Calcul différentiel et intégral

R. Petit

Année académique 2015 - 2016

Contents

1 Intuitions

1.1 Les dérivées

1.1.1 Définition

Alors que l'analyse nait au XVII^e siècle, ce n'est que pendant le XIX^e que les outils nécessaires à une expression rigoureuse ont été à disposition des mathématiciens. L'analyse a pour notion centrale celle de *variation*. Intuitivement, la notion de variation instantanée d'une quantité $f(t)$ peut être décrite de la sorte :

$$V(t, \Delta t) = \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

pour une valeur Δt étant *de plus en petite*. On dit donc que Δt *tend vers 0*. Attention cependant car le manque de rigueur et le manque d'outils adaptés à la manipulation de données infinitésimales amènent à des paradoxes et des résultats illogiques voire inexplicables.

Selon la définition de variation vue ci-dessus, nous pouvons exprimer la dérivée comme étant la variation instantanée d'une fonction $f : \mathbb{R} \to \mathbb{R}$ que nous notons $f'(t)$. Plus précisément, la variation *instantanée* implique que Δt soit *infiniment petit*, ce qui s'écrit ainsi :

$$f'(x) = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

Remarque. Pour des *petites valeurs* de Δt, l'approximation suivante est admissible :

$$f(t + \Delta t) \simeq f(t) + f'(t)\Delta t$$

1.1.2 Exemples

Fonction affine Soit une fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto ax + b$. La dérivée de f est l'expression suivante :

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{a(x+h) + b - (ax+b)}{h} = \lim_{h \to 0} \frac{ax + ah + b - ax - b}{h} = \lim_{h \to 0} \frac{ah}{h} = a$$

Il *doit* sembler intuitif qu'étant donné que la fonction f est une fonction dont le graphe est une droite, sa dérivée (donc sa variation instantanée) est constante sur tout son domaine du fait que la variation d'une fonction uniformément (dé)croissante est constante.

Plus spécifiquement, si $a = 0$, la fonction est une fonction *constante* : $f : \mathbb{R} \to \mathbb{R} : x \mapsto b$. De ce fait, sa dérivée doit être nulle car une fonction croissante n'a pas de variation (par définition). La quantification de cette variation doit donc être représentée par 0. Cela peut se montrer également aisément depuis la définition de la dérivée :

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{b - b}{h} = 0$$

Fonction du second degré Soit une fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto ax^2 + bx + c$. Sa dérivée peut être calculée ainsi :

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{a(x+h)^2 + b(x+h) + c - (ax^2 + bx + c)}{h} \\ &= \lim_{h \to 0} \frac{ax^2 + 2axh + ah^2 + bx + bh + c - ax^2 - bx - c}{h} = \lim_{h \to 0} \frac{2axh + ah^2 + bh}{h} = \lim_{h \to 0} 2ax + ah + b = 2ax + b \end{aligned}$$

Ici, la dérivée est une fonction de x, ce qui indique que selon la valeur de x à laquelle nous voulons évaluer la dérivée, la variation représentée est susceptible de différer.

Fonction valeur absolue Soit une fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} +x \text{ si } x \geqslant 0 \\ -x \text{ si } x < 0 \end{cases}$. Afin de déterminer l'accroissement instantané de f au point d'abscisse $x = 0$, il faut repasser par la définition :

$$f'(a) = \lim_{h \to 0} V_f(a, h)$$

Donc $f'(0)$ peut être déterminé de la manière suivante :

$$f'(0) = \lim_{h \to 0} V_f(0, h) = \lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0} \frac{f(h)}{h}$$

Les cas $h < 0$ et $h \geqslant 0$ doivent être traités séparément de par la définition de la fonction. Pour le premier cas, la dérivée est :

$$f'(0) = \lim_{h \to 0} \frac{h}{h} = 1$$

Et pour le second cas,

$$f'(0) = \lim_{h \to 0} \frac{-h}{h} = -1$$

Le résultat est tout à fait cohérent par rapport au graphique de la fonction $x \mapsto |x|$ car cette fonction a deux accroissements différents : l'un à gauche de 0, l'autre à droite de 0. Cependant, il est impossible d'exprimer la valeur de la dérivée de f au point 0. On dit de f qu'elle n'est pas dérivable en 0.

Fonction exponentielle Soit une fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto a^x$ avec $a \in \mathbb{R}_0^+$. À nouveau, pour définir sa dérivée, il nous faut repasser par la définition :

$$f'(x) = \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} = \lim_{h \to 0} \frac{a^x(a^h - 1)}{h} = a^x \lim_{h \to 0} V(0, h) = a^x f'(0)$$

Ce développement nous indique que pour connaitre la dérivée de f au point x, il nous faut connaitre la dérivée de f au point 0. Il est possible d'observer sur des esquisses de graphique que $f'(0)$ dépend de a de manière croissante. Faire croître a impliquera une croissance de $f'(0)$. Il existe cependant un nombre $\in \mathbb{R}$ tel que $f'(0) = 1$. Ce nombre est $e \simeq 2.718$, ce qui implique $f'(x) = e^x f'(0) = e^x = f(x)$. La fonction $x \mapsto e^x$ (et ses multiples) sont leur propre dérivée.

1.1.3 Dérivée d'ordre supérieur

Étant donné que la dérivée d'une fonction quelconque f (en supposant qu'elle est dérivable) est également une fonction
($f' : \mathbb{R} \to \mathbb{R} : x \mapsto f'(x)$), nous pouvons à nouveau dériver cette fonction (en supposant que la dérivée soit toujours dérivable). Il est alors question de *dérivée seconde*, telle que $f''(x) = (f')'(x) = ((f)')'(x)$. Pour noter la k^e dérivée, il existe la notation suivante : $f^{(k)}$ (car répéter k fois le symbole *prime* (′) est contre-productif en temps de lecture **et** d'écriture).

Ces mêmes dérivées d'ordre supérieur à 1 ont leur importance et leur cohérence : tant $f'(x)$ est la variation de la quantité $f(x)$, tant $f''(x)$ est la variation de $f'(x)$. La dérivée seconde représente donc l'accroissement de l'accroissement (ou la variation de la variation). Elle nous donne donc une information sur comment la variation évolue (en dynamique, si on représente la position d'un mobile par la fonction $x(t)$, $x'(t)$

représente la variation de la position, à savoir la vitesse (donc $x'(t) = v(t)$). Cependant, $x''(t)$ représente la variation de la vitesse, à savoir l'accélération, d'où $x''(t) = v'(t) = a(t)$).

De plus, la dérivée seconde a une autre interprétation graphique : si $f''(x) > 0$, nous savons que $f'(x)$ a une pente positive donc $f'(x)$ est croissante. Cela implique que $f(x)$ est croissante aussi mais de plus en plus croissante. Autrement dit, le graphe de $f(x)$ est *concave* aux alentours de $(x, f(x))$. De manière similaire, lorsque $f''(x) < 0$, $f(x)$ est de moins en moins décroissante et donc le graphe de $f(x)$ est *convexe* aux alentours de $(x, f(x))$.

1.1.4 Notation de Leibniz

En reprenant la définition de la variation donnée plus haut, nous pouvons définir δf et δx comme étant respectivement la variation de la quantité $f(x)$ et la variation de la quantité x. Nous avons donc :

$$V(x, h) = \frac{f(x+h) - f(x)}{h} = \frac{\delta f}{\delta x}$$

Cela implique que nous ayons $\delta f = f(x+h) - f(x)$, ce qui est bien la variation de la quantité $f(x)$, et que nous ayons $\delta x = h = (x+h) - (x)$, ce qui est bien la variation de la quantité x.

Or, nous avions défini la dérivée comme étant la variation instantanée, à savoir $f'(x) = \lim_{h \to 0} V(x, h)$, qui selon la notation de Leibniz correspond à $f'(x) = \lim_{h \to 0} \frac{\delta f}{\delta x}$. Leibniz a cependant instauré une seconde notation correspondant non plus à la variation comme la notation δ mais bien à la dérivée. Cette notation est :

$$\frac{df}{dx} = \lim_{h \to 0} \frac{\delta f}{\delta x}$$

Attention cependant à ne pas utiliser cette valeur comme étant un quotient de nombres réels : la notation $\frac{df}{dx}$ n'a de sens que lorsque la limite a été faite. Et comme $\lim_0 \frac{F(h)}{G(h)} \neq \frac{\lim_0 F(h)}{\lim_0 G(h)}$, nous ne pouvons pas séparer df et dx[1].

Leibniz a également eu besoin d'une notation pour la k^e dérivée. Ayant considéré $\frac{d}{dx}$ comme étant une opération à réaliser k fois, il a noté la k^e dérivée de la sorte :

$$\frac{d^k f}{dx^k}$$

1.2 Règles de dérivation

Somme Soient $f, g : \mathbb{R} \to \mathbb{R}$, deux fonction dérivables en a. De manière intuitive, nous pouvons dire que $(f+g)'(a) = f'(a) + g'(a)$ car l'accroissement de la fonction de somme est la somme des accroissements : au point a, la fonction $(f+g)$ *subit* un accroissement égal à l'accroissement de f plus l'accroissement de g.

Produit Soient $f, g : \mathbb{R} \to \mathbb{R}$, deux fonction dérivables en a. Contrairement à ce que nous serions tentés de dire naïvement, nous ne pouvons pas définir $(fg)'(a) = f'(a)g'(a)$. C'est à Leibniz que l'on doit cette démonstration.

Considérons un rectangle de dimensions $f(a)$ et $g(a)$. La quantité $(fg)(a)$ correspond à l'aire de ce rectangle. Si on passe de a à $a+h$, nous obtenons un nouveau rectangle de dimensions $f(a+h)$ et $g(a+h)$. Selon l'approximation vue au point 1.1.1., nous savons que $f(a+h) \simeq f(a) + f'(a)h$. L'aire du nouveau rectangle

[1] Du moins pas dans le cas d'une dérivée.

est donc $A = f(a+h)g(a+h) \simeq (f(a)+f'(a)h)(g(a)+g'(a)h) = f(a)g(a)+f'(a)g(a)h+f(a)g'(a)h+f'(a)g'(a)h^2$. Rappelons tout de même que la dérivée ici est égale à :

$$\begin{aligned}(fg)'(a) &= \lim_{h\to 0}\frac{(fg)(a+h)-(fg)(a)}{h} = \lim_{h\to 0}\frac{f'(a)g(a)h+f(a)g'(a)h+f'(a)g'(a)h^2}{h} \\ &= \lim_{h\to 0} f'(a)g(a)+f(a)g'(a)+f'(a)g'(a)h = f'(a)g(a)+f(a)g'(a)\end{aligned}$$

Nous avons donc la variation instantanée de la fonction (fg) au point a : $(fg)'(a) = f'(a)g(a)+f(a)g'(a)$.

