'il importe de souligner l'importance de ce résultat d'un point de vue théorique, on doit en contre-partie reconnaître qu'il ne nous donne aucun moyen concret de *trouver explicitement* le maximum ou le minimum de f sur I. Egalement, le lecteur ou la lectrice devra se convaincre, à travers des exemples bien choisis, que la condition de continuité de f et les propriétés de I – fermé et borné – sont essentielles pour la véracité du résultat.

THÉORÈME 19. **(Théorème de la valeur intermédiaire)**
Si $f\colon \mathbb{R} \to \mathbb{R}$ est continue, $a < b$ et que c est un nombre quelconque entre $f(a)$ et $f(b)$, alors il existe $x_0 \in (a, b)$ tel que $f(x_0) = c$.

Preuve. Un cas particulier, qui est en fait équivalent à l'énoncé général, est également connu sous le nom de TVI : si f est continue sur $\mathbb{R}$, $a < b$, $f(a) < 0$ et $f(b) > 0$, alors il existe un $x_0 \in (a, b)$ tel que $f(x_0) = 0$. Si l'on pose $h(x) = f(x) - c$, alors h est continue et la condition $f(a) < c < f(b)$ devient $h(a) < 0$, $h(b) > 0$ et l'on est ramené au cas particulier. Considérons alors l'ensemble

$$E = \{x \in [a, b] \mid h(x) < 0\}$$

et posons $\alpha = \sup E$. On a alors $a < \alpha < b$ car $h(a) < 0$ et $h(b) > 0$. Par définition du supremum d'un ensemble, on alors des suites $\{x_n\} \subset E$ et $\{y_n\} \subset E^c$ convergeant vers α. Par construction on a $h(x_n) < 0$ et $h(y_n) > 0$ pour tout $n \in \mathbb{N}$. La continuité de la fonction h donne donc

$$h(\alpha) = h(\lim_{n\to\infty} \{x_n\}) = \lim_{n\to\infty} \{h(x_n)\} \leq 0,$$
$$h(\alpha) = h(\lim_{n\to\infty} \{y_n\}) = \lim_{n\to\infty} \{h(y_n)\} \geq 0.$$

Ces deux équations imposent donc que $h(\alpha) = 0$ et on a donc trouvé $x_0 = \alpha \in (a, b)$ tel que $h(x_0) = 0$. □

Exercice 4.7. *Soit n un nombre impair quelconque. Montrez que tout polynôme* $p(x) = x^n + a_{n-1}x^{n-1} + ... + a_1x + a_0$ *possède au moins une racine réelle*

Théorème 4.8. **(Théorème du point fixe)**
Soit $f\colon [0, 1] \to [0, 1]$ *continue. Alors* f *possède un point fixe : il existe* $x_0 \in [0, 1]$ *tel que* $f(x_0) = 0$.

Preuve. Un point fixe correspond à l'intersection entre les graphes des fonctions continues f et g où $g(x) = x$. Il s'agit alors de définir une fonction à laquelle on pourra appliquer le TVI pour obtenir le résultat. Les détails sont laissés en exercice. □

Exemple : Fonctions continues sur le cercle et fonctions réelles périodiques
On considère le cercle $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ que l'on paramétrise en utilisant la fonction $\alpha\colon [0, 2\pi] \to S^1$ donnée par $\alpha(t) = (\cos t, \sin t)$. On dira alors qu'une fonction $f\colon S^1 \to \mathbb{R}$ est continue sur S^1 si la fonction $f \circ \alpha\colon [0, 2\pi] \to \mathbb{R}$ est continue comme fonction réelle.

Il est relativement facile de voir que pour toute fonction continue $f\colon S^1 \to \mathbb{R}$, il existe au moins deux points $p_1, p_2 \in S^1$ tels que $f(p_1) = f(p_2)$. Le lecteur ou la lectrice en fournira la preuve. Ainsi, par exemple, sur une tige metallique en forme de cercle où la température serait continment distribuée, il y a toujours deux points ayant la même température... Mais on peut faire bien mieux !

Si l'on appelle p et $-p$ sur S^1 une paire de points *antipodaux*, alors le théorème de la valeur intermédiaire implique le résultat suivant qui peut paraître surprenant à prime abord :

PROPOSITION 4.9. *Toute fonction* $f\colon S^1 \to \mathbb{R}$ *continue possède au moins une paire de points antipodaux* $\{p, -p\}$ *pour lesquels* $f(p) = f(-p)$.

Preuve. On peut supposer que l'on a $f((1,0)) \neq f((-1,0))$ (sinon la preuve est terminée). Sans perte de généralité supposons que $f((1,0)) < f((-1,0))$ et donc $f(\alpha(0)) < f(\alpha(\pi))$. Posons

$$h(t) = f(\alpha(t)) - f(\alpha(\pi + t)),$$

une fonction définie sur $[0, \pi]$, continue et satisfaisant $h(0) < 0$ et $h(\pi) > 0$ par construction. Par le théorème de la valeur intermédiaire, il existe $x_0 \in [0, \pi]$ tel que $h(x_0) = 0$. Alors $p_0 = \alpha(x_0)$ et $\alpha(x_0 + \pi) = -p_0$ sont des points antipodaux pour lesquels on a bien $f(p_0) = f(-p_0)$ tel que désiré. $\square$

On peut alors faire le lien avec les fonctions réelles périodiques. On dit qu'une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est *périodique* de période T si

$$f(x + T) = f(x) \ (\forall x \in \mathbb{R}).$$

Une telle fonction définit naturellement une fonction continue $\tilde{f}\colon S^1 \to \mathbb{R}$ telle que

$$\tilde{f}((\cos(\frac{2\pi}{T}x), \sin(\frac{2\pi}{T}x))) = f(x).$$

Le lecteur ou la lectrice pourra démontrer que $f\colon \mathbb{R} \to \mathbb{R}$ continue et périodique est forcément bornée sur $\mathbb{R}$ et, également, qu'il existe au moins un $x \in \mathbb{R}$ tel que $f(x) = f(x + T/2)$.

5. Théorème de la fonction inverse et conséquences

Etant donné une fonction $f\colon \mathbb{R} \to \mathbb{R}$ qui est supposée injective, on peut considérer sa fonction inverse, $f^{-1}\colon \operatorname{Im} f \to \mathbb{R}$ construite de la façon suivante. Si $x \in \mathbb{R}$ est l'unique élément tel que $y = f(x)$ (unique par hypthèse d'injectivité de f), on pose $f^{-1}(y) = x$. Si l'on suppose que f est continue, on peut se demander si la fonction f^{-1} ainsi obtenue l'est également. Ceci est donné par le résultat suivant :

Théorème 20. **(Théorème de la fonction inverse)**
Soit $f\colon \mathbb{R} \to \mathbb{R}$ *strictement croissante (ou alors strictement décroissante) et continue. Alors la fonction inverse* $f^{-1}\colon \operatorname{Im} f \to \mathbb{R}$ *est elle aussi continue.*

Preuve. Par hypothèse la fonction f est strictement croissante (ou alors strictement décroissante), donc elle est injective et ainsi $f^{-1}\colon \operatorname{Im} f \to \mathbb{R}$ est bien définie. Notons également que cette même hypothèse implique que $f(\mathbb{R})$ est un intervalle ouvert I (possiblement tout $\mathbb{R}$), en utilisant le théorème de la valeur intermédiaire.
Après ces remarques préliminaires, montrons la continuité de $f^{-1}\colon I \to \mathbb{R}$. Soit $y \in I$ et x l'unique réel tel que $f(x) = y$. Fixons $\epsilon > 0$ quelconque. Comme f est continue en x, on a $\delta > 0$ tel que la fonction croissante f satisfasse

$$f(x) - f(x-\delta) < \epsilon \qquad \text{et} \qquad f(x+\delta) - f(x) < \epsilon.$$

On peut sans problème supposer que $\delta < \epsilon$. Posons $y_1 = f(x-\delta)$ et $y_2 = f(x+\delta)$, si bien que l'on a $y_1 < y < y_2$ par croissance de f. Posons $\delta' = \min\{y - y_1, y - y_2\}$. On a alors construit $\delta' > 0$ tel que

$$|y - y'| < \delta' \Rightarrow |f^{-1}(y) - f^{-1}(y')| < \delta' < \epsilon.$$

C'est donc dire que f^{-1} est continue en $y \in I$ quelconque tel que désiré. □

Exemple 5.1. *La fonction* $f(x) = \sqrt{x}$ *est continue sur tout* $\mathbb{R}^+$ *en vertu du théorème de la fonction inverse puisque sa fonction inverse est donnée par la fonction* $g(x) = x^2$ *qui est continue sur tout* $\mathbb{R}$.

Exercice 5.2. *Soit* n *un nombre pair quelconque. Montrez que tout polynôme* $p(x) = x^n + a_{n-1}x^{n-1} + ... + a_1x + a_0$ *ne possède pas de fonction inverse sur* $\mathbb{R}$ *entier mais que la fonction possède forcément un minimum sur* $\mathbb{R}$.

6. Continuité uniforme

Nous avons mis l'emphase sur le fait que la notion de continuité est un concept local : f est continue *en* $a \in \mathbb{R}$. Quand la continuité est vraie pour chaque point, on peut alors parler de continuité dans un sens plus global. La notion importante que nous définissons dans cette section pousse encore plus loin ce caractère global dans le cas de fonctions satisfaisant une condition plus forte que la simple continuité.

Commençons par donner une autre interprétation de la continuité, peut-être plus géométrique que la définition faite à partir des suites ou de l'approche $\epsilon - \delta$. Pour cela, considérons dans $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ l'ensemble suivant autour du point $(a, f(a))$:

$$R_{(\delta,\epsilon)} = \{(x, y) \in \mathbb{R}^2 \mid |x - a| < \delta, |y - f(a)| < \epsilon\}.$$

Géométriquement, $R_{(\delta,\epsilon)}$ est un rectangle centré au point $(a, f(a))$ du graphe de f de longueur 2δ et hauteur ϵ. La définition de la continuité de f par l'approche $\epsilon - \delta$ équivaut alors à pouvoir trouver, pour chaque $\epsilon > 0$, un réel $\delta > 0$ tel que le graphe de f soit inclus dans $R_{(\delta,\epsilon)}$ lorsque x est pris dans l'intervalle $(a - \delta, a + \delta)$. A l'inverse, le fait que la fonction f soit discontinue au point a se traduit par l'existence d'un ϵ_0 tel que pour tout $\delta > 0$ il y ait toujours un $x \in (a - \delta, a + \delta)$ pour lequel le rectangle $R_{(\delta,\epsilon_0)}$ ne contienne pas $(x, f(x))$.

De ce point de vue, étant donné $\epsilon > 0$, le rectangle $R_{(\delta,\epsilon)}$ à choisir pour exprimer la continuité de f en un point a va généralement dépendre du point a auquel on se place. En termes plus analytiques, le nombre δ dans la phrase exprimant la continuité de f en un point a :

$$\forall \epsilon > 0 \; \exists \delta > 0 \text{ tel que } |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon,$$

dépendra généralement du point a. Toute la discussion ci-dessus, assez longue mais ayant pour but d'exprimer intuitivement ces idées, mène donc à la définition suivante :

DÉFINITION 6.1. *Une fonction $f : \mathbb{R} \to \mathbb{R}$ est dite uniformément continue sur $\mathbb{R}$ si*

$$\forall \epsilon > 0 \; \exists \delta > 0 \text{ tel que } |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \epsilon.$$

Alors que la continuité en un point a ne faisait intervenir que le comportement local de la fonction, on voit ci-dessus que la continuité uniforme, elle, fait appel au comportement global de la fonction f sur $\mathbb{R}$. Géométriquement, ceci s'exprime par le fait que pour $\epsilon > 0$ quelconque donné, on peut trouver un $\delta > 0$ tel que, autour de tout $(x', f(x'))$, le graphe de f soit contenu dans $R_{(\delta,\epsilon)}$ pour autant que x soit contenu dans $(x' - \delta, x' + \delta)$.