Composition Soient $f, g : \mathbb{R} \to \mathbb{R}$, deux fonctions telles que g est dérivable en a, et f est dérivable en $g(a)$. Il est toujours possible de trouver la dérivée de la fonction composée $(f \circ g)(x) = f(g(x))$ à l'aide de l'approximation $f(a+h) \simeq f(a)+f'(a)h$ tant que h est *petit*. Notre dérivation devient donc :

$$\begin{aligned}(f\circ g)'(a) &= \lim_{h\to 0}\frac{(f\circ g)(a+h)-(f\circ g)(a)}{h} = \lim_{h\to 0}\frac{f(g(a+h))-f(g(a))}{h} = \lim_{h\to 0}\frac{f(g(a)+g'(a)h)-f(g(a))}{h} \\ &= \lim_{h\to 0}\frac{f(g(a))+f'(g(a))g'(a)h-f(g(a))}{h} = f'(g(a))g'(a)\end{aligned}$$

Réciproque Soit $f : \mathbb{R} \to \mathbb{R}$. Si f est bijective, nous pouvons définir sa fonction réciproque $f^{-1} : \mathbb{R} \to \mathbb{R}$. Si en plus, nous supposons que $f'(x) \neq 0 \forall x \in \mathbb{R}$, nous pouvons exprimer la dérivée de la fonction réciproque comme suit :

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}$$

selon le développement suivant (nécessitant la dérivée d'une composée) : en sachant que la dérivée de la fonction identité est 1 ($(x \mapsto x)'(x) = 1$) et que la composée d'une fonction f avec sa réciproque (f^{-1}) correspond à la fonction identité, nous pouvons trouver la dérivée de la réciproque.

$$\begin{aligned}id'(x) &= 1 \\ (f^{-1}\circ f)'(x) &= 1 \\ (f^{-1})'(f(x))f'(x) &= 1 \\ (f^{-1})'(f(x)) &= \frac{1}{f'(x)}\end{aligned}$$

Ou encore :

$$(f^{-1})'(a) = \frac{1}{f'((f^{-1})(a))}$$

Remarque. De manière graphique, ce résultat peut s'interpréter de la manière suivante : la réciproque d'une fonction bijective est son image par symétrie orthogonale d'axe $y = x$. En traçant cet axe de symétrie, on peut observer qu'aux points $(a, f(a)) \in \Gamma_f$ et $(b, (f^{-1})(b)) \in \Gamma_{f^{-1}}$ tel que $b = f(a)$, nous avons un rapport entre les pentes des tangentes. Plus précisément, nous pouvons observer que la pente de la tangente au second point ($(f^{-1})'(b) = (f^{-1})'(f(a))$) correspond à l'inverse de la pente de la tangente au second point. Donc nous avons bien la même égalité : $(f^{-1})'(f(a)) = \frac{1}{f'(a)}$.

Exemple des fonctions exponentielles et logarithmiques Il a été vu plus haut que la fonction $x \mapsto \exp(x)$ admettait pour dérivée elle-même. Nous avons également vu ci-dessus que $(f^{-1})'(a) = \frac{1}{f'((f^{-1})(a))}$. En étant tenté de trouver $f(x)$ telle que $f'(x) = f(x)\, \forall x \in \operatorname{dom} f$, nous aurions donc $(f^{-1})'(a) = \frac{1}{f((f^{-1})(a))} = \frac{1}{a}$. Cependant, nous connaissons une fonction de ce genre : la fonction exponentielle de base e. Sa fonction inverse, la fonction logarithmique de base e a sa dérivée calculée de la sorte :

$$\log'(x) = \frac{1}{\exp(\log(x))} = \frac{1}{x}$$

1.3 Les intégrales

1.3.1 Définition

En analyse, il existe un outil permettant de calculer l'aire en dessous d'une courbe. Cet outil s'appelle *intégrale*. Soit une fonction $f : \mathbb{R} \to \mathbb{R}$, nous définissons l'aire entre la courbe de f et l'axe horizontal $y = 0$ définie entre les droites $x = a$ et $x = b$ de la manière suivante :

$$A = \int_a^b f(x)\,dx$$

Remarque. Cette aire est dite *signée*. C'est à dire que les parties *en dessous* de l'axe $y = 0$ sont représentés par une valeur réelle négative alors que les parties *au dessus* de l'axe $y = 0$ sont représentés par une valeur réelle positive.

Afin d'approximer l'aire que nous tentons de déterminer, nous découpons l'intervalle $[a, b]$ en n *morceaux* $[x_{i-1}, x_i]$ $1 \leqslant i \leqslant n$ avec $x_i = a + (b-a)\frac{i}{n}$. Ces *arrêtes* nous permettent d'obtenir des rectangles de hauteur respective $f(x_{i-1})$. L'aire de chacun de ces rectangles est donc $b \times h = (x_i - x_{i-1})f(x_{i-1})$. L'approximation de l'aire recherchée (donc de l'intégrale) est la somme des aires de ces rectangles.

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

Pour obtenir l'aire exacte, il faut faire tendre n vers $+\infty$.

Exemple de la fonction carrée Si nous désirons trouver l'aire sous la courbe de $x \mapsto x^2$ sur $[0, t]$, nous la déterminons ainsi :

$$\begin{aligned}
\int_0^t x^2\,dx &= \lim_{n\to\infty} \sum_{i=1}^{n} (x_i - x_{i-1})(x_{i-1})^2 = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{t}{n}\left(t\frac{i-1}{n}\right)^2 = \lim_{n\to\infty} \frac{t^3}{n^3}\sum_{i=1}^{n}(i-1)^2 \\
&= \lim_{n\to\infty} \frac{t^3}{n^3}\frac{n(2n-1)(n-1)}{6} = \lim_{n\to\infty} \frac{t^3(2n-1)(n-1)}{6n^2} = t^3 \lim_{n\to\infty} \frac{2n^2+1-3n}{n^2} \\
&= t^3 \lim_{n\to\infty}\left(2 + \frac{1}{n^2} - \frac{3}{n}\right) = t^3\frac{2}{6} = \frac{t^3}{3}
\end{aligned}$$

1.4 Théorème fondamental du calcul différentiel et intégral

1.4.1 1ère version

En analyse, il existe un théorème (parfois appelé *théorème fondamental de l'analyse*) qui définit la dérivation et l'intégration comme deux opérations inverses l'une de l'autre. Avant d'exprimer le théorème, tentons de donner un peu d'intuition au rapport entre dérivation et intégration.

Soit $F(x)$, une fonction réelle définie par $F(x) = \int_a^x f(t)\,dt$ telle que f est une fonction réelle continue. Admettons que $F(x)$ est dérivable sur l'intégralité de son domaine. Nous pouvons donc déterminer sa dérivée comme étant l'accroissement instantané :

$$F'(x) = \lim_{h\to 0} V_F(x, h) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{1}{h}\int_x^{x+h} f(t)\,dt$$

Ce que représente $\int_x^{x+h} f(t)\,dt$ est l'aire en dessous de la courbe de f entre les points d'abscisse x et $x+h$. h étant tout petit (tendant vers l'infini), nous pouvons approximer $f(t)$ constant entre x et $x+h$. L'aire représentée par l'intégrale est donc $\simeq hf(x)$. Nous pouvons donc poursuivre le calcul de la dérivée de F :

$$F'(x) = \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)\,dt = \lim_{h\to 0} \frac{1}{h} hf(x) = f(x)$$

Nous avons donc $F(x) = \int_a^x f(t)\,dt$ et $f = F'(x)$. Nous pouvons dès à présent énoncer une première version du théorème :

Soit $f : \mathbb{R} \to \mathbb{R}$, *une fonction continue sur* $\mathbb{R}$, *alors la fonction* $F(x)$ *définie par*

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

est dérivable sur l'intégralité de son domaine tel que $F'(x) = f(x)$.

1.4.2 2nde version

Il existe une deuxième manière d'énoncer ce théorème. À nouveau, tentons de le trouver intuitivement. En ayant une fonction $f : \mathbb{R} \to \mathbb{R}$ dérivable en tout point de son domaine, nous pouvons considérer l'intégrale de a en b de $f'(x)$ comme étant la somme des variations instantanées de f sur l'intervalle $[a, b]$. Tentons maintenant de définir plus précisément ce que vaut cette intégrale.

$$\int_a^b f'(t)\,dt = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{b-a}{n} f'(x_{i-1})$$

Rappel Au point 1.1.1., nous avons vu l'approximation suivante : $f(a+h) \simeq f(a) + f'(a)h$. En réorganisant cette approximation, nous avons $f'(a)h \simeq f(a+h) - f(a)$. Dans notre cas, si nous posons $h = \frac{b-a}{n}$, nous pouvons poursuivre notre intégrale.

$$\begin{aligned} \int_a^b f'(t)\,dt &= \lim_{n\to\infty} \sum_{i=1}^{n} hf'(x_{i-1}) = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i-1}+h) - f(x_i) \\ &= \lim_{n\to\infty} f(x_n) - f(x_0) = f(b) - f(a) \end{aligned}$$

Nous avons considéré cette intégrale comme étant la somme des variations instantanées de f entre a et b, et le résultat de ce développement est que la somme des variations instantanées est égale à la variation totale pour aller de a à b.

Le théorème peut donc être exprimé d'une deuxième manière :

Soit $f : \mathbb{R} \to \mathbb{R}$, *une fonction réelle dérivable, alors*

$$\int_a^b f'(t)\,dt = f(b) - f(a).$$

1.5 Primitives

En analyse, il est fréquent de devoir trouver F tel que $F'(x) = f(x)$. C'est donc une recherche de fonction. Cependant, comme nous venons de le voir, trouver une fonction en connaissant sa dérivée revient à réaliser une intégrale (version 1 du théorème fondamental de l'analyse). Une solution F satisfaisant $F'(x) = f(x)$ est appelée une *primitive* de f. $\int f$ ou $\int f(t)\,dt$ sont les manières les plus courantes d'écrire *primitive de* f[2]. Cependant, comme le laisse comprendre la seconde version du théorème fondamental de l'analyse, il existe une infinité de primitives, toutes définies à *une constante près*. Nous généralisons donc « la » primitive de f en $\int f + C$, $C \in \mathbb{R}$.

1.6 Règles d'intégration

Tout comme il y a des règles de dérivation (point 1.1.5.), il existe des règles d'intégration.

Produit Tout comme $(fg)'(x) \neq f'(x)g'(x)$, $\int (fg)(x)\,dx \neq \int f(x)\,dx \int g(x)\,dx$. Pour réussir à intégrer un produit, il faut partir de la règle de Leibniz pour la dérivation :

$$\begin{aligned}
(fg)'(x) &= f'(x)g(x) + f(x)g'(x) \\
\int_a^b (fg)'(x)\,dx &= \int_a^b (f'(x)g(x) + f(x)g'(x))\,dx \\
[(fg)(x)]_a^b &= \int_a^b (f'(x)g(x) + f(x)g'(x))\,dx \\
[(fg)(x)]_a^b &= \int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx \\
\int_a^b f'(x)g(x)\,dx &= [(fg)(x)]_a^b - \int_a^b f(x)g'(x)\,dx
\end{aligned}$$

Il faut donc, pour pouvoir intégrer un produit, considérer un des deux facteurs comme étant une dérivée (qu'il faudra donc intégrer pour avancer).

Composition La règle *d'intégration en chaine* est également appelée *changement de variable*. Elle peut être exprimée comme suit. Soient f et g, deux fonctions continûment dérivables, $g(\alpha) = a$ et $g(\beta) = b$, alors :

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(g(t))g'(t)\,dt.$$

Cela veut dire qu'une intégration peut être résolue en *transformant* l'intégrale de manière à faire apparaitre une composition. La justification de cette formule peut être donnée comme suit :

$$(F(g(t)))'(x) = F'(g(t))g'(t) = f(g(t))g'(t)$$

Donc :

$$\int_\alpha^\beta f(g(t))g'(t)\,dt = F(g(\beta)) - F(g(\alpha)) = F(b) - F(a) = \int_a^b f(t)\,dt.$$

1.7 Les équation différentielles

Une équation différentielle est une équation dont l'inconnue est une fonction f et dans laquelle les dérivées de f apparaissent.

[2] On parle parfois également d'intégrale indéfinie.

1.7.1 équation différentielle linéaire d'ordre 1

Comme le dit la pseudo-définition ci-dessus, une équadiff est une équation où l'inconnue est une fonction. Le cas de $F = \int_a^x f(t)\,dt$ est un cas particulier d'une grande famille d'équadiffs que l'on appelle *équation différentielle linéaire d'ordre 1*. Cette famille est caractérisée par la forme suivante :

$$f'(x) + p(x)f(x) = q(x).$$

Une manière de résoudre une telle équation est de déterminer une fonction auxiliaire $a(x)$ que l'on va multiplier de part et d'autre de l'égalité :

$$a(x)f'(x) + a(x)p(x)f(x) = q(x)a(x)$$

Le but de cette manipulation est de pouvoir faire ressortir $a(x)f'(x) + a'(x)f(x)$, ce qui est égal à $(af)'(x)$. Cependant, pour cela, il faut choisir $a(x)$ telle que $a'(x) = a(x)p(x)$. Il existe une solution :

$$a(x) : x \mapsto e^{\int p(x)\,dx}$$

Maintenant que nous avons cette fonction, il suffit de résoudre $(af)'(x) = (aq)(x)$. Si (aq) est continue, le théorème fondamental de l'analyse assure l'existence d'une primitive $b(x) = \int a(x)q(x)\,dx$. À présent, nous savons que si $(af)'(x) = (aq)(x)$, alors $(af)(x) = b(x)$, ou encore $f(x) = \frac{b(x)}{a(x)}$.

Nous avons donc une solution à l'équation de départ :

$$f(x) = \frac{\int \left(e^{\int p(x)\,dx}\right) q(x)\,dx}{e^{\int p(x)\,dx}}.$$

1.7.2 Unicité de la solution et problème de Cauchy

Lorsque l'on *transforme* le problème initial (équadiff linéaire d'ordre 1) en un problème conditionné (en précisant $f(x_0) = y_0$), nous limitons le nombre de solutions à 1. Si nous avons deux fonctions f_1 et f_2, solutions d'une équadiff linéaire d'ordre 1, et que nous définissons une autre fonction g telle que $g(x) = (f_1 - f_2)(x)$, ladite fonction g est une solution de l'équation suivante :

$$g'(x) + p(x)g'(x) = 0$$

respectant, de plus $g(x_0) = 0$.

Dans l'équation ci-dessus, le fait que $q(x) = 0 \forall x$ implique que $(ag)'(x) = 0 \forall x$ également, donc $(ag)(x) = C \forall x$. Autrement dit, $(ag)(x)$ est une fonction constante telle que $(ag)(x) = C \forall x$. Comme on sait que $g(x_0) = 0$, on sait que $e^{P(x_0)}g(x_0) = 0$ (où $P(x)$ est une primitive de $p(x)$), donc $C = 0$. Cependant, comme $(e^P)(x)$ ne peut s'annuler, il faut $g(x) = 0 \forall x$.

Ce qui veut donc dire que f_1 et f_2, les deux fonctions solutions trouvées pour une équadiff linéaire d'ordre 1, sont les mêmes (vu que leur différence est la fonction nulle). Il existe donc **une et une seule** solution à l'équadiff linéaire d'ordre 1 conditionnée (problème de Cauchy).

1.7.3 équadiffs linéaire d'ordre 2 à coefficients constants

Une équadiff d'ordre 2 est sous la forme suivante :

$$\frac{d^2}{dx^2}f(x) + a\frac{d}{dx}f(x) + bf(x) = 0$$

avec deux paramètres réels $a, b \in \mathbb{R}$. Le terme *linéaire* vient du fait que si $f_1, f_2 \in \mathbb{R}^{\mathbb{R}}$ sont deux solutions de cette équation, alors $c_1f_1 + c_2f_2$ en est également une (avec $C_1, C_2 \in \mathbb{R}$). Le fait que cette équation soit d'ordre 2 veut également dire que l'ensemble des solutions est un espace vectoriel de dimension 2.

Pour résoudre une telle équation, il faut d'abord passer par ce que l'on appelle l'*équation caractéristique*. Cette équation est la suivante :

$$\lambda^2 + a\lambda + c = 0.$$

Cette équation étant du second degré, il faut séparer trois cas possibles :

- l'équation admet deux solutions réelles distinctes λ_1 et λ_2 ;
- l'équation admet une seule solution réelle λ ;
- l'équation n'admet aucune solution réelle.

Dans le premier cas (deux solutions distinctes), l'équation différentielle peut être réécrite sous la forme factorisée suivante :

$$\left(\tfrac{d}{dx} - \lambda_1\right)\left(\tfrac{d}{dx} - \lambda_2\right)f = 0.$$

Les solutions $f_1(x) = e^{\lambda_1 x}$ et $f_2(x) = e^{\lambda_2 x}$ sont possible à *deviner*. Le principe de linéarité exprimé juste au-dessus permet d'affirmer donc que $f(x) = C_1e^{\lambda_1 x} + C_2e^{\lambda_2 x}$ représente la famille des solutions paramétrées par C_1 et C_2.

Dans le second cas (solution unique), l'équadiff peut être réécrite sous la forme factorisée suivante :

$$\left(\tfrac{d}{dx} - \lambda\right)^2 f = \frac{d^2}{dx^2}f + \lambda^2 f - 2\lambda\frac{d}{dx}f = 0.$$

En posant $g(x) = \frac{d}{dx}f(x) - \lambda f(x)$, nous avons $\frac{d}{dx}g(x) = \frac{d^2}{dx^2}f(x) - \lambda\frac{d}{dx}f(x)$. Cela nous permet de réécrire (encore une fois) l'équadiff sous la forme suivante :

$$\frac{d^2}{dx^2}f(x) - \lambda\frac{d}{dx}f(x) - \lambda\left(\lambda\tfrac{d}{dx}f(x) - \lambda f(x)\right) = \frac{d}{dx}g(x) - \lambda(g(x)) = 0.$$

La solution est $g(x) = Ke^{\lambda x}$. Or $g(x) = \frac{d}{dx}f(x) - \lambda f(x)$. Donc il faut encore résoudre l'équation à l'aide de la méthode vue ci-dessus (ordre 1) afin de trouver la solution suivante. La solution finale est :

$$f(x) = \frac{\int\left(e^{\int p(x)\,dx}\right)q(x)\,dx}{e^{\int p(x)\,dx}}$$

avec $p(x) = -\lambda$ et $q(x) = g(x) = K_1e^{\lambda x}$. D'où :

$$f(x) = \frac{\int\left(e^{\int -\lambda\,dx}\right)K_1e^{\lambda x}\,dx}{e^{\int -\lambda x\,dx}} = \frac{\int K_2e^{-\lambda x}K_1e^{\lambda x}\,dx}{K_3e^{-\lambda x}} = \left(K\int dx\right)e^{\lambda x}K_3^{-1} = K_3^{-1}K(x+C)e^{\lambda x} = (C_1x + C_2)e^{\lambda x}.$$

Dans le dernier cas, l'équation caractéristique n'a pas de solution réelle. Il faut donc aller chercher du côté des nombres complexes. Les coefficients étant réels, les deux solutions complexes doivent être conjuguées l'une de l'autre (si $z = \alpha + \beta i$ et $\bar{z} = \alpha - \beta i$, alors $z\bar{z} = \alpha^2 + \beta^2 \in \mathbb{R}$ et $z + \bar{z} = 2\beta \in \mathbb{R}$). L'équadiff devient donc :

$$\left(\frac{d}{dx} - z\right)\left(\frac{d}{dx} - \bar{z}\right) f = 0.$$

L'équation ressemble fortement à celle du premier cas, donc la solution générale est $f(x) = b_1 e^{zx} + b_2 e^{\bar{z}x}$ avec $b_1, b_2 \in \mathbb{C}$. Cependant, comme $\exp(zx) = \exp(\alpha x + i\beta x) = e^{\alpha x}\exp(i\beta x) = e^{\alpha x}(\cos(\beta x) + i\sin(\beta x))$ (et donc $\exp(\bar{z}) = e^{\alpha x}(\cos(\beta x) - i\sin(\beta x))$), en choisissant respectivement $b_1 = b_2 = \frac{1}{2}$ et $b_1 = -b_2 = -\frac{i}{2}$, on obtient respectivement $f(x) = e^{\alpha x}\cos(\beta x)$ et $f(x) = e^{\alpha x}\sin(\beta x)$. Nous avons donc deux solutions, et à nouveau, par linéarité, nous pouvons exprimer la famille des solutions paramétrée par $C_1, C_2 \in \mathbb{R}$:

$$f(x) = C_1 e^{\alpha x}\cos(\beta x) + C_2 e^{\alpha x}\sin(\beta x).$$

Ici, notre solution ne fait plus intervenir quoi que ce soit de complexe ($\in \mathbb{C}$), tous les coefficients sont réels (α et β sont des *constantes* dépendantes de l'équation caractéristique).

1.7.4 équations de Newton

Après avoir étudié des équations linéaires, regardons une autre famille d'équations : les **équations de Newton**. Cette famille est représentée par la forme suivante :

$$\frac{d^2}{dt^2}x(t) = f(x(t))$$

avec $x : \mathbb{R} \to \mathbb{R} : t \mapsto x(t)$, l'inconnue et f, la fonction **continue** de *force*. Une telle équation représente la position $x(t)$ en fonction du temps d'un mobile (de masse unitaire) soumis à une force f dépendant de la position. Si on multiplie l'équation de part et d'autre par la quantité $x'(t)$, on obtient

$$x^{(2)}(x)x'(t) - f(x(t))x'(t) = 0$$

Ce que l'on peut intégrer afin d'avoir

$$\frac{1}{2}(x'(t))^2 - F(x(t)) - K = 0,$$

avec F, une primitive de f, ou encore

$$(x'(t))^2 - 2F(x(t)) = E$$

avec E, la constante d'intégration. Si l'équation différentielle est conditionnée (problème de Cauchy) telle que $x(0) = x_0$ et $x'(0) = v_0$, alors une solution pour cette équation est $x'(t) = \sqrt{E_0 + 2F(x(t))}$ avec $E_0 = v_0^2 - 2F(x_0)$.

Exemple : les équations de Fisher Les équations de Fisher sont sous la forme suivante : $u''(t) = (u - u^3)(t)$. La fonction f de force est ici $f : \mathbb{R} \to \mathbb{R} : u(t) \mapsto u(t) - u^3(t)$, et a pour primitive $F : \mathbb{R} \to \mathbb{R} : \frac{1}{2}u^2(t) - \frac{1}{4}u^4(t) + C$. Pour des raisons de simplicité, ici, C se verra attribuer la valeur $-\frac{1}{4}$. Donc $F(u(t)) = -\frac{1}{4}(u^2(t) - 1)^2$.

Étant donné les solutions constantes à l'équation $u : t \mapsto K_u$ avec $K_u \in \{-1, 0, 1\}$, nous cherchons u telle que :

$$\lim_{t \to \pm\infty} u(t) = \pm 1,$$

et :

$$\lim_{t\to\pm\infty} u'(t) = 0.$$

L'intégration première nous donne $u'(t)^2 = \frac{1}{2}(u^2(t)-1)^2$, ou encore :

$$u'(t) = \frac{\sqrt{2}}{2}(u^2(t)-1).$$

Dans les intervalles de temps tels que $u'(t) > 0$ et $-1 \neq u(t) \neq 1$, nous avons :

$$\int_{t_0}^{t} \frac{1}{u^2-1}\,du = \int_{t_0}^{t} \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{2}(t-t_0)$$
$$\operatorname{arctanh}(u(t)) - \operatorname{arctanh}(u(t_0)) = \frac{\sqrt{2}}{2}(t-t_0)$$

Cependant, nous savons qu'$\exists t_0$ tel que $u(t_0) = 0$ vu que nous cherchons $-1 < u(t) < 1$. Supposons $t_0 = 0$, de manière à ce que l'équation devienne $\operatorname{arctanh}(u(t)) = \frac{\sqrt{2}}{2}t$ d'où $u(t) = \tanh(\frac{t}{\sqrt{2}})$.