Comme le montrent les exemples ci-dessous, cette condition est en fait plutôt contraignante, si bien qu'il existe beaucoup de fonctions continues sur tout $\mathbb{R}$ qui ne sont pas uniformément continues sur tout $\mathbb{R}$.

Exemple 6.2. *La fonction $f(x) = x^2$ est continue sur tout $\mathbb{R}$ mais elle n'y est pas uniformément continue. Nous donnons deux approches pour montrer ceci. D'une part, fixons $\epsilon_0 = 1$ (par exemple). Pour $\delta > 0$ quelconque supposons que $|x - y| < \delta$. On a alors*

$$|f(x) - f(y)| = |x^2 - y^2| = |x - y||x + y|.$$

Mais alors par la propriété archimédienne de $\mathbb{R}$, il est possible de trouver x_0 et y_0 tels que $|x_0 - y_0| < \delta$ mais $|x_0 - y_0||x_0 + y_0| > 1 = \epsilon_0$. Il s'en suit que f ne peut être uniformément continue.
Dans la seconde approche, soit toujours $\epsilon_0 = 1$ et pour tout $\delta > 0$ on construit $x_0 = \frac{1}{\delta} + \delta$ et $y_0 = \frac{1}{\delta}$. On a alors la condition $|x_0 - y_0| < \delta$ mais

$$|f(x_0) - f(y_0)| = |(\frac{1}{\delta} + \delta)^2 - \frac{1}{\delta^2}| = |2 + \delta^2| > \epsilon_0,$$

et donc f n'est pas uniformément continue sur $\mathbb{R}$.

On pourrait croire, naïvement, que le fait que la fonction f soit uniformément continue ou non est lié au fait que la fonction soit bornée ou non. En fait, il n'en est rien comme le montrent les fonctions dans l'exercice suivant.

Exercice 6.3. *Démontrez que les fonction non-bornées $f(x) = x$ et $g(x) = \sqrt{x}$ sont uniformément continues sur $\mathbb{R}$, alors que la fonction $h(x) = \sin(\frac{1}{x})$ est bornée et continue sur $(0, +\infty)$ mais elle n'y est pas uniformément continue.*

On a établi que les notions de continuité et de continuité uniforme étaient bel et bien distinctes sur $\mathbb{R}$, mais on peut légitimement se poser la même question sur des ensembles aux caractéristiques particulières. Les exemples ci-dessus montrent que les notions sont distinctes sur un exemble fermé quelconque ($\mathbb{R}$ est fermé !) ainsi que sur un ensemble borné ($(0, 1)$ est un ensemble borné). C'est une propriété remarquable des ensembles compacts (fermés et bornés) de $\mathbb{R}$ que continuité et continuité uniforme y soient équivalentes.

Théorème 21. **(Théorème de la continuité uniforme)**
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue et E un ensemble fermé et borné dans $\mathbb{R}$. Alors f est uniformément continue sur E : $\forall \epsilon > 0\ \exists \delta$ tel que pour $x, y \in E$, $|x - y| < 0 \Rightarrow |f(x) - f(y)| < \epsilon$.

Preuve. Supposons que f n'est pas uniformément continue sur E : en particulier ,il existe alors $\epsilon > 0$ tel que pour tout $n \in \mathbb{N}$ on ait $x_n, y_n \in E$ satisfaisant

$$|x_n - y_n| < \frac{1}{n} \text{ mais } |f(x_n) - f(y_n)| \geq \epsilon.$$

Etant donné que E est fermé et borné, on sait qu'il existe une sous-suite convergente $\{x_{n_k}\}$ (disons vers $x \in E$). Comme on a $|x_{n_k} - y_{n_k}| < \frac{1}{n_k}$, on a également $\{y_{n_k}\} \to x$.
Or puisque $|f(x_{n_k}) - f(y_{n_k})| \geq \epsilon$, on ne peut en particulier avoir

$$\lim_{n_k \to \infty} \{f(x_{n_k})\} = \lim_{n_k \to \infty} \{f(y_{n_k})\},$$

ce qui contredit la continuité de f en $x \in E$. Contradiction. $\square$

EXERCICE 6.4. *En utilisant l'interprétation géométrique de la continuité uniforme et la notion de saut sur un intervalle, donnez une autre preuve qui s'appuiera également sur le fait qu'une fonction continue sur un intervalle fermé et borné est toujours bornée.*
Egalement, en utilisant le théorème de Borel pour un intevalle fermé et borné, essayez de fournir une troisième preuve de ce résultat fondamental. Une fois que vous aurez fait adéquatement cet exercice, vous aurez une bonne idée de ce que représente le concept de continuité uniforme.

Nous donnons une application de ces idées en démontrant un théorème d'approximation des fonctions continues sur un intervalle par des fonctions particulièrement simples, les fonctions escalier et les fonctions linéaires par morceaux. Ce résultat sera d'ailleurs utilisé au chapitre portant sur l'intégration pour obtenir rapidement une preuve du théorème fondamental du calcul différentiel et intégral.

DÉFINITION 6.5. *On appelle $f\colon [a,b] \to \mathbb{R}$ fonction escalier s'il existe une partition de l'intervalle $[a,b]$, $a = x_0 < x_1 < ... < x_n = b$ telle que*

$$f(x) = \begin{cases} c_1 & \textit{sur } [x_0, x_1) \\ c_2 & \textit{sur } [x_1, x_2) \\ ... & \\ c_n & \textit{sur } [x_{n-1}, x_n) \end{cases}$$

où $c_1, c_2, ..., c_n$ sont des constantes.

DÉFINITION 6.6. *Une fonction $f\colon [a,b] \to \mathbb{R}$ est dite linéaire par morceaux s'il existe une partition de l'intervalle $[a,b]$, $a = x_0 < x_1 < ... < x_n = b$ telle que sur chaque $[x_{i-1}, x_i]$ f soit linéaire $(i = 1, 2, ..., n)$.*

THÉORÈME 6.7. **(Théorème d'approximation)**
Soit $f\colon \mathbb{R} \to \mathbb{R}$ continue et I un intervalle fermé et borné. Alors f peut être approximée sur I par des fonctions escalier (resp. linéaires par morceaux) : $\forall \epsilon > 0$ $\exists g\colon I \to \mathbb{R}$ fonction

escalier (resp. linéaire par morceaux) telle qu'on ait $|f(x) - g(x)| < \epsilon$ *pour tout* x *dans l'intervalle* I.

Preuve. Comme f est uniformément continue sur I, on peut trouver une partition de l'intervalle $a = x_0 < x_1 < ... < x_n = b$ telle que le saut de la fonction f satisfasse

$$w(f, [x_{i-1}, x_i]) < \epsilon \ (i = 1, 2, ..., n).$$

On a alors sur chaque sous-intervalle $[x_{i-1}, x_i]$ une approximation par la fonction constante $c_i = f(x_i)$ pour laquelle $|f(x) - c_i| < \epsilon$ si $x \in [x_{i-1}, x_i]$. Il en résulte une fonction escalier qui approxime f sur I dans un voisinage $\epsilon > 0$ de f pour chaque $x \in I$. De même on peut construire sur chaque $[x_{i-1}, x_i]$ la fonction linéaire reliant les points $(x_{i-1}, f(x_{i-1}))$ et $(x_i, f(x_i))$ qui donnera l'approximation désirée de f par une fonction linéaire par morceaux sur I. $\square$

CHAPITRE 4

Quelques fonctions remarquables

Il est indéniable que le développement de l'Analyse réelle doit, outre les concepts théoriques développés au cours des chapitres précédents, également faire une place adéquate aux nombreux exemples de fonctions qui ont motivé le développement de la théorie et ses applications. Alors que ces fonctions peuvent souvent être introduites intuitivement dans des cours élémentaires de Calcul Différentiel, le lecteur ou la lectrice se rendra compte au fil de ce chapitre que l'approche formelle développée dans un cours d'Analyse requiert plus de rigueur. Le but de ce chapitre est donc d'introduire rigoureusement, une fois pour toutes pourrait-on dire, quelques unes des fonctions qui sont à la base de l'étude de la théorie des fonctions d'une variable réelle.

1. Fonctions linéaires

On connait d'un cours d'Algèbre Linéaire les fonctions linéaires $f\colon \mathbb{R} \to \mathbb{R}$:

$$f(x+y) = f(x) + f(y) \ (\forall x, y \in \mathbb{R})$$

$$f(\alpha x) = \alpha f(x) \ (\forall \alpha, x \in \mathbb{R}).$$

Un exemple de fonction linéaire est donné par le choix de $\alpha \in \mathbb{R}$ en définissant $g\colon \mathbb{R} \to \mathbb{R}$ par $g(x) = \alpha x$. Il est immédiat que g est linéaire et, de plus, elle est continue puisque l'on a que $\{x_n\} \to x \iff \{\alpha x_n\} \to \alpha x$.

Proposition 1.1. *Soit $f\colon \mathbb{R} \to \mathbb{R}$ linéaire et continue. Alors il existe $\alpha \in \mathbb{R}$ tel que $f(x) = \alpha x$ pour tout $x \in \mathbb{R}$.*

Preuve. On donne brièvement l'idée et le lecteur ou la lectrice complètera l'argument. Posons $\alpha = f(1)$. On utilise alors la condition de linéarité de f pour montrer que $f(k) = f(1) \cdot k$ pour tout $k \in \mathbb{Z}$ et ensuite $f(r) = f(1) \cdot r$ pour tout $r \in \mathbb{Q}$. Jusqu'ici tout est parfaitement algébrique. La continuité de f est ensuite utilisée lorsque l'on montre qu'il découle de la densité des rationnels dans $\mathbb{R}$ que $f(x) = f(1) \cdot x$ pour tout $x \in \mathbb{R}$. $\square$

Remarque 1.2. *Il est possible de construire des fonctions linéaires qui ne sont pas continues sur $\mathbb{R}$, mais la construction est délicate et dépend de l'étude de $\mathbb{R}$ en tant qu'espace vectoriel sur le corps $\mathbb{Q}$.*

2. Fonctions exponentielles

On veut dans cette section définir formellement des fonctions $f\colon \mathbb{R} \to \mathbb{R}$ de la forme $f(x) = a^x$, où $a \in \mathbb{R}$ est une constante positive. Dans le cas où $r \in \mathbb{Q}$ on sait déjà ce qu'est a^r. En effet, si $r = \frac{p}{q}$, alors $a^r = (a^p)^{1/q}$ ou encore $(a^{1/q})^p$. Cette construction donne immédiatement l'identié $a^{r+s} = a^r \cdot a^s$ $(\forall r, s \in \mathbb{Q})$. Dans le cas de $x \in \mathbb{R}$ quelconque, on définit

$$a^x = \sup \{a^r \mid r < x\ ,\ r \in \mathbb{Q}\}.$$

Ceci définit une fonction f de $\mathbb{R}$ vers $\mathbb{R}$. On montre aisément que si $a > 1$ alors f est strictement croissante : par définition du supremum, si $x > r$ on a $a^x > a^r$ et si $y < s$ alors $a^y < a^s$. Alors pour x et y réels quelconques satisfaisant $x > y$, si l'on choisit $r, s \in \mathbb{Q}$ tels que $x > r > s > y$ on aura, par ce qui précède, $a^x > a^y$ tel que voulu.

Un cas particulièrement important est celui où $a = e = \lim_{n\to\infty}(1 + \frac{1}{n})^n$, qui définit la fonction exponentielle usuelle $f(x) = e^x$.