1.8 Problèmes et paradoxes

La majeure partie de l'analyse mathématique est basée sur la notion de limite, et surtout de limite infinie. Si cette notion est utilisée sans suffisamment de rigueur, un certain nombre de paradoxes peuvent apparaitre. Par exemple, soit la série $x, x^2, x^3, ...$ telle que $x_i = x^i$. Pour en connaitre la limite infinie, nous pouvons faire $L = \lim_{n\to\infty} x^n = \lim_{m\to\infty} x^{m+1} = x \lim_{m\to\infty} x^m = xL$. Nous avons donc $L = xL$, ou encore, $\forall x \neq 1, L = 0$. Ce qui peut sembler contre-intuitif pour $x = 2$ ou $x = 3$ par exemple. Pareillement, si l'on désire mesurer l'hypoténuse d'un triangle ABC avec $A = (0,0), B = (1,0), C = (0,1)$, on peut approximer la longueur de l'hypoténuse comme la longueur d'un « escalier » de N marches. Chaque marche est composée de deux segments de longueur $\frac{1}{N}$. La longueur de l'escalier est donc $\frac{2N}{N}$, ou encore 2. Alors que le théorème de Pythagore nous dit que l'hypoténuse est de longueur $\sqrt{2}$.

2 Les nombres réels

Le traitement de l'analyse peut être très puissant, cependant comme vu ci-dessus, un manque de rigueur peut amener à des contradictions voire à des résultats insensés. C'est pour cette raison qu'il faut instaurer cette rigueur dès les notions élémentaires telles que les nombres.

2.1 Axiomatique des nombres

Rappel Les nombres sont organisés de la sorte : $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Il faut savoir que chaque ensemble est défini selon des axiomes, et que c'est à partir de ces axiomes et de déductions logiques que sont construits les ensembles *plus gros*. Ici, les axiomes correspondant aux ensembles $\mathbb{N}, \mathbb{Z}$ et $\mathbb{Q}$ sont considérés comme *évidents* et seuls ceux de $\mathbb{R}$ sont explicités.

2.1.1 Axiomes de $\mathbb{R}$

Avant de citer les axiomes de $\mathbb{R}$, il faut introduire la notion de majorant et de minorant (pour l'axiome de complétude).

Soit $A \subset \mathbb{R}$. On dit que A est majoré $\iff \exists M \in \mathbb{R} | \forall a \in A, a \leqslant M$.

De manière similaire, on dit que A est minoré $\iff \exists m \in \mathbb{R} | \forall a \in A, a \geqslant m$.
Remarque. Le minorant/majorant ne doit pas nécessairement appartenir à A.

1. $\mathbb{R}$ est un *corps*
   - $(\mathbb{R}, +)$ est un *groupe commutatif*, donc la loi d'addition satisfait les conditions suivantes : associativité, commutativité, existence d'un élément neutre et existence d'un inverse ;
   - $(\mathbb{R} \setminus \{0\}, .)$ est un *groupe commutatif* ;
   - La multiplication est distributive sur l'addition.
2. $(\mathbb{R}, \leqslant)$ est un corps entièrement ordonné
   - la relation d'ordre $\leqslant$ satisfait les propriétés suivantes : réflexivité, transitivité, antisymétrie et ordre total ;
   - $a \leqslant b \Rightarrow a + z \leqslant b + z \forall z \in \mathbb{R}$;
   - $a, b \leqslant 0 \Rightarrow ab \geqslant 0$.
3. $\mathbb{R}$ satisfait l'axiome de complétude. C'est cet axiome qui différencie grandement $\mathbb{R}$ de $\mathbb{Q}$. Cet axiome de complétude dit ceci :
   - $\forall A \subset \mathbb{R}$, si A est non-vide et majoré, alors A possède un majorant minimum appelé *supremum* de A et noté $\sup A$.
   - $\forall A \subset \mathbb{R}$, si A est non-vide et minoré, alors A possède un minorant maximum appelé *infimum* de A et noté $\inf A$.

   Et l'ensemble des rationnels $\mathbb{Q}$ ne respecte pas cet axiome.

2.1.2 Résultat de l'axiome : les racines

Commençons par énoncer la propriété d'Archimède qui dit que $\forall y \in \mathbb{R}, \exists n \in \mathbb{N} \,|\, n > y$. La preuve de ce principe fait intervenir la notion de majorant vue ci-dessus.

Soit $S = \{n \in \mathbb{N} \,|\, n \leqslant y\}$ avec $y \in \mathbb{R}$. Par définition, y majore S. Il faut différencier les cas où $\#S = 0$ et $\#S > 0$. Dans le premier cas, $0 > y$, donc $n = 0$ est le nombre naturel recherché. Dans le second, posons $s = \sup S$. Comme $s \in S$, $s - 1$ n'est pas un majorant de S, $\Rightarrow \exists m \in S \,|\, s - 1 < m$, ou encore $m + 1 > s$. Or s majore S donc $m + 1 \notin S$. Si $m + 1 \notin S$, alors $m + 1 > y$. Donc $n = m + 1$ est le nombre naturel recherché.

S'en suit le corollaire suivant : $\forall y \in \mathbb{R}^+, \exists n \in \mathbb{N} \,|\, y > \frac{1}{n}$. La preuve repose sur le lemme précédent : soient $y \in \mathbb{R}^+, x = \frac{1}{y}$. Le lemme d'Archimède dit qu'$\exists n \in \mathbb{N} \,|\, x < n$. Donc $\frac{1}{y} < n$, ou encore $\frac{1}{n} < y$. Il faut effectivement $y > 0$ car sinon lors de la dernière étape, nous avons $\frac{1}{n} > y$, ce qui n'est pas possible pour $n \in \mathbb{N}$.

Maintenant, intéressons-nous aux racines carrées et à l'affirmation de leur existence. Tentons de démontrer qu' $\exists x \in \mathbb{R} \,|\, x^2 = N \forall N \in \mathbb{R}$. Afin de démontrer ceci, procédons par l'absurde. Soient $S = \{y \in \mathbb{R} \,|\, y^2 \leqslant N\}$ et $x = \sup S$. Prouvons maintenant que $x^2 = N$.

Si $x^2 < N$, posons $x_n = x + \frac{1}{n}$ avec $n \in \mathbb{N}_0$. Donc $x_n^2 = x^2 + \frac{2x}{n} + \frac{1}{n^2} \leqslant x^2 + \frac{2x+1}{n} \, \forall n \in \mathbb{N}_0$. Comme $N - x^2 > 0$ et $2x + 1 > 0$ (vu que $x \geqslant 0$ du fait que $0 \in S$), la quantité $\frac{N-x^2}{2x+1}$ est strictement positive également, donc $\exists n \in \mathbb{N} \,|\, \frac{1}{n} \leqslant \frac{N-x^2}{2x+1}$. Pour ce même n, nous avons $x_n^2 < N$ car $x_n^2 < x^2 + (2x+1)\frac{1}{n} < x^2 + (2x+1)\frac{N-x^2}{2x+1} = x^2 + N - x^2 = N$. Donc $x_n^2 \in S$, cependant $x_n = x + \frac{1}{n} > x$, ce qui n'est pas possible.

Inversement, si $x^2 > N$, on définit $x_n = x - \frac{1}{n}$. D'où $x_n^2 > x^2 - \frac{2x+1}{x^2-N}$. De plus, les quantités $x^2 - N$ et $2x + 1$ sont toutes deux strictement positives, donc $\exists n \in \mathbb{N} \,|\, \frac{1}{n} < \frac{x^2-N}{2x+1}$. Ce qui mène à $x_n^2 > N$. Or $x_n = x - \frac{1}{n} < x$. Donc x_n ne peut être un majorant de S. Cela implique qu'$\exists y \in S \,|\, y \geqslant x_n$, ou encore $y^2 \geqslant x_n^2 > N$, ce qui n'est pas possible car $y \in S \Rightarrow y^2 \leqslant N$.

Donc si $x^2 \not< N$ et $x^2 \not> N$, alors $x^2 = N$. De plus, nous avons une définition d'une racine carrée (que l'on peut étendre à la racine n^e) :

$$x^{\frac{1}{n}} = \sqrt[n]{x} = \sup\{y \in \mathbb{R}^+ \,|\, y^n \leqslant x\}.$$

Lemme 2.1. *Il s'en déduit que* $\forall q = \frac{m}{n} \in \mathbb{Q}, x^q = (x^{\frac{1}{n}})^m$.

Remarque. Ici, les puissances irrationnelles ne sont pas définies. Une telle définition viendra par la suite.

2.2 Densité des rationnels

Nous avons vu l'ensemble $\mathbb{R}$ et sa partie rationnelle ($\mathbb{Q}$). Cependant, quelle est la proportion de ces nombres rationnels et donc quelle est la proportion de nombres irrationnels étant donné le résultat suivant : $\forall x < y \in \mathbb{R}, \exists q \in \mathbb{Q} \,|\, x < q < y$?

Commençons par prouver ce résultat. Séparons le problème en deux: le cas où $y - x > 1$ et le cas où $x < y$ quelconque. Dans le premier cas, $\exists n \in \mathbb{N} \,|\, n > x$ où n est le plus petit entier plus grand que x. On en déduit que $n - 1 \leqslant x$, ou encore $n \leqslant x + 1 < y$, ou encore $x < n < y$, prenons $q = n$. Dans le second cas, Archimède dit qu'$\exists m \in \mathbb{N} \,|\, m > \frac{1}{y-x}$. Cette inégalité peut se réécrire $my - mx > 1$, ce qui correspond au cas précédent. Prenons $q = \frac{n}{m}$, et nous avons $x < q < y$.

Cela permet de se dire qu'il n'y a pas trop de points qui ont été ajoutés pour passer de $\mathbb{Q}$ à $\mathbb{R}$. Cependant, il faut savoir que $\#\mathbb{R} > \#\mathbb{Q}$, bien que ces deux ensembles soient infinis.

2.3 Inégalité triangulaire

L'inégalité triangulaire dit que $\forall x, y \in \mathbb{R}, |x+y| \leqslant |x|+|y|$. Pour le prouver, il faut savoir que $ab \geqslant 0 \Rightarrow |a+b| = |a|+|b|$. Cependant, quand $a \leqslant 0 \leqslant b$, alors $a-|b| = -|a|-|b| \leqslant a+b \leqslant |a|+b = |a|+|b|$. Or, pour avoir $x \leqslant -y$ et $x \geqslant y$, il faut $|x| \leqslant |y|$. Donc $|a+b| \leqslant ||a|+|b|| = |a|+|b|$.

De manière similaire, on peut dire que $\forall x, y \in \mathbb{R}, |x-y| \geqslant |x|-|y|$. Il suffit pour le prouver de poser $x = a-b$ et $y = b$. De là, le lemme précédent s'applique de manière à ce que $|a| \leqslant |a-b|+|b|$, ce qui peut se réécrire comme suit : $|a-b| \geqslant |a|-|b|$.

2.4 Autres corps

Avant tout, il faut définir ce qu'est un corps.

L'ensemble $\mathbb{K}$ est un corps si, muni des opérations d'addition et de produit, il respecte les trois propriétés suivantes :

- $(\mathbb{K}, +)$ est un groupe commutatif ;
- $(\mathbb{K}\backslash\{0\}, .)$ est un groupe commutatif ;
- le produit est distributif sur l'addition.

C'est principalement le corps des nombres réels qui sera étudié dans ce cours, cependant il en existe plein d'autres.

3 Les suites

Après avoir utilisé la notion de « limite » pour *définir* les notions de dérivée et d'intégrale, il est nécessaire de définir précisément cette notion de limite (ou de *convergence*).