Proposition 2.1. *Pour tout $a > 0$, la fonction exponentielle $f(x) = a^x$ est continue sur $\mathbb{R}$*

Preuve. On donne l'idée de base et les détails seront matière à réflexion. On montre dans un premier temps la continuité de f en 0. Ceci découle essentiellement de la définition ci-dessus de a^x comme supremum (donnez les détails). Ensuite, il s'agit d'établir la formule $a^{x+y} = a^x \cdot a^y$ pour $x, y \in \mathbb{R}$ quelconques, et ceci est à nouveau établi à partir de la définition au moyen des suprema. Pour compléter la preuve dans le cas de la continuité en un point x_0 quelconque, on utilise alors le fait que

$$a^{x_0+h} - a^{x_0} = a^{x_o}(a^h - 1)$$

converge vers 0 par la continuité de a^x en 0. □

Pour généraliser notre construction au cas $a > 0$ (nous rappelons que nous avons supposé $a > 1$ ci-dessus), on utilise le fait que si $0 < a < 1$, alors $\frac{1}{a} > 1$ et le même genre d'idées que ci-dessus donne la continuité de $f(x) = a^x$ pour $a > 0$. On peut également montrer que si $a < 1$, alors la fonction $f(x) = a^x$ est strictement décroissante.

Comme dans le cas des fonctions linéaires continues, on peut se demander s'il existe une caractérisation des fonctions continues $g\colon \mathbb{R} \to \mathbb{R}$ satisfaisant $g(x+y) = g(x)g(y)$. Les fonctions $f(x) = a^x$ sont de telles fonctions par ce qui précède. En fait, on a

THÉORÈME 22. *Soit $g\colon \mathbb{R} \to \mathbb{R}$ continue, non-identiquement nulle et pour laquelle $g(x+y) = g(x)g(y)$. Alors il existe un $a > 0$ tel que $g(x) = a^x$ pour tout $x \in \mathbb{R}$*

Preuve. On ne donne que l'idée de base : montrez que $g(1) > 0$ et que pour tout $x \in \mathbb{Q}$ on a alors $g(x) = f(1)^x$, en utilisant les propriétés algébriques. La preuve est ensuite complétée en utilisant la continuité de la fonction g. $\square$

3. Fonctions logarithmiques

On rappelle que pour $a > 1$, la fonction $f(x) = a^x$ est continue, strictement croissante et elle prend toutes les valeurs strictement positives par le théorème de la valeur intermédiaire. Conséquemment, en invoquant le théorème de la fonction inverse, on sait qu'il existe une fonction inverse $f^{-1}\colon \mathbb{R}_*^+ \to \mathbb{R}$. On dénote cette fonction inverse $f^{-1}(x) = \log_a x$, et l'on parle du *logarithme en base a de x*. Cette fonction est ainsi strictement croissante et continue sur $\mathbb{R}_*^+$. Elle prend toutes les valeurs de $\mathbb{R}$ et on remarque également que, par construction, la fonction logarithme satisfait

$$\log_a(xy) = \log_a x + \log_a y \ (\forall x, y \in \mathbb{R}_*^+).$$

L'analogue de la caractérisation des fonctions exponentielles est alors énoncé sous la forme suivante :

THÉORÈME 23. *Soit $g\colon \mathbb{R}_*^+ \to \mathbb{R}$ continue, non-identiquement nulle et telle que $g(xy) = g(x) + g(y)$ pour tous $x, y \in \mathbb{R}_*^+$. Alors il existe un $a > 0$ tel que $g(x) = \log_a x$.*

Preuve. A nouveau, on donne l'idée de base et on laisse les détails en exercice. Il s'agit premièrement de montrer que g doit satisfaire $g(1) = 0$, $g(\frac{1}{x}) = -g(x)$ et qu'alors le théorème de la valeur intermédiaire ainsi que la propiété $g(x^n) = ng(x)$ impliquent que g prend toutes les valeurs de $\mathbb{R}$. Soit alors $a > 0$ tel que $g(a) = 1$. Alors pour tout $r = \frac{m}{n} \in \mathbb{Q}$, on a $g(a^r) = r$ et donc, par continuité de la fonction g on obtient $g(a^x) = x$ $(\forall x \in \mathbb{R})$. Ceci démontre bien que g est la fonction logarithme en base a tel que désiré. $\square$

4. Longueur d'arc d'une courbe

Soit $f\colon [a,b] \to \mathbb{R}$ continue. Son graphe $\mathcal{C} = \{(x, f(x)) \mid x \in [a,b]\}$ définit alors une courbe dans $\mathbb{R}^2$ et l'on aimerait avoir une notion de longueur de cette courbe. Soit Π une partition de $[a,b]$ en $a = x_0 < x_1 < ... < x_n = b$. et posons

$$l(f,\Pi) = \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2}.$$

Ceci n'est rien d'autre que la longueur de la courbe polygonale - donc linéaire par morceaux - associée à la fonction f et la partition Π. On remarquera qu'il découle immédiatement de l'inégalité du triangle que plus la partition est choisie *fine*, plus le nombre $l(f,\Pi)$ croit.

DÉFINITION 4.1. *La longueur de la courbe $\mathcal{C}$ est*

$$l(\mathcal{C}) = l(f) = \sup_{\Pi \text{ partition de } [a,b]} \{l(f,\Pi)\}.$$

Lorsque $l(\mathcal{C}) < \infty$ on dit que la courbe a une longueur finie.

EXEMPLE 4.2. *Si $f\colon [a,b] \to \mathbb{R}$ est continue et monotone, alors on a que son graphe est de longueur finie puisque*

$$l(f) \le |f(b) - f(a)| + |b - a| < \infty.$$

EXERCICE 4.3. *Démontrez que la fonction*

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

est une fonction continue sur $(0,1]$ dont la longueur satisfait $l(f) = \infty$.

Une classe importante de fonctions continues ayant une longueur finie est donnée par la notion de fonction *convexe* : $f\colon I \to \mathbb{R}$ continue est dite convexe si pour tous $x, y \in I$ et $0 \le t \le 1$, on a

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y). \tag{2}$$

Géométriquement, cela signifie qu'entre tous les x et y de I, le graphe de la fonction f est *en dessous* du segment de droite joignant les points $(x, f(x))$ et $(y, f(y))$ dans $\mathbb{R}^2$. (Faites un dessin si vous ne saisissez pas ceci).

THÉORÈME 4.4. *Soit $f\colon [a,b] \to \mathbb{R}$ une fonction continue et convexe. Alors $l(f) < \infty$.*

Preuve. On donne les éléments essentiels. Si f satisfait l'équation (2), alors toute fonction linéaire par morceaux g qui est inscrite dans le graphe de f satisfait également cette inégalité de convexité. Dans le cas d'une fonction linéaire par morceaux et convexe, on remarque que la pente des segments consécutifs ne peut qu'augmenter lorsqu'on se déplace de gauche à droite.
Il existe donc pour la fonction f un point $\xi \in [a, b]$ tel que :
- f soit décroissante sur $[a, \xi]$,
- f soit croissante sur $[\xi, b]$.

Cette observation implique alors le résultat puisqu'on a

$$l_{[a,b]}(f) = l_{[a,\xi]}(f) + l_{[\xi,b]}(f) \leq |f(\xi) - f(a)| + |\xi - a| + |f(b) - f(\xi)| + |b - \xi| < \infty.$$

$\square$

5. Fonction trigonométriques

Soit $f(x) = \sqrt{1 - x^2}$ considérée comme fonction définie sur $[-1, 1]$. Son graphe décrit la moitiée supérieure du cercle unité dans $\mathbb{R}^2$. Pour $x \in [-1, 1]$, soit $t = t(x)$ la longueur de f sur $[x, 1]$. On a alors

$$t \colon [-1, 1] \to [0, \pi]$$

et cette formule est prise comme une *définition* du nombre π, c'est-à-dire que $\pi = l_{[-1,1]}(f)$, la longueur d'arc de f. Notons également que, par construction, la fonction t est continue et bijective entre $[-1, 1]$ et $[0, \pi]$.

Définition 5.1. *Soit* $\sin \colon [0, \pi] \to \mathbb{R}$ *donnée par* $\sin t = \sin t(x) = f(x)$. *On appelle cette fonction la fonction sinus sur* $[0, \pi]$.

On étend alors la construction par symétrie autour de l'origine dans $\mathbb{R}^2$ pour obtenir $\sin \colon [-\pi, \pi] \to \mathbb{R}$, et ensuite par périodicité sur tout $\mathbb{R}$ en posant que $\sin t = \sin(t + 2\pi)$. On a alors par construction que

$$|\sin t| \leq 1 \ (\forall t \in \mathbb{R}),$$

$$|\sin s - \sin t| \leq |s - t| \ (\forall s, t \in \mathbb{R}).$$

Proposition 5.2. *La fonction sinus satisfait la limite* $\lim\limits_{t \to 0} \frac{\sin t}{t} = 1$.

Preuve. On donne l'idée essentielle. Soit (x, y) dans le premier quadrant du cercle unité $\{(x, y) \mid x^2 + y^2 = 1\}$. Comme la fonction f définie en début de section est décroissante sur $[x, 1]$, on établit à partir des définitions que

$$\sin t = y \leq t \leq y + (1 - x) = \sin t + 1 - x$$

et donc on en déduit que

$$1 \leq \frac{t}{\sin t} \leq 1 + \frac{1 - x}{\sqrt{1 - x^2}} \text{ si } x \neq 1.$$

Mais alors le fait que

$$\lim_{x \to 1} \frac{1 - x}{\sqrt{1 - x^2}} = 0$$

implique bien le résultat. □

Une fois la fonction sinus définie, on peut alors formellement introduire la fonction cosinus en posant

$$\cos t = \sqrt{1 - (\sin t)^2}.$$

Egalement, la fonction tangente pourra être définie par

$$\tan t = \frac{\sin t}{\cos t},$$

sur l'intervalle ouvert $(-\frac{\pi}{2}, \frac{\pi}{2})$ et ensuite étendue par linéarité à $\mathbb{R} - \{\pm\frac{\pi}{2} + 2k\pi \mid k \in \mathbb{Z}\}$.

Exercice 5.3. *Démontrez que la fonction cosinus satisfait*

$$\lim_{t \to 0} \frac{1 - \cos t}{t} = 0.$$

CHAPITRE 5

Fonctions dérivables

1. Définitions et propriétés élémentaires des fonctions dérivables

On se souvient qu'une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est continue en $a \in \mathbb{R}$ si l'on a

$$\lim_{x \to a} f(x) = f(a),$$

c'est-à-dire

$$\forall \epsilon > 0 \ \exists \delta > 0 \text{ tel que } |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Une autre façon d'écrire la même chose est :

$$\lim_{h \to 0} f(a + h) = f(a),$$

ou plus explicitement

$$\forall \epsilon > 0 \ \exists \delta > 0 \text{ tel que } |h| < \delta \Rightarrow |f(a + h) - f(a)| < \epsilon.$$

Ou encore

$$\lim_{h \to 0} f(a + h) - f(a) = 0.$$

Ceci devrait motiver quelque peu la définition suivante :

Définition 1.1. *Une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est dérivable en $a \in \mathbb{R}$ si la limite suivante existe :*

$$\lim_{h \to 0} \frac{f(a + h) - f(a)}{h}.$$

On appelle la fraction ci-dessus le quotient différentiel de la fonction f en $a \in \mathbb{R}$. On dénote alors le nombre réel correspondant à cette limite du quotient différentiel par

$$f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h},$$

que l'on appelle la *dérivée* de f en $a \in \mathbb{R}$. Formellement ceci signifie qu'il existe un nombre réel $\alpha \in \mathbb{R}$ tel que

$$\forall \epsilon > 0 \ \exists \delta > 0 \text{ tel que } 0 < |h| < \delta \Rightarrow \left|\frac{f(a + h) - f(a)}{h} - \alpha\right| < \epsilon$$

et le nombre α est alors noté $f'(a)$.

Remarque 1.2. *Pour qu'une fonction soit dérivable en $a \in \mathbb{R}$, il faut en particulier qu'elle y soit continue. En effet, si*

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

existe, comme on a que le dénominateur h tend vers 0, il faut au moins que que l'on aie

$$\lim_{h \to 0} f(a+h) - f(a) = 0,$$

ce qui est exactement la condition de continuité en a.

Il est facile de construire des exemples de fonctions qui sont continues sans être dérivables en un point $a \in \mathbb{R}$, par exemple la fonction $f(x) = |x|$ en $a = 0$, dont la limite du quotient différentiel n'existe pas puisque selon que l'on approche 0 par $h > 0$ ou $h < 0$, on obtient un résultat différent.

On pourra interpréter la dérivabilité d'une autre façon une fois que nous aurons introduit les notations "o" et 'O" qui donnent une idée de l'ordre de grandeur relatif de deux fonctions autour d'un point x_0 choisi :

On écrit $f(x) = o(g(x))$ si $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$.
On écrit $f(x) = O(g(x))$ si $\frac{f(x)}{g(x)}$ est bornée dans un voisinage de x_0.

Les mêmes notions s'appliquent lorsque l'on considère $x \to -\infty$ ou $x \to +\infty$. Donnons quelques exemples pour illustrer ces notions et leur lien avec la continuité et la dérivabilité des fonctions.

Exemple 1.3. *La notation $f(x) = o(1)$ est équivalente au fait que $\lim_{x \to x_0} f(x) = 0$. De même $f(x) = O(1)$ équivaut à dire que f est bornée autour de x_0. Le cas f bornée sur tout $\mathbb{R}$ est celui pour lequel $f(x) = O(1)$ sur tout $\mathbb{R}$. Donnez les détails des preuves pour vous familiariser avec ces notions.*

Exemple 1.4. *Soit g une fonction positive sur $\mathbb{R}$. Alors la notation $f(x) = O(g(x))$ signifie*

$$m \leq \frac{f(x)}{g(x)} \leq M.$$

Ce qui peut se réécrire (g partout positive !) comme

$$mg(x) \leq f(x) \leq Mg(x).$$

C'est-à-dire que f et g ont un comportement du même ordre.

EXEMPLE 1.5. *Soit $f(x) = 3x^2 + 2$. Alors $f(x) = o(x^3)$ mais $f(x) \neq o(x^2)$ lorsque $x \to \infty$:*

$$\frac{f(x)}{x^3} = \frac{3x^2+2}{x^3} \to 0 \qquad \textit{si } x \to \infty.$$
$$\frac{f(x)}{x^2} = \frac{3x^2+2}{x^2} \to 3 \qquad \textit{si } x \to \infty.$$

En termes de cette notation, on remarque que f est continue en a si et seulement si $f(x) = f(a) + o(1)$ lorsque x tend vers a. De même, nous pouvons interpréter la dérivabilité en ces termes, comme l'indique la proposition suivante :

PROPOSITION 1.6. *Une fonction $f : \mathbb{R} \to \mathbb{R}$ est dérivable en $a \in \mathbb{R}$ avec dérivée $f'(a)$ si et seulement si*

$$f(a+h) = f(a) + f'(a)h + o(h) \ \textit{autour de } a.$$

Preuve. f dérivable en a signifie que

$$\begin{aligned}
&\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = f'(a),\\
\iff\ &\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} - f'(a) = 0,\\
\iff\ &\frac{f(a+h)-f(a)}{h} - f'(a) = o(1),\\
\iff\ &f(a+h) - f(a) - f'(a)h = o(1)h,\\
\iff\ &f(a+h) = f(a) + f'(a)h + o(h).
\end{aligned}$$

□

Cette proposition nous amène vers une interprétation géométrique de la dérivée. En effet, on peut voir l'expression $f(a) + f'(a)h$ comme une application affine (c'est-à-dire une application linéaire plus d'une translation) et puisque $o(h)$ tend vers 0 lorsque $h \to 0$, la dernière formule dans la proposition précédente implique que la droite de $\mathbb{R}^2$ donnée par l'équation

$$y = f(a) + f'(a)x$$

est une approximation linéaire de f autour de a. Que veut-on dire exactement par *approximation linéaire* ? Prenons le cas où la fonction f serait linéaire au voisinage de $a \in \mathbb{R}$: $f(x) = \alpha x$ pour un certain $\alpha \in \mathbb{R}$ fixé. Alors on a

$$\begin{aligned} f(a+h) &= \alpha(a+h), \\ &= \alpha a + \alpha h, \\ &= f(a) + \alpha h. \end{aligned}$$

Donc on obtient

$$\frac{f(a+h)-f(a)}{h} = \frac{\alpha h}{h} = \alpha,$$

si bien que f est dérivable en a et $f'(a) = \alpha$. Dans ce cas-ci on a

$$f(a+h) = f(a) + f'(a)h,$$

si bien que $o(h) \equiv 0$ et l'approximation est *exacte*, ce qui correspond au fait que le graphe de f est exactement égal à la droite tangente.

Dans le cas général, la formule $f(a+h) = f(a)+f'(a)h+o(h)$ nous dit que l'application affine $f(a) + f'(a)h$ est une approximation de $f(a + h)$ à l'ordre 1 : la différence $g(h) = f(a + h) - (f(a) + f'(a)h)$ satisfait la condition

$$\lim_{h \to 0} \frac{g(h)}{h} = 0.$$

Donnons maintenant quelques exemples de fonctions dérivables car, si l'emphase est ici plutôt sur la théorie, on doit également être capable de vérifier la condition de dérivabilité dans des cas pratiques.

Exemple 1.7. *Commençons par un exemple simple... $f(x) = x^2$ dont on analyse le quotient différentiel en un point a quelconque :*

$$\frac{f(a+h)-f(a)}{h} = \frac{(a+h)^2 - a^2}{h} = 2a + h.$$

Cette expression a une limite lorsque h end vers 0 et donc f est dérivable en a et on trouve que $f'(a) = 2a$. Le fait que f soit dérivable sur tout $\mathbb{R}$ nous permet de construire sa fonction dérivée*, $f' \colon \mathbb{R} \to \mathbb{R}$ et dans ce cas-ci on a $f'(x) = 2x$ pour tout $x \in \mathbb{R}$.*

Exemple 1.8. **(Dérivabilité de la fonction exponentielle)**
On a rencontré au chapitre précédent la fonction exponentielle $f(x) = e^x$. Nous avons vu qu'elle satisfait $f(x+y) = f(x)f(y)$. Nous savons en outre que c'est la seule forme possible

pour une fonction continue satisfaisant cette dernière équation et dont la valeur en $x = 1$ *est le nombre* e*. Nous allons montrer que* $f(x) = e^x$ *est dérivable partout sur* $\mathbb{R}$ *et que* $f' \equiv f$ *sur tout* $\mathbb{R}$*. Contrairement à ce que l'on pourrait croire, ceci n'a rien d'évident... comme le lecteur ou la lectrice le verra ci-dessous.*

Premièrement, on invoque un peu de la théorie des séries développée au cours du chapitre 2. Considérons la fonction donné par une série de puissances

$$g(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ... = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Comme il s'agit d'une série qui converge absolument, le produit formel ci-dessous peut être évalué par un réarrangement quelconque :

$$(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ...) \cdot (1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + ...).$$

Pour les termes dont le degré total est n*, le produit ci-dessus donne*

$$\begin{aligned} \frac{x^n}{n!} + \frac{x^{n-1}y}{(n-1)!} + \frac{x^{n-2}y^2}{(n-2)!2!} + ... + \frac{y^n}{n!} &= \frac{1}{n!}(x^n + nx^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 + ... + y^n), \\ &= \frac{1}{n!}(x+y)^n. \end{aligned}$$

Donc la fonction continue g *satisfait* $g(x)g(y) = g(x + y)$*. De plus,* $g(1) = e$ *ce qui établit une nouvelle forme de notre fonction exponentielle :*

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ... = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Evaluons maintenant la dérivé de la fonction exponentielle à un point x_0 *en utilisant cette expression en série de puissances. Comme*

$$\frac{e^{x_0+h} - e^{x_0}}{h} = e^{x_0} \cdot \frac{e^h - 1}{h},$$

en remplaçant e^h *par la valeur de sa série de puissance, on obtient*

$$\begin{aligned} \frac{e^{x_0+h} - e^{x_0}}{h} &= e^{x_0}(1 + \frac{h}{2!} + \frac{h^2}{3!} + \frac{h^3}{4!} + ...) \\ &= e^{x_0} + e^{x_0}(\frac{h}{2!} + \frac{h^2}{3!} + \frac{h^3}{4!} + ...) \end{aligned}$$

Et donc

$$\frac{e^{x_0+h} - e^{x_0}}{h} - e^{x_0} = e^{x_0}(\frac{h}{2!} + \frac{h^2}{3!} + \frac{h^3}{4!} + ...).$$

Or si $|h| < 1$, *on a* $|\frac{h}{2!} + \frac{h^2}{3!} + \frac{h^3}{4!} + ...| < |h|$, *si bien que*

$$\frac{e^{x_0+h} - e^{x_0}}{h} - e^{x_0} = o(1).$$

Ce qui signifie exactement que la dérivée de e^x *en* x_0 *est* e^{x_0} *tel qu'annoncé.*

Etudions maintenant la classe des fonctions dérivables en nous intéressant au comportement des fonctions dérivables sous les opérations usuelles sur les fonctions, comme nous l'avons fait au cours du chapitre sur les fonctions continues. Ceci nous donnera un moyen efficace et rapide de démontrer que plusieurs fonctions sont dérivables sans trop d'efforts.

PROPOSITION 1.9. *Soient* $f: \mathbb{R} \to \mathbb{R}$ *et* $g: \mathbb{R} \to \mathbb{R}$ *fonctions dérivables. Alors*

(1) $\alpha f + g$ *est dérivable pour tout* $\alpha \in \mathbb{R}$ *et* $(\alpha f + g)'(x) = f'(x) + g'(x)$.

(2) fg *est dérivable et* $(fg)'(x) = f(x)g'(x) + f'(x)g(x)$.

(3) *Si* $f(x) \neq 0$ *alors* $\frac{1}{f}$ *est dérivable et* $(\frac{1}{f})'(x) = \frac{f'(x)}{f(x)^2}$.

Preuve.

(1) La preuve est immédiate et laissée en exercice.

(2) On a

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \phi(x),$$
$$\frac{g(x+h) - g(x)}{h} = g'(x) + \psi(x),$$

où l'on a $\lim_{h\to 0} \phi(h) = 0 = \lim_{h\to 0} \psi(h)$.

Alors

$$\begin{aligned}
\lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} &= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x+h)g(x)}{h} \\
&\qquad + \frac{f(x+h)g(x) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} [f(x+h)(g'(x) + \psi(h)) + g(x)(f'(x) + \phi(h))] \\
&= f(x)g'(x) + g(x)f'(x) \\
&\quad + \lim_{h\to 0} [(f(x+h) - f(x))g'(x) + f(x+h)\psi(h) + g(x)\phi(h)]
\end{aligned}$$

Le dernier terme tend vers 0 par la continuité de f et du fait que ϕ et ψ tendent vers 0 avec h, par hypothèse de dérivabilité de f et g. Ceci donne bien la formule de la dérivée d'un produit de fonctions.

(3) On regarde le quotient différentiel de $\frac{1}{f}$:

$$\begin{aligned}\frac{1/f(x+h)-1/f(x)}{h} &= \frac{(f(x)-f(x+h))/(f(x+h)f(x))}{h}\\ &= -\frac{f(x+h)-f(x)}{h}\cdot\frac{1}{f(x+h)f(x)}.\end{aligned}$$

Par continuité de f en x et du fait que $f(x)\neq 0$, on a alors effectivement la formule désirée lorsque l'on laisse h tendre vers 0 à la dernière équation.