Une suite réelle (dans $\mathbb{R}$) est une liste infinie de $x_i \in \mathbb{R}$ indexés par $i \in \mathbb{N}$. Cette suite se note (x_n), $(x_n)_n$, ou encore $(x_n)_{n\in\mathbb{N}}$.

3.1 Convergence et divergence

On dit d'une suite (x_n) qu'elle converge en $a \in \mathbb{R} \iff \forall\epsilon > 0, \exists N \in \mathbb{N} \mid n \geqslant N \Rightarrow |x_n - a| < \epsilon$. Cela se note :

$$\lim_{n\to\infty} x_n = a.$$

Il reste cependant à prouver qu'une suite convergente n'a qu'une seule limite. Pour ce faire, procédons par l'absurde. Supposons que $x_n \to a$ et $x_n \to b$ quand $n \to \infty$ et supposons que $a \neq b$. Selon la définition ci-dessus, $\exists N_1 \in \mathbb{N} \mid n \geqslant N_1 \Rightarrow |x_n - a| < \epsilon$ et $\exists N_2 \in \mathbb{N} \mid n \geqslant N_2 \Rightarrow |x_n - b| < \epsilon$. En prenant $\epsilon = \frac{1}{2}|a-b| > 0$ car $a \neq b$ et $N = \max\{N_1, N_2\}$, l'inégalité triangulaire dit que $|a - x_N + x_N - b| \leqslant |a - x_N| + |x_N - b| < 2\epsilon$. Donc $|a - b| < |a - b|$. L'hypothèse disant $\epsilon > 0$ est fausse, ce qui implique $\epsilon = 0$, ou encore $a = b$.

Pour définir une convergence en l'infini positif, on procède de la sorte : $\lim_{n\to\infty} x_n = \infty \iff \forall K \in \mathbb{R}^+, \exists N \in \mathbb{N} \mid n \geqslant N \Rightarrow x_n > K$. De manière similaire, pour l'infini négatif : $\lim_{n\to\infty} x_n = -\infty \iff \forall K \in \mathbb{R}^+, \exists N \in \mathbb{N} \mid n \geqslant N \Rightarrow x_n < K$.

On dit d'une suite qui ne converge en aucun réel qu'elle est divergente ($\infty \notin \mathbb{R}$!).

3.1.1 Techniques de démonstration de divergence ou de convergence

Soit (x_n), une suite convergente. Alors (x_n) est bornée. Autrement dit, $\exists K \in \mathbb{R}^+ \mid \forall n \in \mathbb{N}, |x_n| \leqslant K$.

Pour prouver ceci, il faut d'abord montrer que pour $a = \lim x_n$, $\forall\epsilon > 0, \exists N \in \mathbb{N} \mid n \geqslant N \Rightarrow |x_n - a| < \epsilon$. Donc $|x_n| = |x_n - a + a| \leqslant |x_n - a| + |a| < \epsilon + |a|$. Autrement dit, à partir de $n = N$, (x_n) est borné par $\epsilon + |a|$. Il reste donc un nombre fini d'éléments à traiter. Donc $K = \max\{|x_0|, ..., |x_{N-1}|, \epsilon + |a|\}$ borne l'entièreté de (x_n).

De plus, les opérations sur suites sont définies telles que, pour (x_n) et (y_n) convergentes respectivement en a et b :

- $(x_n + y_n)$ est convergente en $a + b$;
- $(x_n y_n)$ est convergente en ab ;
- $b \neq 0 \Rightarrow (\exists M \mid n \geqslant M \Rightarrow y_n \neq 0) \wedge (x_n/y_n)_{n\geqslant M}$ est convergente en $\frac{a}{b}$.

Ces assertions se démontrent comme suit.

Somme de suites Soit $\epsilon > 0$, $\exists N_1 \mid \forall n \geqslant N_1, |x_n - a| < \frac{\epsilon}{2}$. Pareillement pour $N_2 \mid \forall n \geqslant N_2, |y_n - b| < \frac{\epsilon}{2}$. En prenant $N = \max\{N_1, N_2\}$, on a $|(x_n + y_n) - (a + b)| \leqslant |x_n - a| + |y_n - b| < \epsilon$.

Produit de suites La suite $(x_n y_n)$ est bornée, donc $\exists K \mid |x_n| \leqslant K$. Donc $|x_n y_n - ab| = |x_n(y_n - b) - b(x_n - a)|$. Autrement dit, $|x_n y_n - ab| \leqslant |x_n||y_n - b| + |b||x_n - a| = K|y_n - b| + |b||y_n - b|$. Avec $\epsilon > 0$, $\exists N_1 \mid \forall n \geqslant N_1, |y_n - b| < \frac{\epsilon}{2K}$ et $\exists N_2 \mid \forall n \geqslant N_2, |x_n - a| l \frac{\epsilon}{|b|+1}$ (il faut mettre $|b| + 1$ dans le cas où $b = 0$. Avec $N = \max\{N_1, N_2\}$, $\forall n \geqslant N$, on a :

$$|x_n y_n - ab| \leqslant K|y_n - b| + |b||x_n - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Quotient de suites En prouvant que $(\frac{1}{y_n}) \to \frac{1}{b}$, le quotient découle du produit. Soit $\epsilon = \frac{|b|}{2} > 0, \exists M \mid n \geqslant M \Rightarrow |y_n - b| < \epsilon$. Donc, par l'inégalité triangulaire, $|y_n| \geqslant |b| - |y_n - b| > |b| - \epsilon = \frac{|b|}{2}$. La suite $(y_n)_{n \geqslant M}$ est bien définie.Maintenant prouvons sa convergence en $\frac{1}{b}$:

$$\left|\frac{1}{y_n} - \frac{1}{b}\right| = \left|\frac{b - y_n}{by_n}\right| \leqslant \frac{2}{|b|^2}|b - y_n|.$$

Pour $\epsilon > 0$ fixé, $\exists N \geqslant M \mid \forall n \geqslant N, |y_n - b| < \frac{|b|^2}{2}\epsilon$. Donc $|\frac{1}{y_n} - \frac{1}{b}| < \frac{2}{|b|^2}\frac{|b|^2}{2}\epsilon = \epsilon$.

Exemple du quotient de polynômes de degré k Soient $a_i, b_i \in \mathbb{R}$ avec $0 \leqslant i \leqslant k$. La suite (x_n) définie par $x_n = \frac{\sum_{i=0}^k a_i n^i}{\sum_{i=0}^k b_i n^i}$ converge en $\frac{a_k}{b_k}$. Pour le prouver, il faut diviser le numérateur et le dénominateur par n^k. La suite devient $x_n = \frac{\sum_{i=0}^k a_i n^{i-k}}{\sum_{i=0}^k b_i n^{i-k}}$. Or $\frac{1}{n^K}$ converge en 0 $\forall K > 0$. Donc $\lim_{n\to\infty} x_n = \lim_{n\to\infty} \frac{\sum_{i=0}^k a_i n^{i-k}}{\sum_{i=0}^k b_i n^{i-k}} = \frac{a_k}{b_k}$.

Théorème du sandwich Soient (a_n) et (b_n), deux suites réelles convergentes vers l. Si (x_n) est une suite satisfaisant $a_n \leqslant x_n \leqslant b_n \forall n \geqslant N_0$, alors $x_n \to l$.

Soit $\epsilon > 0$. Par définition, on sait :

$$\exists N_1, N_2 \mid (n \geqslant N_1 \Rightarrow |a_n - l| < \epsilon) \wedge (n \geqslant N_2 \Rightarrow |b_n - l| < \epsilon).$$

Donc $\forall n \geqslant N = \max\{N_0, N_1, N_2\}, -\epsilon < a_n - l \leqslant x_n - l \leqslant b_n - l < \epsilon$. Autrement dit, $|x_n - l| < \epsilon$.

Il en découle un corollaire disant que si (y_n) est une suite bornée (pas forcément convergente) et $z_n \to 0$, alors $y_n z_n \to 0$. Pour le prouver, on sait qu'$\exists K > 0 \mid |y_n| < K$. Autrement dit, $-K < |y_n| < K$, d'où $-K|z_n| \leqslant |y_n||z_n| \leqslant K|z_n|$. Puisque $|y_n||z_n| = |y_n z_n|$ et $z_n \to 0$, alors $|y_n z_n| \to 0$.

Règle de l'exponentielle Cette règle dit que si (a_n) est une suite réelle positive convergent en 0, alors (a_n^p) avec $p \in \mathbb{R}$ converge également en 0. Pour le prouver, il faut prendre $\epsilon > 0$ et $\epsilon' = \epsilon^{p^{-1}}$. Par définition, $\exists N \mid n \geqslant N \Rightarrow |a_n| < \epsilon'$. Cependant, $a_n \geqslant 0$, donc $0 \leqslant a_n < \epsilon' = \epsilon^{p^{-1}}$. D'où $0 \leqslant a_n^p < \epsilon$.

Suites convergentes en 0 Les suites suivantes convergent en 0, la plupart se démontrent avec le binôme de Newton et/ou le théorème du sandwich.

- $(\frac{1}{n^p})$ avec $p > 0$;
- (c_n) avec $|c| < 1$;
- $(n^p c^n)$ avec $p \in \mathbb{R}, |c| < 1$;
- $(\frac{c^n}{n!})$ avec $p \in \mathbb{R}$;
- $\frac{n^p}{n!}$ avec $p \in \mathbb{R}$.

La première proposition se démontre avec la règle de l'exponentielle vu que $(n^{-1}) \to 0$. La seconde se démontre avec le théorème du sandwich en posant $c = \frac{1}{1+a}$ avec $a > 0$. La troisième se démontre de manière similaire. La quatrième se démontre avec le principe d'Archimède et le théorème du sandwich. La dernière se réécrit $\frac{n^p}{n!} = \frac{n^p}{2^n}\frac{2^n}{n!}$ et est donc un produit de deux suites convergentes en 0.

Le procédé pour déterminer la convergence d'une suite est donc de trouver un maximum de suites convergentes en 0 puis d'appliquer les opérations sur les limites de suite. Regardons maintenant du côté des suites divergentes.

Soit $x_n \to \infty$ et $(y_n) \subset \mathbb{R}$. S'il $\exists A \in \mathbb{R} \mid \forall n, y_n > A$, alors $x_n + y_n \to \infty$. Également, si $A > 0$, alors $x_n y_n \to \infty$. La première affirmation se démontre comme suit : soit $K > 0, \exists N \mid \forall n \geqslant N, x_n > K - A$, donc $x_n + y_n > K$. Pour le produit, $\exists N \mid \forall n \geqslant N, x_n > \frac{K}{A}$ donc $x_n y_n > K$.

Règle de la réciproque Soit (x_n) une suite réelle qui tend vers $\pm\infty$. Alors $\frac{1}{x_n} \to 0$. De plus, soit (x_n), une suite réelle non-nulle. S'$\exists M \,|\, n \geqslant M \Rightarrow x_n > 0$, alors $x_n \to \infty$. Similairement, s'$\exists M \,|\, n \geqslant M \Rightarrow x_n < 0$, alors $x_n \to -\infty$. Pour démontrer la première affirmation, il faut avoir $\epsilon > 0$, donc $\frac{1}{\epsilon} > 0$. On sait qu'$\exists N \,|\, n \geqslant N \Rightarrow x_n > \frac{1}{\epsilon}$. ou encore $0 < \frac{1}{x_n} < \epsilon$ (car $x_n > \frac{1}{\epsilon} > 0$). Pour démontrer la seconde, il faut avoir $K > 0$. On sait qu'$\exists N \,|\, n \geqslant N \Rightarrow x_n < \frac{1}{K}$. Ou encore $\frac{1}{x_n} > K$.

Soient (x_n), une suite réelle et une suite strictement croissante $n_1 < n_2 < \dots$. La suite (x_{n_k}) est une *sous-suite* de (x_n).