□

THÉORÈME 24. **(Règle de dérivation en chaîne)**
Soient f et g fonctions dérivables sur $\mathbb{R}$. Alors la fonction composée $g\circ f$ est également dérivable sur $\mathbb{R}$ et sa dérivée satisfait

$$(g\circ f)'(x) = g'(f(x))f'(x).$$

Preuve. Cette preuve contient sa part de subtilités... Il est normal qu'elle soit difficilement digérée. Il est pratique de distinguer deux cas au cours de la preuve. Supposons premièrement que $f(x+h)-f(x)\neq 0$ si $0<|h|<h_0$ pour un certain $h_0\in\mathbb{R}$. Dans ce cas, pour le quotient différentiel de la fonction composée, on a

$$\begin{aligned}\lim_{h\to 0}\frac{(g\circ f)(x+h)-(g\circ f)(x)}{h} &= \frac{g(f(x+h))-g(f(x))}{h},\\ &= \lim_{h\to 0}\frac{g(f(x+h))-g(f(x))}{f(x+h)-f(x)}\cdot\frac{f(x+h)-f(x)}{h},\\ &= \lim_{h\to 0}\,[g'(f(x))+\psi(f(x+h)-f(x))]\cdot[f'(x)+\phi(h)],\end{aligned}$$

où la dernière ligne découle du fait que g est dérivable en $f(x)$ et f est dérivable en x (ceci requiert tout de même quelques moments de réflexion dans le cas du premier facteur). On écrit alors la dernière ligne ci-dessus comme

$$g'(f(x))f'(x) + [g'(f(x))\phi(h) + \psi(f(x+h)-f(x))\cdot(f'(x)+\phi(h))].$$

Comme l'expression entre crochets tend vers 0 (car f est continue et $\phi\to 0,\ \psi\to 0$ lorsque h tend vers 0), on a bien le résultat.

Dans le second cas, supposons que l'on a $f(x+h_n) = f(x)$ pour $\{h_n\} \to 0$. Alors comme f est dérivable, la condition ci-dessus implique que $f'(x) = 0$ (précisez pourquoi exactement). De plus $g \circ f(x+h_n) = g \circ f(x)$ implique de la même manière que, si $g \circ f$ est dérivable en x, alors $(g \circ f)'(x) = 0$ tel que voulu. Il ne reste donc qu'a montrer que $g \circ f$ est dérivable en x dans ce second cas. Lorsque l'on évalue le quotient différentiel de $g \circ f$ en x pour h tendant vers 0, on a deux types de valeurs de h : celles pour lesquelles $f(x+h) = f(x)$ et celles pour lesquelles $f(x+h) \neq f(x)$. Dans le premier cas, le quotient différentiel est nul pour ces valeurs de h. Dans le second type de valeurs de h, on peut utiliser de premier cas de notre preuve et obtenir

$$\frac{g \circ f(x+h) - g \circ f(x)}{h} = [g'(f(x) + \psi(f(x+h) - f(x))] \cdot \phi(h),$$

expression qui tend vers 0 lorsque $h \to 0$ (le lecteur ou la lectrice devra réfléchir quelques minutes aux hypothèses utilisées pour cette dernière affirmation). Dans tous les cas, le quotient différentiel tend vers la même limite, c'est-à-dire vers 0, et donc $g \circ f$ est dérivable en x. □

Remarque 1.10. **(Règle de chaîne et Algèbre Linéaire)**
Une façon de bien comprendre la règle de chaîne est de se rappeler quelques notions d'Algèbre Linéaire. On se souvient de deux choses. Premièrement, un réel $\alpha \in \mathbb{R}$ définit naturellement une transformation linéaire $\alpha\colon \mathbb{R} \to \mathbb{R}$ de l'espace vectoriel de dimension 1, $\mathbb{R}$, dans lui même par la formule $\alpha(v) = \alpha \cdot v$ pour $v \in \mathbb{R}$.
Ensuite, étant donné deux applications linéaires $\alpha, \beta\colon \mathbb{R} \to \mathbb{R}$, la composition $\alpha \circ \beta$ est tout simplement donnée par

$$\alpha \circ \beta(v) = \alpha(v) \cdot \beta(v).$$

Etant donné une fonction f dérivable en a, sa dérivée $f'(a) \in \mathbb{R}$ induit donc une application linéaire par la première observation. De même pour g en $f(a)$ et pour $g \circ f$ en a. De ce point de vue, le théorème de la règle de chaîne nous dit que l'application linéaire associée à la dérivée d'une composition de fonctions f et g est tout simplement égale à la composition des applications linéaires engendrées par f' et g' respectivement.

2. Propriétés de la dérivation

On verra plus tard dans le cours que toute fonction continue sur $\mathbb{R}$ est la dérivée d'une fonction sur $\mathbb{R}$. Pour l'instant, on remarquera qu'il y a des fonctions dont la dérivée existe mais n'est pas continue. En d'autres termes, g dérivable n'implique pas g' continue.

Exemple 2.1. *Considérons la fonction*

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}) & si\ x \neq 0 \\ 0 & si\ x = 0. \end{cases}$$

Cette fonction est partout dérivable avec dérivée donnée par

$$f'(x) = \begin{cases} 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}) & si\ x \neq 0 \\ 0 & si\ x = 0. \end{cases}$$

On remarque que f' n'est pas continue au point $x = 0$ puisque $\lim\limits_{x\to 0} \cos(\frac{1}{x})$ n'existe pas (on trouve facilement deux suites convergeant vers 0 pour lesquelles f égale $+1$ et -1 respectivement).

Considérons des fonctions $f_n\colon \mathbb{R} \to \mathbb{R}$ dépendant d'un paramètre $n \in \mathbb{N}$. On obtient alors une suite de fonctions $\{f_n(x)\} = \{f_n\}$. On dit que la suite de fonctions f_n converge vers la fonction f si l'on a, pour chaque $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \{f_n(x)\} = f(x).$$

Exemple 2.2. *On pourra étudier la limite des suites de fonctions suivantes :*

(1) $\{f_n(x)\} = \{e^{-x^2/n}\}$ *converge vers $f \equiv 1$ sur tout $\mathbb{R}$.*

(2) $\{f_n(x)\} = \{\sum\limits_{k=0}^{n} \frac{x^k}{k!}\}$ *converge vers e^x sur tout $\mathbb{R}$.*

Il est important de remarquer que si $\{f_n\}$ est une suite de fonctions continues qui a la propriété de converger, alors la fonction $\lim\limits_{n\to\infty} \{f_n\}$ n'est pas forcément continue :

Exemple 2.3. *Les fonctions $f_n(x) = x^n$ définissent une suite de fonctions convergente sur $[0,1]$ dont la limite est calculée de la façon suivante : pour $0 \leq x < 1$, on a $\lim\limits_{n\to\infty} \{f_n(x)\} = 0$, alors que pour $x = 1$ on a $\lim\limits_{n\to\infty} \{f_n(x)\} = 1$. La fonction limite est donc*

$$f(x) = \begin{cases} 0 & si\ x \in [0,1) \\ 1 & si\ x = 1 \end{cases}$$

et elle n'est pas continue en $x = 1$.

Proposition 2.4. *Soit $f\colon \mathbb{R} \to \mathbb{R}$ une fonction qui est la dérivée d'une fonction $g\colon \mathbb{R} \to \mathbb{R}$: $f = g'$. Alors f est limite de suite de fonctions continues.*

Preuve. On a g dérivable donc elle est continue. Posons $g_n(x) = g(x + \frac{1}{n})$ $(n \in \mathbb{N})$. Alors les fonctions g_n sont également continues par construction. Il en découle que la fonction $n(g_n - g)$ est également continue pour chaque $n \in \mathbb{N}$. Mais alors comme f est dérivée de g, on a, en particulier, que

$$\begin{aligned} f(x) &= \lim_{n\to\infty} \frac{g(x + \frac{1}{n}) - g(x)}{\frac{1}{n}} \\ &= \lim_{n\to\infty} n\ (g_n(x) - g(x)) \end{aligned}$$

ce qui prouve le résultat. □

Théorème 25. **(Théorème de Rolle)**
Soit $f\colon [a,b] \to \mathbb{R}$ un fonction dérivable telle que $f(a) = f(b)$. Alors il existe un point $\xi \in (a,b)$ tel que $f'(\xi) = 0$.

Preuve. Si f est constante sur $[a,b]$, c'est terminé puisqu'alors $f' \equiv 0$. Si f n'est pas constante, alors elle atteint son minimum ou son maximum en $\xi \in (a,b)$ puisque $f(a) = f(b)$. Supposons sans perte de généralité que ξ est un maximum. Soit $\{x_n\} \to \xi$ suite croissante et $\{y_n\} \to \xi$ suite décroissante. Comme ξ est maximum on a alors

$$\begin{aligned} \frac{f(x_n) - f(\xi)}{x_n - \xi} &\geq 0\ (\forall n > N), \\ \frac{f(y_n) - f(\xi)}{y_n - \xi} &\leq 0\ (\forall n > N), \end{aligned}$$

par construction des suites $\{x_n\}$ et $\{y_n\}$ et du fait que ξ est maximum. Ces deux inégalités sont vraies pour tout n assez grand, donc elles demeurent vraies à la limite. Dans les deux cas, le membre de gauche est égal à $f'(\xi)$ et les deux inégalités prises ensemble donnent donc bien $f'(\xi) = 0$. □

On note au passage que la preuve démontre de façon implicite le fait suivant : si $f\colon \mathbb{R} \to \mathbb{R}$ possède un maximum (ou un minimum) en un point $a \in \mathbb{R}$ et qu'elle est dérivable en a, alors on a $f'(a) = 0$. C'est-à-dire que la recherche d'extrema d'une fonction dérivable commence par la recherche des zéros de sa dérivée.

Théorème 26. **(Théorème des accroissements finis)**
Soit $f\colon [a,b] \to \mathbb{R}$ un fonction dérivable. Alors il existe un $\xi \in (a,b)$ tel que

$$f(b) - f(a) = f'(\xi)(b - a).$$

Preuve. Définissons la fonction auxilière

$$F(x) = (f(b) - f(a))(x - a) - (b - a)(f(x) - f(a)).$$

Alors F est dérivable sur $[a, b]$ puisque f l'est et elle satisfait, en outre, $F(a) = 0 = F(b)$. Par le théorème de Rolle, on a alors l'existence d'un $\xi \in (a, b)$ tel que $F'(\xi) = 0$. En remplaçant dans l'équation figurant dans l'énoncé du théorème, on obtient

$$\begin{aligned} & f(b) - f(a) - (b - a)f'(\xi) = 0, \\ \Rightarrow\ & f(b) - f(a) = f'(\xi)(b - a). \end{aligned}$$

□

On se souvient que si f est dérivable en a, on avait obtenu à la Proposition (1.6) que

$$f(a + h) = f(a) + f'(a)h + o(h).$$

Le théorème des acroissements finis nous dit que l'on peut prendre $o(h) = 0$ si l'on accepte de remplacer $f'(a)$ par $f'(\xi)$, où $\xi \in (a, a + h)$ ou alors $\xi \in (a + h, a)$. Ce résultat est important et permet de démontrer facilement plusieurs énoncés où de l'information sur la dérivée f' donne de l'information sur f.

Corollaire 2.5. *Soit $f\colon [a, b] \to \mathbb{R}$ dérivable avec $f' \equiv 0$ sur $[a, b]$. Alors f est constante sur $[a, b]$.*

Preuve. Pour tout $x \in [a, b]$ on a, pour un certain ξ entre a et x,

$$f(x) - f(a) = f'(\xi)(x - a) = 0.$$

Donc $f \equiv f(a)$ sur tout $[a, b]$. □

Corollaire 2.6. *Soient $f, g\colon [a, b] \to \mathbb{R}$ deux fonctions dérivables telles que $f' \equiv g'$ sur $[a, b]$. Alors pour tout $x \in [a, b]$, on a $f(x) = g(x) + c$ où c est une constante réelle.*

Preuve. En effet, $f' \equiv g'$ et la linéarité de l'opération de dérivation impliquent que $(f - g)' \equiv 0$. Donc $f - g$ est fonction constante sur $[a, b]$ par le dernier corollaire et le résultat est démontré. □

Corollaire 2.7. *Si $f\colon [a, b] \to \mathbb{R}$ est dérivable avec $f'(x) > 0$ pour tout $x \in [a, b]$, alors f est strictement croissante sur $[a, b]$. De même si $f'(x) < 0$ pour tout $x \in [a, b]$, alors f est strictement décroissante.*

Preuve. Soient $x, y \in [a,b]$ avec $x < y$. Par le théorème des accroissements finis, on a $\xi \in (x,y)$ tel que

$$f'(\xi) = \frac{f(y) - f(x)}{y - x}.$$

Comme $f'(\xi) > 0$ par hypothèse, on a bien $f(x) < f(y)$ tel que voulu. La preuve est identique dans le cas où f' est strictement négative sur $[a,b]$. □

EXERCICE 2.8. *Parmi les trois derniers corollaires, quels sont ceux dont la réciproque est également vraie ?*

3. Théorème de Taylor et fonctions infiniment dérivables

Soit $f\colon \mathbb{R} \to \mathbb{R}$ dérivable, donc $f'\colon \mathbb{R} \to \mathbb{R}$ existe. Si f' est dérivable, on peut définir $f''\colon \mathbb{R} \to \mathbb{R}$. Si f'' est dérivable, on peut définir $f'''\colon \mathbb{R} \to \mathbb{R}$. Et ainsi de suite... Soit donc

$$f^{(n)}\colon \mathbb{R} \to \mathbb{R}$$

la $n^{\text{ième}}$ dérivée de f si f est dérivable au moins n fois. On remarque que si f est au moins n fois dérivable, alors automatiquement $f', f'', ..., f^{(n-1)}$ sont des fonctions continues, en vertu du fait que les fonctions dérivables sont également continues. Dans le cas important où une fonction $f\colon \mathbb{R} \to \mathbb{R}$ possède des dérivées d'ordre quelconque, on dit que f est *infiniment dérivable* ou de classe $C^\infty(\mathbb{R})$.