Il en découle le lemme suivant : Soit (x_n), une suite réelle convergente en a. Alors toute sous-suite de (x_n) converge également en a. Pour le démontrer, il faut repartir de la définition et donc, puisque $x_n \to a$, alors $\exists n \,|\, n \geqslant N \Rightarrow |x_n - a| < \epsilon$ avec $\epsilon \in \mathbb{R}_0^+$. Or, comme (n_k) est une série naturelle, $n_k \geqslant k \; \forall k$ (preuve par récurrence). Si $k \geqslant N$, alors $|x_{n_k} - a| < \epsilon$. Donc $x_{n_k} \to a$.

On peut en déduire le corollaire suivant : si (x_n) a deux sous-suites convergentes mais ayant des limites différentes, alors (x_n) ne converge pas.

3.1.2 Les suites monotones

Une suite est dite *croissante* si $x_n \leqslant x_{n+1} \; \forall n$ et est dite *décroissante* si $x_n \geqslant x_{n+1} \; \forall n$. Si une suite est soit croissante, soit décroissante, elle est dite *monotone*. Sur base de cette définition, il existe un théorème important, celui de la convergence des suites monotones :

Soit (x_n), une suite monotone bornée. (x_n) est convergente telle que quand (x_n) est croissante, $\lim x_n = \sup\{x_n \,|\, n \in \mathbb{N}\}$ et quand (x_n) est décroissante, $\lim x_n = \inf\{x_n \,|\, n \in \mathbb{N}\}$.

Pour le démontrer, il faut définir $S = \{x_n \,|\, n \in \mathbb{N}\}$, borné par hypothèse. De là, il faut supposer (x_n) soit croissante soit décroissante (la démonstration est similaire) et prouver que $a = \sup\{x_n \,|\, x \in \mathbb{N}\}$ si $x_n \to a$. Soit $\epsilon > 0$. Par définition, a est le majorant minimal de S. Donc $a - \epsilon$ ne majore pas S, donc $\exists x_N > a - \epsilon$. Or, comme (x_n) est croissante, $n \geqslant N \Rightarrow x_n \geqslant x_N > a - \epsilon$. Mais $x_n \leqslant a$ car a majore S. Donc $a - \epsilon < x_n \leqslant a$. Ou encore $-\epsilon < x_n - a < 0$, d'où $|x_n - a| < \epsilon$.

On sait donc qu'une suite monotone non bornée diverge en $\pm\infty$ et qu'une suite bornée converge en $\sup\{x_n \,|\, n \in \mathbb{N}\}$ ou en $\inf\{x_n \,|\, n \in \mathbb{N}\}$.

3.2 Théorème de Bolzano-Weierstrass

Regardons comment construire deux suites monotones à partir d'une suite quelconque. Soit une suite (x_n), commençons par définir :

$$S_m := \{x_n | m \leqslant n\}.$$

Si (x_n) est majorée, alors $s_n = \sup S_n$ est fini pour tout n et (s_n) est une suite décroissante car $m \geqslant n \Rightarrow S_m \subseteq S_n \Rightarrow \sup S_m \leqslant \sup S_n$. De même, si (x_n) est minorée, alors $i_n = \inf S_n$ est fini pour tout n et (i_n) est une suite croissante.

Définition 3.1. Soit $(x_n) \subseteq \mathbb{R}$.

- Si (x_n) est majorée, alors (s_n) est bien définie, et on écrit :

$$\limsup_{n \to \infty} x_n := \lim_{n \to \infty} s_n$$

  que l'on appelle la *limite supérieure* de (x_n).

- Si (x_n) n'est pas majorée, alors on écrit $\limsup_{n\to\infty} x_n = \infty$.
- Si (x_n) est minorée, alors (i_n) est bien définie, et on écrit :

$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty} i_n$$

  que l'on appelle la *limite inférieure* de (x_n).
- Si (x_n) n'est pas minorée, alors on écrit $\liminf_{n\to\infty} x_n = -\infty$.
- Constatons donc que pour une suite (x_n) bornée, $\limsup_{n\to\infty} x_n$ et $\liminf_{n\to\infty} x_n$ sont tous les deux finis.

Théorème 3.2 (Bolzano-Weierstrass). *Soit (x_n) une suite bornée. Alors il existe une sous-suite de (x_n) qui converge en $\limsup_{n\to\infty} x_n$ et une autre qui converge en $\liminf_{n\to\infty} x_n$.*

De plus, pour une sous-suite convergente (x_{n_k}) quelconque, $\liminf_{n\to\infty} x_n \leqslant \lim_{k\to\infty} x_{n_k} \leqslant \limsup_{n\to\infty} x_n$.

$$\liminf_{n\to\infty} x_n \leqslant \lim_{k\to\infty} x_{n_k} \leqslant \limsup_{n\to\infty} x_n.$$

Démonstration. Soit (x_n) une suite bornée. Montrons qu'il existe un sous-suite convergent en $\limsup_{n\to\infty} x_n$. La démonstration pour la sous-suite convergent en $\liminf_{n\to\infty} x_n$ étant similaire. Définissons cette suite (x_{n_k}) terme à terme.

Prenons $n_0 = 0$. C'est à dire que la sous-suite (x_{n_k}) débute en x_0. Regardons maintenant l'ensemble $S_{0+1} = \{x_1, x_2, \ldots\}$. Comme s_1 est le supremum de S_1, il doit exister $x_m \in S_1$ tel que $s_1 - 1 < x_m \leqslant s_1$. Prenons alors pour n_1 le plus petit m satisfaisant cette condition. Comme $\forall \alpha \in S_1, \alpha > x_0 = n_0$, il est évident que $n_1 > n_0$.

Regardons maintenant l'ensemble $S_{1+n_1} = \{x_{1+n_1}, x_{2+n_1}, \ldots\}$. De manière similaire au point précédent, il est possible de trouver pour n_2 un m tel que $x_m \in S_{1+n_1}$ et $s_{1+n_1} - \frac{1}{2} < x_m \leqslant s_{1+n_1}$. Il est tout aussi évident que $n_2 > n_1$.

En réitérant ce procédé k fois, on obtient $n_0 < n_1 < n_2 < \ldots < n_k$ satisfaisant : $s_{1+n_i} - \frac{1}{i+1} < x_{n_{i+1}} \leqslant s_{1+n_i}$.

On obtient donc au final une sous-suite (x_{n_k}) de (x_n) telle que $s_{1+n_k} - \frac{1}{k+1} < x_{n_{k+1}} \leqslant s_{1+n_k}$. De plus, la suite (s_{1+k_k}) est une sous-suite de (s_n). On a donc $s_{1+n_k} \to \limsup_{n\to\infty} x_n$ et $s_{1+n_k} - \frac{1}{k+1} \to \limsup_{n\to\infty} x_n$. Par le théorème du sandwich, on peut déterminer que $x_{n_k} \to \limsup_{n\to\infty} x_n$ également.

De plus (en admettant l'existence de (i_{n_k}) car la preuve est similaire à celle de (s_{n_k})), on a :

$$\forall k \in \mathbb{N}, i_{n_k} \leqslant x_{n_k} \leqslant s_{n_k}.$$

Et comme (i_{n_k}) et (s_{n_k}) sont des sous-suites d'une suite convergente, elles sont elles-mêmes convergentes. Donc par un résultat précédent, $\lim_{k\to\infty} i_{n_k} \leqslant \lim_{k\to\infty} x_{n_k} \leqslant \lim_{k\to\infty} s_{n_k}$, ou encore $\liminf_{n\to\infty} x_n \leqslant \lim_{k\to\infty} x_{n_k} \leqslant \limsup_{n\to\infty} x_n$. Ce qui prouve la deuxième partie du théorème. □

Corollaire 3.3. *On peut établir un corollaire de ce théorème : une suite $(x_n) \subseteq \mathbb{R}$ converge si et seulement si :*

$$\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = L \in \mathbb{R},$$

en quel cas, $x_n \to L$.

Démonstration. La démonstration se fait par le théorème du sandwich dans un sens et par le théorème de Bolzano-Weierstrass dans l'autre sens. □

Remarque. Nous avons donc maintenant trois résultats primordiaux sur la convergence des suites :

1. Toute suite convergente est bornée ;
2. toute suite monotone et bornée est convergente ;
3. Toute suite bornée possède une sous-suite convergente.

3.3 Le critère de Cauchy

Jusqu'ici, la notion de convergence fait intervenir *directement* la notion de limite, ce qui sous-entend qu'il faut connaitre la limite avent de commencer la preuve de la convergence. Il existe donc une notion permettant de prouver la convergence sans expliciter la limite.

Définition 3.4. Une suite (x_n) est dite *de Cauchy* si $\forall \epsilon > 0, \exists N > 0 \mid m, n \geqslant N \Rightarrow |x_n - x_m| < \epsilon$.

Cette définition n'implique pas uniquement qu'il faut que $|x_n - x_{n+1}|$ tende vers zéro mais bien qu'il faut que **pour tout** $m, n \geqslant N$, ces deux éléments tendent vers zéro.

Lemme 3.5. *Soit* $(x_n) \subseteq \mathbb{R}$ *une suite convergente.* (x_n) *est de Cauchy.*

Démonstration. Soit une suite (x_n) convergente en $a \in \mathbb{R}$. Par définition, $\forall \epsilon > 0, \exists N > 0 | n \geqslant N \Rightarrow |x_n - a| < \epsilon$. Soit $\epsilon' = \frac{\epsilon}{2}$. Soit ce N découlant de la définition de convergence. Prenons $m, n \geqslant N$. On a donc $|x_m - x_n| = |x_m - a + a - x_n| < |x_m - a| + |x_n - a| = \epsilon'$. □

Théorème 3.6 (Critère de Cauchy)**.** *Une suite* (x_n) *converge si et seulement si elle est de Cauchy.*

Démonstration. L'implication $\Rightarrow$ est donnée par le lemme précédent. Il faut encore prouver l'implication $\Leftarrow$.

Commençons par montrer que (x_n) est bornée et puis appliquons Bolzano-Weierstrass. Par hypothèse, (x_n) est de Cauchy. Donc $\forall \epsilon > 0, \exists N > 0 \mid m, n \geqslant N \Rightarrow |x_n - x_m| < \epsilon$. Soit ϵ, prenons ce N qui découle de la définition de suite de Cauchy. On a donc $\forall n \geqslant N, |x_n| = |x_n + x_N - x_N| \leqslant |x_n - x_N| + |x_N| < \epsilon + |x_N|$. Construisons $K = \max(\{|x_0|, |x_1|, |x_2|, \ldots, \epsilon + |x_N|\})$. On a donc $|x_n| < K \forall n \geqslant N$, ce qui implique que (x_n) est bornée.

Appliquons maintenant le théorème de Bolzano-Weierstrass qui dit qu'il existe une sous-suite $(x_{n_k})_k$ qui converge en une valeur $a \in \mathbb{R}$ quand $k \to \infty$. Par la convergence, il existe N_1 tel que $\forall k \geqslant N_1, |x_{n_k} - a| < \frac{\epsilon}{2}$. Et par la définition de suite de Cauchy, il existe N_2 tel que $m, n \geqslant N_2 \Rightarrow |x_n - x_m| < \frac{\epsilon}{2}$. Soit δ tel que $\delta \geqslant N_1$ et $n_\delta \geqslant N_2$. On a alors $\forall n \geqslant N_2, |x_n - a| = |x_n - x_\delta + x_\delta - a| \leqslant |x_n - x_\delta| + |x_\delta - a| < \epsilon$.

Il y a donc également convergence de la suite (x_n) en $a \in \mathbb{R}$ quand $n \to \infty$. □

4 Fonctions continues

4.1 Limite d'une fonction en un point

Définition 4.1. Soient $a, b \in \mathbb{R}$ tels que $a < b$. On note :

$$\begin{aligned}
[a, b] &:= \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\} \\
]a, b] &:= \{y \in \mathbb{R} \mid a < x \leqslant b\} \\
[a, b[&:= \{x \in \mathbb{R} \mid a \leqslant x < b\} \\
]a, b[&:= \{x \in \mathbb{R} \mid a < x < b\} \\
[a, \infty[&:= \{x \in \mathbb{R} \mid a \leqslant x\} \\
]a, \infty[&:= \{x \in \mathbb{R} \mid a < x\} \\
]-\infty, b] &:= \{x \in \mathbb{R} \mid x \leqslant b\} \\
]-\infty, b[&:= \{x \in \mathbb{R} \mid x < b\} \\
]-\infty, \infty[&:= \mathbb{R}\,.
\end{aligned}$$

Ces ensembles sont appelés *intervalles*.

Définition 4.2. Soit $U \subseteq \mathbb{R}$. U est dit ouvert $\iff \forall x \in U, \exists \epsilon > 0 \,|\,]x - \epsilon, x + \epsilon[\subseteq U$. U est dit fermé $\iff \mathbb{R} \setminus U$ est ouvert.

Définition 4.3. Soient $A \subseteq \mathbb{R}$ et $a \in A$. a est dit intérieur à A s'$\exists \delta > 0 \,|\,]a - \delta, a + \delta[\subseteq A$. L'ensemble des points a tels que a est intérieur à A est noté $\operatorname{int} A$.

Remarque. $\forall A \subseteq \mathbb{R}$, $\operatorname{int} A \subseteq A$. De plus, un ensemble peut être simultanément ouvert et fermé : $]-\infty, \infty[$ est ouvert car $\forall x \in R, \exists \epsilon > 0 \,|\,]x - \epsilon, x + \epsilon[\in \mathbb{R}$ et est fermé car $\mathbb{R} \setminus]-\infty, \infty[=]0, 0[$ est ouvert.

Définition 4.4. Soit $a \in \mathbb{R}$. Un voisinage de a est un ensemble contenant un intervalle de la forme $]c, d[$ avec $c < a < d$.

Définition 4.5. Soient $A \subseteq \mathbb{R}$ et $a \in \mathbb{R}$. a est adhérent à A si $\forall \delta > 0,]a - \delta, a + \delta[\cap A \neq \emptyset$. On note $\operatorname{adh} A$ l'ensemble des points adhérents à A.

Définition 4.6. Soient $f : U \subseteq\to \mathbb{R}$, $B \subseteq \mathbb{R}$ et $a \in \operatorname{adh}(U \cap B)$. La limite de f dans B lorsque x tend vers a existe dans $\mathbb{R}$ et vaut $L \in \mathbb{R}$ si :

$$\forall \epsilon > 0, \exists \delta > 0 \,|\, \forall x \in (U \cap B), |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon.$$

Cela se note $f(x) \to L$ lorsque $x \to a$ dans B ou :

$$\lim_{\substack{x \to a \\ x \in B}} f(x) = L.$$

Remarque. Ici, ϵ permet de déterminer un voisinage autour de L, la limite, alors que δ permet de déterminer un voisinage autour de a.

Théorème 4.7 (Unicité de la limite)**.** *Soient* $f : U \subseteq \mathbb{R} \to \mathbb{R}$, $B \subseteq \mathbb{R}$, $a \in \operatorname{adh}(B \cap U)$. *Soient* $L_1, L_2 \in \mathbb{R}$ *tels que :*

$$\lim_{\substack{x \to a \\ x \in B}} f(x) = L_1$$

$$\lim_{\substack{x \to a \\ x \in B}} f(x) = L_2.$$

Alors $L_1 = L_2$.

Démonstration. Soient $f : U \subseteq \mathbb{R} \to \mathbb{R}$, $B \subseteq \mathbb{R}$ et $a \in \text{adh}(U \cap B)$. Prenons $\epsilon = \frac{|L_1 - L_2|}{3}$. Par hypothèse, on sait qu'il existe $\delta_1 \mid \forall x \in U \cap B, |x - a| < \delta_1 \Rightarrow |f(x) - L_1| < \epsilon$ et $\delta_2 \mid \forall x \in U \cap B, |x - a| < \delta_2 \Rightarrow |f(x) - L_2| < \epsilon$.

Soit $x_0 \in U \cap B$ tel que $|x_0 - a| < \min(\{L_1, L_2\})$. On sait alors que $f(x_0) \in]L_1 - \epsilon, L_1 + \epsilon[$ et $f(x_0) \in]L_2 - \epsilon, L_2 + \epsilon[$. Or, par choix de ϵ, ces deux ensembles sont d'intersection vide. Il y a donc contradiction et $L_1 = L_2$. □

Définition 4.8. La limite de $f : U \subseteq \mathbb{R} \to \mathbb{R}$ en $x \to a$ existe et vaut $L \in \mathbb{R}$ si la limite de f en $x \to a$ dans $B = \mathbb{R}$ existe et vaut L. On note cela :

$$\lim_{x \to a} f(x) = L \in \mathbb{R}.$$

4.1.1 Limites pointées, gauches et droites

Voyons dans quels contextes il est intéressant de manipuler le $B \subseteq \mathbb{R}$.

Définition 4.9. Soit $a \in \mathbb{R}$. Un voisinage de a est pointé s'il contient un intervalle $]c, d[\setminus\{a\}$. Un voisinage de a est *de droite* s'il contient un intervalle $[a, d[$ et peut être pointé si l'intervalle est sous la forme $]a, d[$. De manière similaire, un voisinage est *de gauche* si l'intervalle est sous la forme $]c, a]$ et peut également être pointé.

Définition 4.10 (Définition des limites à gauche, à droite et pointées)**.** Soit $f : U \to \mathbb{R}$ où U est un intervalle contenant un voisinage pointé de a. La limite à gauche, respectivement à droite, respectivement pointée de f en $x \to a$ existe et vaut $L \in \mathbb{R}$ si la limite de f dans $B =]-\infty, a[$, respectivement $]a, \infty[$, respectivement $\mathbb{R} \setminus \{a\}$ en $x \to a$ existe et vaut L.

Cela se note respectivement :

$$\lim_{\substack{x \to a \\ <}} f(x) = L, \quad \lim_{\substack{x \to a \\ >}} f(x) = L, \quad \lim_{\substack{x \to a \\ \neq}} f(x) = L.$$

Remarque. Pour pouvoir parler de limites de f soit à gauche, soit à droite, soit pointée, il faut impérativement que f soit définie dans les alentours de a. Plus précisément, il faut $a \in \text{adh}(A \cap B)$ avec B défini selon le cas.

Lemme 4.11. *Soient $a \in \mathbb{R}$ et $f : U \to \mathbb{R}$ tels que U définit un voisinage pointé de a. Alors f possède une limite pointée en a si et seulement si f possède une limite à gauche en a et une limite à droite en a telles que* $\lim_{x \to a^+} f(x) = \lim_{x \to a^-} f(x) = L \in \mathbb{R}$. *Dans ce cas, on a* $\lim_{\substack{x \to a \\ \neq}} f(x) = L$.

Proposition 4.12. *Soient $f : U \to \mathbb{R}$, $B \subseteq \mathbb{R}$ et $a \in \text{adh}(U \cap B)$. La limite de f dans B en $x \to a$ existe et vaut $L \in \mathbb{R}$ si et seulement si $\forall A \subseteq B$, la limite de f dans A en $x \to a$ existe et vaut L.*

Démonstration. Pour la condition suffisante ($\Leftarrow$), on sait que pour **tout** $A \subseteq B$, la limite existe. En prenant $A = B$, on sait que la limite existe également dans B et vaut $L \in \mathbb{R}$.

Pour la condition nécessaire ($\Rightarrow$), observons que $\forall \epsilon > 0, \exists \delta > 0 \mid x \in (B \cap U) \wedge (|x - a| < \delta) \Rightarrow (|f(x) - L| < \epsilon)$. On sait alors que pour $\epsilon \in \mathbb{R}_0^+$ fixé et pour $A \subseteq B$, il existe $\delta \in \mathbb{R}_0^+$ tel que $x \in A \Rightarrow |f(x) - L| < \epsilon$. □

Proposition 4.13. *Soient $f : U \to \mathbb{R}$, $B \subseteq \mathbb{R}$ et $a \in \text{adh}(U \cap B)$. La limite de f dans B en $x \to a$ existe et vaut $L \in \mathbb{R}$ si et seulement si pour toute suite $(x_n)_{n \in \mathbb{N}} \subseteq B$ telle que $x_n \to a$, la suite $(f(x_n))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ converge en L.*

Démonstration. Soit $\epsilon > 0$. On sait par hypothèse qu'$\exists \delta > 0$ tel que $x \in (B \cap U) \wedge |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon$. Soit $(x_n) \subseteq B$ telle que $x_n \to a$. On sait donc qu'$\exists N > 0$ tel que $n > N \Rightarrow |x_n - a| < \epsilon$. Si ceci est vrai pour tout ϵ, ça l'est plus précisément pour δ déterminé par l'hypothèse. On a donc $\forall \epsilon > 0, \exists N > 0 | n > N \Rightarrow |f(x_n) - L| < \epsilon$ ou encore $f(x_n) \to L$ pour $n \to \infty$.

Montrons maintenant que si toutes les suites de B convergentes en a implique $f(x_n) \to L$, alors la limite de f en a existe et vaut L. Fonctionnons par l'absurde : soit $(x_n) \subseteq B$ une suite dans B telle que $x_n \to a$ quand $n \to \infty$ et $(f(x_n)) \to L \in \mathbb{R}$ mais alors que $f(x) \not\to L$ pour $x \to a$. On sait alors qu'$\exists \epsilon_0 > 0 | \forall \delta > 0$ il n'y a pas de convergence de $f(x)$ en L. Soit ϵ_0. Prenons $\delta = \frac{1}{n}$. On en déduit que $\forall n \geqslant 1$, il existe $x_n \in U$ tel que $|x_n - a| < \frac{1}{n}$ et $|f(x_n) - L| \geqslant \epsilon_0$ donc tel que $x_n \to a$ mais $f(x_n) \not\to L$ lorsque $n \to \infty$. Or par hypothèse, on sait que $(f(x_n)) \to L$. Ce qui est une contradiction avec $|f(x_n) - L| \geqslant \epsilon_0$. Donc $f(x) \to L$ pour $x \to a$. □

4.1.2 Règles de calcul de limite de fonctions

Théorème 4.14. *Soient* $f, g : U \subseteq \mathbb{R} \to \mathbb{R}$ *et* $a \in \mathbb{R}$ *tels que* f *et* g *sont définies dans le voisinage de* a. *Soient* L_f *et* L_g *tels que* $\lim_{x \to a} f(x) = L_f$ *et* $\lim_{x \to a} g(x) = L_g$. *alors :*

$$\lim_{x \to a} (f+g)(x) = L_f + L_g$$
$$\lim_{x \to a} (fg)(x) = L_f L_g.$$

Si $L_g \neq 0$, *la fonction* $\frac{f}{g}(x)$ *existe dans le voisinage de* a *telle que :*

$$\lim_{x \to a} \frac{f}{g}(x) = \frac{L_f}{L_g}.$$

Démonstration. Soient $f, g : U \to \mathbb{R}$ convergents respectivement en L_f et L_g quand $x \to a$.