DÉFINITION 3.1. *Si f est n fois dérivable et que $f^{(n)}$ est continue, on dit alors que f est de classe $C^n(\mathbb{R})$.*

Notons que l'on a immédiatement les inclusions $C^k(\mathbb{R}) \subset C^n(\mathbb{R})$ pour tout $k < n$. Le lecteur ou la lectrice démontrera que ces inclusions sont strictes.

Pour les fonctions de classe $C^n(\mathbb{R})$, on peut démontrer une version forte du théorème des accroissements finis :

THÉORÈME 27. **(Théorème de Taylor)**
Si $f\colon \mathbb{R} \to \mathbb{R}$ est dérivable $n+1$ fois, alors pour tous x et a dans $\mathbb{R}$, il existe un ξ entre ces deux points tel que

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Preuve. Nous appliquons le théorème de Rolle à la fonction auxiliaire suivante :

$$F(y) = f(x) - [f(y) + (x-y)f'(y) + ... + \frac{(x-y)^n}{n!} f^{(n)}(y)] + C(x-y)^{n+1},$$

où C est une constante (par rapport à la variable y) donnée par

$$C = \frac{1}{(x-a)^{n+1}} \cdot [-f(x) + f(a) + (x-a)f'(a) + ... + \frac{(x-a)^n}{n!} f^{(n)}(a)].$$

Il s'agit d'une fonction dérivable par construction satisfaisant en outre $F(a) = 0 = F(x)$, ce dont le lecteur ou la lectrice se convaincra après quelques instants de réflexion. On peut alors invoquer le théorème de Rolle et trouver ξ entre x et a tel que $F'(\xi) = 0$. Voici maintenant la subtilité dans cette preuve : en dérivant en ξ par rapport à y l'expression ci-dessus pour F, on obtient alors

$$0 = f'(\xi) - f'(\xi) + (x-\xi)f''(\xi) - (x-\xi)f''(\xi) + ... - \frac{(x-\xi)^n}{n!} f^{(n+1)}(\xi) - C(n+1)(x-\xi)^n$$

une expression pour laquelle on constate que tous les termes sont télescopés sauf pour le dernier et ceci nous donne

$$C = -\frac{1}{(n+1)!} f^{(n+1)}(\xi).$$

En combinant ceci avec la définition de C, on obtient

$$-\frac{1}{(n+1)!} f^{(n+1)}(\xi) = \frac{1}{(x-a)^{n+1}} \cdot [-f(x) + f(a) + (x-a)f'(a) + ... + \frac{(x-a)^n}{n!} f^{(n)}(a)]$$

qui peut être réarrangée pour donner

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

□

On remarque que ce résultat nous dit que l'on peut approcher une fonction $C^\infty(\mathbb{R})$ par des polynômes de degré n, où l'erreur commise est

$$\frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\xi).$$

EXEMPLE 3.2. *Autour de $a = 0$ on peut donner une expression pour $f(x) = \sin x$:*

$$f'(x) = \cos x \ , \ f''(x) = -\sin x \ , \ f^{(3)}(x) = -\cos x \ , \ ...$$

Ainsi

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^8}{8!} \sin \xi,$$

où ξ est entre 0 et x.

On en arrive alors à l'idée de série de Taylor associée à une fonction infiniment dérivable, donnée comme la série de puissances

$$\sum_{n=1}^{\infty} \frac{(x-a)^n}{n!} f^{(n)}(a).$$

Plusieurs questions naturelles se posent à propos de la série de Taylor associée à une fonction : (1) est-ce que la série converge ? (2) converge-t-elle vers $f(x)$?

Une première observation est de constater que $\sum_{n=1}^{\infty} \frac{(x-a)^n}{n!} f^{(n)}(a)$ converge vers $f(x)$ si et seulement si

$$\lim_{n\to\infty} \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\xi_n) = 0.$$

En effet, si l'on regarde l'expression

$$f(x) - \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) = \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\xi_n),$$

en prenant la limite de chaque côté lorsque n tend ver l'infini, on obtient dans le membre gauche la différence entre $f(x)$ et la série de Taylor de f en a. Cette différence sera nulle exactement lorsque

$$\lim_{n\to\infty} \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\xi_n) = 0.$$

Exemple 3.3. *Considérons $f(x) = e^x$. La série de Taylor de cette fonction en $a = 0$ peut être immédiatement calculée et est égale à*

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ... + \frac{x^n}{n!} + ...$$

Une application directe du test du quotient donne que pour tout $x \in \mathbb{R}$ fixé, on a

$$\lim_{n\to\infty} \frac{x^{n+1}}{(n+1)!} e^{\xi_n} = 0.$$

Donc la série de Taylor de $f(x)$ converge vers $f(x)$ pour tout réel x.

Exemple 3.4. *Considérons $f(x) = \ln x$. Sa série de Taylor autour de $a = 1$ est donnée par*

$$(x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} + ... + (-1)^{n+1} \frac{(x-1)^n}{n} + ...$$

Cette série a la propriété quelle converge pour certaines valeurs de x et diverge pour d'autres :

(1) $|x-1| \leq 1$ *: la série de Taylor converge par comparaison avec la série harmonique alternée.*

(2) $|x-1| > 1$ *: la série diverge puisque son terme général ne tend même pas vers* 0*.*

EXEMPLE 3.5. *Nous construisons maintenant une fonction f infiniment dérivable dont la série de Taylor converge en tout $x \in \mathbb{R}$, mais pas vers la valeur $f(x)$ de la fonction ! Soit en effet*

$$f(x) = \begin{cases} e^{-1/x^2} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

Si $x > 0$, on a

$$\begin{aligned} f'(x) &= \frac{1}{x^3} e^{-1/x^2} \\ f''(x) &= \frac{-3}{x^4} e^{-1/x^2} + \frac{1}{x^6} e^{-1/x^2} \end{aligned}$$

et de façon générale, on voit après quelques instants de réflexion que $f^{(n)}(x)$ est de la forme $p(x)e^{-1/x^2}$, où $p_n(x)$ est un certain polynôme. Or puisque pour chaque $n \in \mathbb{N}$ et chaque $x \in \mathbb{R}$, on a

$$e^x > \frac{x^n}{n!},$$

on obtient facilement que

$$\lim_{x \to 0^+} p(1/x) e^{-1/x^2} = 0.$$

Par ailleurs, pour $x \leq 0$ on a que f est identiquement nulle. Ces deux observations mènent à la conclusion que f est infiniment dérivable partout sur $\mathbb{R}$ et en 0, on a calculé que

$$f^{(n)}(0) = 0 \ (\forall n \in \mathbb{N}).$$

Ceci implique que la série de Taylor de f converge vers la fonction nulle autour de 0, une fontion qui est distincte de f dans tout voisinage de 0 par construction.

CHAPITRE 6

Introduction à l'intégration

Dans le dernier chapitre, nous voulons introduire ce qui sera, en quelque sorte, l'opération inverse de la dérivation des fonctions. Il y a de nombreuses approches pour traiter de la théorie élémentaire de l'intégration. Par tradition plus qu'autre chose, une introduction à l'Analyse réelle privilégie en général le traitement de l'intégration par l'entremise de l'intégrale de Riemann. Dans cette optique, on laisse pour plus tard l'intégration au sens de Lebesgue qui peut même souvent être introduite en tant qu'exemple de théorie de la mesure abstraite.
Dans ce chapitre, nous prenons parti pour l'intégrale de Lebesgue. En effet la théorie de Lebesgue a aujourd'hui un siècle et il ne nous semble pas qu'il y ait de raisons valables pour ne pas l'introduire à un niveau élémentaire comme dans un cours d'Analyse de première année. Outre cette raison de nature plutôt philosophique, il est également important de remarquer que l'approche proposée ici (dont l'origine remonte aux travaux de Riesz et Nagy dans les années 1950) utilise très peu d'outils et que c'est probablement la façon la plus rapide d'arriver à un énoncé et une preuve du Théorème fondamental du calcul différentiel et intégral. Nous invitons le lecteur ou la lectrice curieux à comparer cette approche à la théorie de Riemann en consultant une des références données à la fin de ce cours. Pour une approche plus abstraite de l'intégrale de Lebesgue nous recommandons d'attendre un cours plus avancé...

1. Fonctions escalier

Commençons par quelques notions préliminaires. D'une part, étant donné un intervalle borné $I = (a, b)$ on a précédemment dans le cours rencontré la notion de *longueur* de l'intervalle i, $l(I) = b - a$. Si $I' = [a, b]$, on a alors $l(I') = l(I)$, donc pour l'utilisation de cette notion on considèrera toujours des intervalles ouverts.

PROPOSITION 1.1. *La longueur d'un intervalle satisfait les propriétés élémentaires suivantes :*

(1) $I_1 \subseteq I_2 \;\Rightarrow l(I_1) \leq l(I_2)$.

(2) $I = I_1 \cup I_2 \Rightarrow l(I) \leq l(I_1) + l(I_2)$ *avec égalité si et seulement si* $I_1 \cap I_2 = \varnothing$. *En général, on a*

$$l(I) = l(I_1) + l(I_2) - l(I_1 \cap I_2).$$

On a ensuite besoin de la notion de fonction caractéristique que nous introduisons dans le contexte des intervalles mais qui peut être aisément étendue à tout sous-ensemble de $\mathbb{R}$.

DÉFINITION 1.2. *On appelle fonction caractéristique d'un intervalle* (a, b) *la fonction*

$$\chi_{(a,b)}(x) = \begin{cases} 1 & si\ x \in (a, b) \\ 0 & si\ x \notin (a, b). \end{cases}$$

On remarque que $\chi_{(a,b)}$ est une fonction définie sur tout $\mathbb{R}$. Elle possède des points de discontinuité en a et b puisque $w(f, a) = 1 = w(f, b)$, mais elle est continue partout ailleurs. Cette notion très simple nous permet de définir ce qui pourrait être considéré comme l'objet central de ce chapitre :

DÉFINITION 1.3. *On dit que* $\varphi \colon \mathbb{R} \to \mathbb{R}$ *est une fonction escalier si on peut l'exprimer comme*

$$\varphi = \sum_{k=1}^{n} c_k \chi_{I_k},$$

où $c_1, c_2, ..., c_n \in \mathbb{R}$ *et* $I_1, I_2, ..., I_n$ *sont des intervalles bornés.*

L'ensemble $\{I_n\}$ est appelé le *support* de φ. Il découle de la définition qu'une fonction escalier est précisément une fonction qui peut s'exprimer comme combinaison linéaire de fonctions caractéristiques d'intervalles dans $\mathbb{R}$. Bien que très simple, cette notion a le léger désavantage suivant : une fonction escalier φ peut s'exprimer de multiples façons comme combinaison linéaire de fonctions caractéristiques, en variant à la fois les coefficient et les intervalles utilisés dans la définition. Ce problème serait aisément résolu en imposant que les intervalles I_k soient tous disjoints, mais on verra un autre façon de procéder plus tard.

On peut très aisément définir l'intégrale d'une fonction escalier de la façon suivante :

DÉFINITION 1.4. *L'intégrale d'une fonction escalier* $\varphi = \sum\limits_{k=1}^{n} c_k \chi_{I_k}$ *est définie comme*

$$\int \varphi = \sum_{k=1}^{n} c_k l(I_k).$$

Il est important de saisir que, à priori, cette notion pourrait dépendre de la représentation comme combinaison linéaire de fonctions caractéristiques. Une partie de la théorie développée dans les pages qui suivent est implicitement dédiée à cette question.

Soit $\varphi = \sum_{k=1}^{n} c_k \chi_{I_k}$ où $I_k = (a_k, b_k)$ $(k = 1, 2, ..., n)$. Désignons par $\alpha_0 < \alpha_1 < ... < \alpha_m$ les points terminaux des intervalles $I_1, I_2, ..., I_n$. On a tout simplement réorganisé les a_i et b_j de façon croissante. Ceci nous donne ce que nous appellerons la *partition standard* associée a ϕ. On peut alors démontrer la proposition suivante.

Proposition 1.5. *Soit* $\varphi = \sum_{k=1}^{n} c_k \chi_{I_k}$ *où* $I_k = (a_k, b_k)$ $(k = 1, 2, ..., n)$ *avec sa partition standard donnée par* $\alpha_0 < \alpha_1 < ... < \alpha_m$. *Alors*

(1) φ *est constante sur chaque* (α_{i-1}, α_i) $(i = 1, 2, ..., m)$.

(2) $\varphi \equiv 0$ *à l'extérieur de* (α_0, α_m).

(3) φ *prend un nombre fini de valeurs (et donc* φ *est fonction bornée).*

(4) φ *possède un nombre fini de points de discontinuité.*

Preuve. On laisse la preuve en exercice. □

Les propriétés (1) et (2) peuvent être prises comme définition pour la notion de fonction escalier, c'est-à-dire qu'une fonction f est fonction escalier si et seulement si il existe $\beta_0 < \beta_1 < ... < \beta_m$ tels que (1) et (2) soient satisfaites. On appelle alors cet ensemble $\beta_0 < \beta_1 < ... < \beta_m$ une *partition* associée à la fonction escalier f.

2. La notion de raffinement

Soit $\alpha_0 < \alpha_1 < ... < \alpha_m$ partition standard pour une fonction escalier $\varphi = \sum_{k=1}^{n} c_k \chi_{I_k}$. Si l'on rajoute des subdivisions de (α_{i-1}, α_i), la fonction φ demeure localement constante sur chaque sous-intervalle par construction. C'est-à-dire que si l'on a, pour $\{\alpha_0, \alpha_1, ..., \alpha_m\} \subseteq \{\beta_0, \beta_1, ..., \beta_p\}$, que l'ensemble $\{\beta_0, \beta_1, ..., \beta_p\}$ constitue également une partition pour φ. On dit alors que l'on a construit un *raffinement* de $\{\alpha_0, \alpha_1, ..., \alpha_m\}$ pour la fonction escalier φ.

Notons par ailleurs que l'on peut rendre disjoints les supports de $\varphi = \sum_{k=1}^{n} c_k \chi_{I_k}$ de la façon suivante : si $\alpha_0 < \alpha_1 < ... < \alpha_m$ est une partition pour φ, on définit $K_1, K_2, ..., K_{2m+1}$

comme les intervalles $(\alpha_0, \alpha_1), ..., (\alpha_{m-1}, \alpha_m), [\alpha_o, \alpha_0], ..., [\alpha_m, \alpha_m]$. Alors par construction

$$\varphi = \sum_{k=1}^{2m+1} b_k \chi_{K_k},$$

où $K_1, ..., K_{2m+1}$ sont des ensembles disjoints.

L'intérêt de la notion de raffinement est qu'étant donné deux représentations de la même fonction escalier $\varphi\colon \mathbb{R} \to \mathbb{R}$

$$\sum_{k=1}^{n} c_k \chi_{I_k} \ , \ \sum_{l=1}^{m} b_l \chi_{J_l},$$

on peut trouver un raffinement commun $K_1, ..., K_p$ de $I_1, ..., I_n$ et de $J_1, ..., J_m$ tel que la fonction escalier s'exprime sous la forme

$$\varphi = \sum_{i=1}^{p} e_i \chi_{K_i}.$$

De la même manière, étant donné deux fonctions escalier φ et ψ, on peut trouver un raffinement commun de leurs supports tel que

$$\varphi = \sum_{i=1}^{p} e_i \ \chi_{K_i} \text{ et } \psi = \sum_{i=1}^{p} f_i \ \chi_{K_i}.$$

Définition 2.1. *Soit $L^0(\mathbb{R})$ l'ensemble de toutes les fonctions escalier sur $\mathbb{R}$.*

Proposition 2.2. *Si l'on a $\varphi, \psi \in L^0(\mathbb{R})$, alors*

(1) $\alpha\varphi + \beta\psi \in L^0(\mathbb{R})$ $(\forall \alpha, \beta \in \mathbb{R})$.

(2) $\varphi\psi \in L^0(\mathbb{R})$.

(3) $|\varphi|, \max\{\varphi, \psi\}, \min\{\varphi, \psi\} \in L^0(\mathbb{R})$.

Preuve. Exercice. □

3. Lemme fondamental et propriétés de l'intégrale

Lemme 3.1. **(Lemme fondamental)**
Soit φ une fonction escalier admettant deux descriptions

$$\sum_{k=1}^{n} c_k \chi_{I_k} \ , \ \sum_{l=1}^{n'} c'_k \chi_{I'_k}.$$

Alors on a

$$\sum_{k=1}^{n} c_k l I_k = \sum_{l=1}^{m} c'_k l(I'_k).$$

Preuve. A COMPLÉTER. □

L'importance de ce résultat est claire : il permet de conclure que la définition que l'on a faite de l'intégrale des fonctions dans $L^0(\mathbb{R})$ est bien définie, c'est-à-dire qu'elle ne dépend pas de la représentation choisie pour le calcul de l'intégrale.

PROPOSITION 3.2. **(Linéarité de l'intégrale)**

$$\int \alpha\varphi + \beta\psi = \alpha \int \varphi + \beta \int \psi.$$

Preuve. La preuve découle immédiatemet de la définition de l'intégrale et d'un choix de représentant pour φ et ψ. □

PROPOSITION 3.3. *Soient $\varphi, \psi \in L^0(\mathbb{R})$ telles que $\varphi \geq \psi$ partout sur $\mathbb{R}$. Alors*

$$\int \varphi \geq \int \psi.$$

Preuve. Par soustraction il suffit de démontrer que $\varphi \geq 0$ implique $\int \varphi \geq 0$. Soit alors

$$\varphi = \sum_{k=1}^{n} c_k \chi_{I_k},$$

où les I_k sont disjoints par hypothèse. Alors $\varphi \geq 0$ est équivalent à dire que $c_k \geq 0$ pour $1 \leq k \leq n$. Mais alors on a bien

$$\int \varphi = \sum_{k=1}^{n} c_k l(I_k) \geq 0.$$

□

PROPOSITION 3.4. *Pour toute $\varphi \in L^0(\mathbb{R})$, on a $|\int \varphi| \leq \int |\varphi|$.*

Preuve. En effet, comme on a $|\varphi| \pm \varphi \geq 0$, on a par la Proposition (3.3) que $\int |\varphi| \geq \pm \int \varphi$ ce qui donne immédiatement le résultat. □

PROPOSITION 3.5. *Si $\varphi \in L^0(\mathbb{R})$, alors on peut l'exprimer comme $\varphi = \varphi^+ - \varphi^-$, où φ^+, φ^- sont des fonctions positives, et l'on a alors*

$$\int \varphi = \int \varphi^+ - \int \varphi^-.$$

Preuve. En effet, si l'on pose $\varphi^+ = \max\{0, \varphi\}$ ainsi que $\varphi^- = \max\{0, -\varphi\}$, alors par construction on a deux fonctions escalier positives dont la différence est clairement égale à φ. La formule pour l'intégrale découle alors de la Proposition (3.2). □

Il est important de remarquer que dans la théorie que l'on développe ici, si l'on modifie $\varphi \in L^0(\mathbb{R})$ en des points *isolés*, la nouvelle fonction $\widetilde{\varphi}$ aura la même intégrale. Ceci s'explique aisément par la fait qu'un tel changement introduit une nouvelle valeur pour chaque point x_0 isolé pour lequel on a modifié φ, mais également un nouveau sous-intervalle $[x_0, x_0]$ dans la partition exprimant la nouvelle fonction. Comme la longueur de ce nouvel intervalle est nulle, on n'a pas changé la valeur de l'intégrale selon la définition utilisée. En fait, plus tard on verra que la condition de point isolé est bien plus forte que nécessaire... la notion appropriée dans notre contexte sera celle d'un *ensemble de mesure nulle* sur laquelle on reviendra.

Corollaire 3.6. *Soit $\varphi \in L^0(\mathbb{R})$ une fonction escalier positive. Alors on a $\int \varphi = 0$ si et seulement si $\varphi \equiv 0$ sauf peut-être en un nombre fini de points.*

Preuve. En effet, si l'on a $\int \varphi = 0$ pour $\varphi = \sum_{k=1}^{n} c_k \chi_{I_k}$, alors le fait que la fonction soit positive par hypothèse implique que : soit $c_k = 0$, c'est-à-dire que φ est nulle sur I_k, ou alors que $l(I_k) = 0$, c'est-à-dire que l'intervalle I_k ne consiste que d'un point $x_k \in \mathbb{R}$. Dans tous les cas, on obtient exactement la conclusion du corollaire. □

Corollaire 3.7. *Si $\varphi \geq \psi$ sauf en un nombre fini de points, alors $\int \varphi \geq \int \psi$.*

Preuve. Clair. □

4. Intégration des fonctions continues

On passe maintenant à l'étude de l'intégration des fonctions continues sur un intervalle $[a, b]$. On dénote par $C([a, b])$ l'espace vectoriel de telles fonction continues, et soit $L^0([a, b])$ l'espace vectoriel des fonctions escalier sur $\mathbb{R}$ qui sont nulles à l'extérieur de $[a, b]$.

Etant donné $f \in C([a.b])$ considérons l'ensemble suivant :

$$E_f = \{\varphi \in L^0([a, b]) \mid \varphi \leq f\}.$$

L'ensemble E_f décrit donc toutes les fonctions escalier sur $[a, b]$ qui sont bornées supérieurement par f. On peut alors donner une définition de l'intégrale d'une fonction continue :

DÉFINITION 4.1. *L'intégrale d'une fonction $f \in C([a,b])$ est définie comme*

$$\int_a^b f = \sup_{\varphi \in E_f} \int \varphi.$$

L'intuition derrière ceci est qu'en approchant inférieurement une fonction continue f par des fonctions escalier possédant toutes une intégrale et en prenant le supremum de ces intégrales sur toutes les fonctions escaliers bornées supérieurement par f, on devrait converger vers un réel pouvant être interprété comme l'intégrale de f. Mais il faut tout d'abord montrer que cette définition a du sens !
Comme f est continue sur $[a,b]$ ensemble fermé et borné, on sait qu'il existe $M \in \mathbb{R}$ tel que

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

Alors la fonction $M\chi_{[a,b]}$ appartient à E_f, qui n'est donc pas l'ensemble vide. De plus, pour toute $\varphi \in E_f$, on a par construction que $\varphi \leq M\chi_{[a,b]}$, si bien que $\int \varphi \leq M(b-a)$. Cela signifie que l'ensemble

$$\{\int \varphi \mid \varphi \in E_f\}$$

est borné supérieurement et donc il possède un supremum et ainsi la définition (4.1) est rigoureuse.

On remarque que $\int_a^a f = 0$ par construction de l'intégrale. Posons également par convention que

$$\int_b^a f = -\int_a^b f.$$

On est alors en mesure de donner plusieurs propriétés de l'intégrale pour les fonctions continues. Ces propriétés découlent, pour l'essentiel de la définition de l'intégrale, de la continuité de f et de propriétés analogues pour les fonctions escalier.

PROPOSITION 4.2. *Soient $f, g \in C([a,b])$. Alors*

(1) *Si $f \geq g$, on a $\int_a^b f \geq \int_a^b$.*
(2) *Si $m \leq f \leq M$ sur $[a,b]$, alors*

$$m(b-a) \leq \int_a^b f \leq M(b-a).$$

(3) *Si* $f \geq 0$ *et que* $\int_a^b f = 0$, *alors* $f \equiv 0$ *sur* $[a, b]$

On remarquera que la propriété (3) est plus forte que celle démontrée pour les fonctions escaliers, et c'est ici la continuité de f qui permet cela. Ce n'est pas la seule différence d'ailleurs. Ainsi, la linéarité de l'intégrale des fonctions continues n'est pas si facile à démontrer à partir de la définition de l'intégrale donnée ci-dessus. Pour comprendre ceci, l'étudiant ou l'étudiante pourra, par exemple, essayer de se convaincre que les ensembles E_{f+g}, E_f et E_g ou encore les ensembles E_f et E_{-f} sont de nature bien différente. On reviendra donc à la question de la linéarité de l'intégrale lorsque l'on aura quelques outils de plus.

PROPOSITION 4.3. *Soient* $f \in C([a, b])$ *et* $a < c < b$. *Alors*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Preuve. On démontre l'égalité en deux temps grâce à deux inégalités réciproques. D'une part, si $\varphi \in L^0(\mathbb{R})$, on observe que l'on peut écrire φ comme

$$\varphi = \varphi\chi_{[a,c]} + \varphi\chi_{[c,b]}.$$

Pour l'intégrale on a alors

$$\begin{aligned}\int_a^b f &= \int_a^b \varphi\chi_{[a,c]} + \int_a^b \varphi\chi_{[c,b]} \\ &= \int_a^c \varphi\chi_{[a,c]} + \int_c^b \varphi\chi_{[c,b]} \\ &\leq \int_a^b f + \int_c^b f.\end{aligned}$$

Comme ceci est vrai pour toute fonction $\varphi \in E_f$, en prenant le supremum l'inégalité demeure vraie et donc

$$\int_a^b f \leq \int_a^b f + \int_c^b f.$$

Par ailleurs, si $\psi \in E_f$ sur $[a, c]$ et $\theta \in E_f$ sur $[c, b]$, alors on a que $\psi + \theta \in E_f$ sur $[a, b]$ (peut-être après avoir ajusté la valeur au point c, ce qui n'a aucune influence sur le calcul d'intégrales). On a alors

$$\int_a^b f \geq \int_a^c \varphi + \int_c^b \theta,$$

et en prenant le supremum pour chaque membre de droite, on obtient

$$\int_a^b f \geq \int_a^b f + \int_c^b f.$$

$\square$

On arrive alors au résultats central de ce chapitre. Le premier théorème sera seulement énoncé pour l'instant puisqu'on l'applique immédiatement pour donner une preuve du théorème fondamental du calcul différentiel et intégral. On revient alors sur la preuve du résultat sur lequel ce dernier théorème est fondé.

THÉORÈME 28. **(Théorème de l'intégrale indéfinie)**
Soit $f \in C([a, b])$. On appelle intégrale indéfinie $F\colon [a, b] \to \mathbb{R}$ de f la fonction

$$F(x) = \int_a^x f.$$

Alors la fonction F est dérivable sur $[a, b]$ et satisfait $F' = f$ sur $[a, b]$.

Ce théorème peut être interprété de la façon suivante : l'intégrale indéfinie d'une fonction continue est en quelque sorte l'opération inverse de la dérivation des fonctions. Ceci n'est pas tout-à-fait exact puisque F est continûment dérivable, alors que la dérivée d'une fonction quelconque n'est pas forcément continue. Mais on comprend tout de même le principe général mis de l'avant.

THÉORÈME 29. **(Théorème fondamental du calcul différentiel et intégral)**
Soit $f\colon [a, b] \to \mathbb{R}$ dérivable avec $f'\colon [a, b] \to \mathbb{R}$ continue. Alors

$$\int_a^b f' = f(b) - f(a).$$

Preuve. Posons $F(x) = \int_a^x f'$. Par le Théorème de l'intégrale indéfinie, on a F dérivable et $F' \equiv f'$ sur $[a, b]$. Il s'en suit que $F - f$ est fonction constante sur $[a, b]$, et comme $F(a) = 0$, on doit avoir $F(x) = f(x) - f(a)$ pour tout $x \in [a,b]$. Mais alors on obtient bien

$$\int_a^b f' = F(b) = f(b) - f(a).$$

$\square$

En principe, ceci réduit l'intégration des fonctions à la recherche de primitives... pour autant que de telles primitives puissent être trouvées ! Or rien ne garantit qu'une fonction quelconque admette une primitive. Le cas des fonctions continues en est un très avantageux de ce point de vue puisqu'on garantit l'existence d'une primitive. Mais certaines fonctions très simples à définir ne possèdent pas de primitive et elles sont pourtant intégrables :

Exemple 4.4. *Considérons la fonction définie sur tout* $\mathbb{R}$ *par*

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

La fonction f *ne possède pas de primitive : en effet, une primitive devrait avoir une dérivé étant continue au sens de Darboux. Or la fonction* f *n'est pas continue au sens de Darboux, puisque la valeur* $+1$*, par exemple, n'est pas atteinte par* f *sur* $\mathbb{R}$*. Remarquons qu'il est aisé de démontrer que* f *est intégrable sur tout segment* $[-a,a]$ *en décomposant le segment en* $[-a,0)$ *et* $(0,a]$ *sur lesquels la fonction est constante, donc égale à une fonction escalier.*

Donnons maintenant la preuve du théorème de l'intégrale indéfinie.

Preuve. On doit montrer que l'on a la limite suivante

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = f(x),$$

où $F(x) = \int_a^x f$. Pour $x, x+h \in [a,b]$ et $h > 0$ (le cas $h < 0$ est similaire) on a

$$F(x+h) - F(x) = \int_x^{x+h} f.$$

Pour le membre de droite, on a la double inégalité suivante qui est claire de par la définition de l'intégrale :

$$h \times \inf\{f(t) \mid t \in [a, a+h]\} \leq \int_x^{x+h} f \leq h \times \sup\{f(t) \mid t \in [a, a+h]\}.$$

Etant donné que l'on a supposé f continue, le théorème de la valeur intermédiaire peut être appliqué pour obtenir l'existence d'un élément $\xi \in [a, a+h]$ tel que

$$hf(\xi) = \int_x^{x+h} f.$$

Ceci peut être réécrit comme

$$\frac{F(x+h) - F(x)}{h} = f(\xi).$$

Le nombre h étant arbitraire, en prenant la limite lorsque h tend vers 0 et en utilisant à nouveau la continuité de f on a bien

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = f(x).$$

$\square$

Une conséquence immédiate du théorème de l'intégrale indéfinie est la preuve de la linéarité de l'intégrale de fonctions continues :

PROPOSITION 4.5. *Si $f, g \colon [a, b] \to \mathbb{R}$ sont continues et que $\alpha \in \mathbb{R}$, alors*

$$\int_a^b (f + \alpha g) = \int_a^b f + \alpha \int_a^b g.$$

Preuve. Par la linéarité de la dérivation de fonctions dérivables et le théorème de l'intégrale indéfinie, on a

$$\begin{aligned} \frac{d}{dx}(\int_a^x f + \alpha \int_a^x g) &= \frac{d}{dx} \int_a^x f + \alpha \, \frac{d}{dx} \int_a^x g \\ &= f(x) + \alpha g(x) \\ &= \frac{d}{dx} \int_a^x (f + \alpha g). \end{aligned}$$

Donc les expressions

$$\int_a^x f + \alpha \int_a^x g \qquad \text{et} \qquad \int_a^x (f + \alpha g)$$

diffèrent d'une constante C (indépendante de x). En posant $x = a$ on trouve que $C = 0$ ce qui complète la preuve du résultat.

CHAPITRE 7

Problèmes supplémentaires

Allant au-delà des exercices permettant aux étudiants de vérifier leur compréhension de base de la matière et de développer une certaine intuition analytique, les problèmes rassemblés dans ce chapitre on plutôt pour but d'approfondir certaines notions au-delà d'un premier courts d'Analyse mathématique et ils devraient constituer de jolis défis pour les étudiants avancés qui auraient encore plus de temps ` consacrer à cette matière. Nous insistons toutefois sur une chose : il ne faut pas se décourager si certains problèmes viennent à bout des efforts les plus coriaces !

Problème 1. **(Principe d'induction de Artin)** *Pour étudier des propriétés de la fonction Gamma, Emil Artin a introduit le principe d'induction suivant :*

Soit $S \subseteq \mathbb{N}$ *un ensemble tel que*

(1) $1 \in S$.

(2) $k \in S \Rightarrow 2k \in S$, *pour tout* k.

(3) $k + 1 \in S \Rightarrow k \in S$, *pour tout* k.

Alors on a $S = \mathbb{N}$.

Démontrez que le Principe d'induction de Artin est équivalent au Principe d'induction mathématique introduit au Chapitre 1.

Problème 2. **(Equivalences de Principes d'induction)** *Nous avons énoncé au Chapitre 1 deux formes du Principe d'induction mathématique :*
(I) *Principe d'induction sur les sous-ensembles* $S \subseteq \mathbb{N}$.
(II) *Principe d'induction sur les propriétés indexées par* $\mathbb{N}$.

Voici d'autres formes en apparence plus fortes :

(III) *Principe d'induction complète sur les sous-ensembles $S \subseteq \mathbb{N}$: comme (I) mais on suppose que les entiers inférieurs ou égaux à k sont dans S.*

(IV) *Principe d'induction complète sur les propriétés indexées par /mathbbN : comme (II) mais on suppose la propriété $\mathcal{P}$ vraie pour tout entier inférieur ou égal à k.*

(V) *Principe du bon ordre dans $\mathbb{N}$: Tout sous-ensemble non-vide $S \subseteq \mathbb{N}$ possède un plus petit élément.*

Montrez que ces 5 formulations sont toutes équivalentes. Certaines sont immédiates, comme **(II)** $\iff$ **(II)**, *ou encore* **(III)** $\iff$ **(IV)**. *Pour compléter l'équivalence, montrez par exemple le cycle d'implications :* **(I)** $\Rightarrow$ **(III)** $\Rightarrow$ **(V)** $\Rightarrow$ **(I)**.

PROBLÈME 3. *Démontrez le Théorème de Dedekind : Si E et F sont deux sous-ensembles disjoints non-vides de $\mathbb{R}$ pour lesquels :*

(1) $\mathbb{R} = E \cup F$.

(2) $x \in E$ *et* $y \in F \Rightarrow x < y$.

montrez qu'il existe un unique $a \in \mathbb{R}$ tel que : $x \in E, y \in F \Rightarrow x \leq a \leq y$.

PROBLÈME 4.

PROBLÈME 5.

PROBLÈME 6.

PROBLÈME 7.

Remarques d'édition :

Les coupures de Dedekind sont bien couvertes dans Rudin, m'elanger cela à la référence.

Justifier l'absence de topologie.

Parler d'ensemble de Cantor.

Bibliographie

[Boa] R.P. Boas, *A primer on real numbers*, Wiley and sons, New York (1960)

[God] R. Godement, *Analyse mathématique I*, Springer, Berlin (1998)

[Har] G. H. Hardy, *A course of pure mathematics*, Cambridge University Press, Cambridge (1952)

[Pri] H. A. Priestley *Introduction to integration*, Oxford University Press, Oxford (1997)

[Rud] W. Rudin, *Principles of mathematical analysis*, McGraw-Hill, New York (1976)

[Spi] M. Spivak, *Calculus*, Publish or Perish, Houston (1984)