Addition Montrons que la limite de la somme vaut la somme des limites. Soit $\epsilon > 0$. On sait par la définition des limites de f et g qu'il existe δ_1 tel que $\forall x \in U, |x - a| < \delta_1 \Rightarrow |f(x) - L_f| < \frac{\epsilon}{2}$ et δ_2 tel que $\forall x \in U, |x - a| < \delta_2 \Rightarrow |g(x) - L_g| < \frac{\epsilon}{2}$. Prenons alors $\delta = \min(\{\delta_1, \delta_2\})$. On sait dès lors que $\forall x \in U, |f(x) + g(x) - L_f - L_g| = |f(x) - L_f + g(x) - L_g| \leqslant |f(x) - L_f| + |g(x) - L_g| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$.

Multiplication Montrons que le produit des limite vaut le produit des limites. Montrons tout d'abord qu'il existe un voisinage autour de a tel que $f(x)$ est bornée.

Soit $\epsilon = 1$. On sait alors qu'il existe δ_1 tel que $\forall x \in U, |x - a| < \delta_1 \Rightarrow |f(x) - L_f| < 1$. Ce δ_1 définit un voisinage de a où $|f(x)| < |L_f| + 1 \forall x$.

Montrons ensuite que $\lim_{x \to a} (fg)(x) = L_f L_g$.

Soit $\epsilon \in \mathbb{R}_0^+$. On sait qu'il existe δ_2 tel que $\forall x \in U, |x - a| < \delta_2 \Rightarrow |f(x) - L_f| < \frac{\epsilon}{2(|L_g|+1)}$ et δ_3 tel que $\forall x \in U, |x - a| < \delta_3 \Rightarrow |g(x) - L_g| < \frac{\epsilon}{2|L_f|}$. Prenons alors $\delta = \min(\{\delta_i | i \in [3]\})$. On a alors $\forall x \in U$ tel que $|x - a| < \delta$:

$$\begin{aligned} |f(x)g(x) - L_f L_g| &= |f(x)(g(x) - L_g) + L_g(f(x) - L_f)| \leqslant |f(x)||g(x) - L_g| + |L_g||f(x) - L_f| \\ &< (|L_f| + 1)\frac{\epsilon}{2(|L_f| + 1)} + |L_g|\frac{\epsilon}{2|L_g|} = 2\frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Quotient Montrons que la limite du quotient vaut le quotient des limites. Commençons par montrer qu'il existe un voisinage de a où $g(x) \neq 0$.

Soit $\epsilon = \frac{|L_g|}{2}$. On sait alors qu'il existe δ_V tel que $\forall x \in U, |x - a| < \delta_V \Rightarrow |g(x) - L_g| < \frac{|L_g|}{2}$. Ou encore pour ce même δ_V, $|g(x)| > \frac{|L_g|}{2}$. Donc $|g(x)|$ est strictement positif. Définissons alors $V := U \cap]a - \delta_V, a + \delta_V[$, un voisinage de a sur lequel la fonction $\frac{f}{g}(x)$ est bien définie.

Prouvons maintenant que $\frac{1}{g}(x) \to \frac{1}{L_g}$ quand $x \to a$, et le résultat découlera de la proposition précédente.

Soit $\epsilon > 0$. On sait qu'il existe $\delta > 0$ tel que $\forall x \in U, |x-a| < \delta \Rightarrow |g(x) - L_g| < \frac{|L_g|^2\epsilon}{2}$. On sait dès lors :

$$\left|\frac{1}{g(x)} - \frac{1}{L_g}\right| = \frac{|g(x)-L_g|}{|g(x)||L_g|} < \frac{|g(x)-L_g|}{\frac{|L_g|}{2}|L_g|} = \frac{2|g(x)-L_g|}{|L_g|^2} < \frac{2(\frac{|L_g|^2\epsilon}{2})}{|L_g|^2} = \epsilon.$$

□

Théorème 4.15. *le théorème du sandwich des suites a un homologue pour les fonctions : soient* $f, g, h : \mathbb{R} \to \mathbb{R}$ *trois fonctions telles que* $\forall x \in \mathbb{R}, f(x) \leqslant g(x) \leqslant h(x)$ *avec* $\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L \in \mathbb{R}$. *Alors* $\lim_{x\to a} g(x) = L$.

Démonstration. Définissons les suites $(f(x_n))_n, (g(x_n))_n$ et $(h(x_n))_n$. On sait que $\forall k \in \mathbb{R}^{\mathbb{R}}, \lim_{x\to a} k(x)$ existe et vaut L_k si et seulement si pour toute suite $(x_n) \subseteq \mathbb{R}$ convergente en $a \in \mathbb{R}$, on a $k(x_n) \to L_k$ quand $n \to \infty$.

On sait dès lors que $f(x_n) \to L$ et $h(x_n) \to L$. Par le théorème du sandwich, on sait que $g(x_n) \to L$ également. □

Théorème 4.16 (Conservation des inégalités)**.** *Soient* $f, g : U \subseteq \mathbb{R} \to \mathbb{R}$, $B \subseteq \mathbb{R}$ *et* $a \in \mathrm{adh}(U \cap B)$. *Si* $\exists \delta > 0$ *tel que* $\forall x \in]a-\delta, a+\delta[\cap U \cap B, f(x) \leqslant g(x)$ *et* $\lim_{x\to a} f(x)$ *et* $\lim_{x\to a} g(x)$ *existent dans* $\mathbb{R}$, *alors* $\lim_{x\to a} f(x) \leqslant \lim x \to a g(x)$.

4.1.3 Limites infinies et limites à l'infini

Définition 4.17. Soient $f, g : U \subseteq \mathbb{R} \to \mathbb{R}$, $B \subseteq \mathbb{R}$, $a \in \mathrm{adh}(U \cap B)$. La limite de f dans B en a existe et vaut $+\infty$ si $\forall K > 0, \exists \delta > 0 \,|\, \forall x \in (U \cap B), |x-a| < \delta \Rightarrow f(x) > K$. Cela se note :

$$\lim_{x\to a} f(x) = \infty.$$

De manière similaire, la limite de f dans B en a vaut $-\infty$ si $\forall K > 0, \exists \delta > 0 \,|\, \forall x \in (U \cap B), |x-a| < \delta \Rightarrow f(x) < -K$. Ce la se note :

$$\lim_{x\to a} f(x) = -\infty.$$

Remarque. les cas particuliers où $B = \mathbb{R}, B =]-\infty, a], B = [a, +\infty[, B = \mathbb{R} \setminus \{a\}$ donnent les définitions de limite standard infinie, limite à gauche, limite à droite et limite pointée infinies.

Définition 4.18. Soit $f : U \subseteq \mathbb{R} \to \mathbb{R}$ où U n'est pas majoré. La limite de f en « l'infini » vaut $L \in \mathbb{R}$ si $\forall \epsilon > 0, \exists M > 0 \,|\, \forall x \in U, x > M \Rightarrow |f(x) - L| < \epsilon$. Cela se note :

$$\lim_{x\to\infty} f(x) = L.$$

De manière similaire, si U n'est pas minoré, la limite de f en « moins l'infini » vaut $L \in \mathbb{R}$ si $\forall \epsilon > 0, \exists M > 0 \,|\, \forall x \in U, x < -M \Rightarrow |f(x) - L| < \epsilon$. Cela se note :

$$\lim_{x\to-\infty} f(x) = L.$$

4.2 Définition de la continuité

Définition 4.19. Soient $a \in \mathbb{R}$, $f : I \subseteq \mathbb{R} \to \mathbb{R}$ une fonction définie sur un voisinage de a. La fonction f est continue en a si

$$\lim_{x \to a} f(x) = f(a).$$

Ou encore si

$$\forall \epsilon > 0, \exists \delta > 0 \text{ t. q.} |x - a| < \delta \Rightarrow \left| f(x) - f(a) \right| < \epsilon.$$

Cette définition comporte trois points importants : il faut que la limite de f en a existe, que f soit définie en a et que la fonction prenne la valeur de sa limite en a.

Proposition 4.20. *Venant de la définition de limite de fonction sur base des suites, il est possible d'exprimer la continuité d'une fonction en terme de suites. Soient $a \in \mathbb{R}$, $f : I \subseteq \mathbb{R} \to \mathbb{R}$ une fonction définie sur un voisinage de a. Alors la fonction f est continue en a si et seulement si pour toute suite (x_n) convergente en a, la suite $(f(x_n))$ converge en $f(a)$.*

Proposition 4.21. *Les règles de calcul concernant la continuité sont assez évidents. Soient $a \ni \mathbb{R}$, $f, g : I \subseteq \mathbb{R} \to \mathbb{R}$ deux fonctions définies sur un voisinage de$a$ et continues en a. Alors $f + g$ est également continue en a, fg est également continue en a et si $g(a) \neq 0$, alors $\frac{f}{g}$ est continue en a. De plus soit $h : V \subseteq \mathbb{R} \to \mathbb{R}$ une fonction définie sur un voisinage de $f(a)$ et contenant $f(I)$, l'image de f. Si h est continue en $f(a)$, alors $h \circ f : I \to \mathbb{R}$ est continue en a.*

Démonstration. Les premières propositions viennent directement des règles sur les limites. Pour la composée, prenons (x_n) une suite convergent en a. Par continuité de f, on sait que $f(x_n)$ converge en $f(a)$ et par continuité de h, on sait que $h(f(x_n))$ converge en $h(f(a))$. Dès lors, on sait que $h \circ f$ est continue en a. □

Définition 4.22. Tout comme pour les limites, on parle de continuité à gauche (respectivement à droite) si la limite de x tendant vers a à gauche (respectivement à droite) vaut $f(a)$.

Définition 4.23. Une fonction $f : [a, b] \to \mathbb{R}$ est *continue* si elle est continue en tout point $c \in]a, b[$, continue à droite en a et à gauche en b.

4.3 Théorème des bornes atteintes

Théorème 4.24. *Soit $f : [a, b] \to \mathbb{R}$ une fonction continue. Alors f est bornée **et** f atteint ses bornes. Dès lors :*

$$\exists m, M \in [a, b] \text{ t. q.} \forall x \in [a, b] : f(m) \leqslant f(x) \leqslant f(M).$$

Démonstration. Montrons d'abord que f est bornée et puis montrons qu'elle atteint ses bornes.

Supposons par l'absurde que f n'est pas majorée. Dès lors, $\forall n \in \mathbb{N} : \exists x_n \in [a, b] \text{ t. q.} f(x_n) > n$. La suite (x_n) est bornée car $\forall n : x_n \in [a, b]$. Par le théorème de Bolzano-Weierstrass, on sait qu'il existe $(x_{n_k})_k$ une sous-suite de (x_n) convergeant en $x \in \mathbb{R}$. Par la continuité de f en x (hypothèse), la suite $(f(x_{n_k}))_k$ converge en $f(x)$. Or par construction de (x_n), la suite $(f(x_{n_k}))_k$ diverge vers $+\infty$. Il y a donc contradiction. Idem pour le minorant.

Montrons maintenant que f atteint ses bornes. Par l'absurde, supposons que f n'atteint pas ses bornes. Prenons $M \coloneqq \sup\{f(x) \text{ t. q. } x \in [a, b]\}$ qui existe et est fini. Supposons que $\nexists c \in [a, b] \text{ t. q. } f(c) = M$. Posons alors :

$$g(x) \coloneqq \frac{1}{M - f(x)}.$$

g est continue et donc bornée par $K \in \mathbb{R}$. Dès lors,

$$\forall x \in [a, b] : \frac{1}{M - f(x)} \leqslant K.$$

Ce qui implique (du fait que $f(x) < A$ et $K > 0$) que $M - f(x) \geqslant \frac{1}{K}$ ou encore $f(x) \leqslant M - \frac{1}{K}$. Ce qui contredit le fait que M est le majorant de $\{f(x) \text{ t. q. } x \in [a, b]\}$. Idem pour la borne inférieure. □

4.4 Théorème de la valeur intermédiaire

Théorème 4.25. *Soit* $f : [a, b] \to \mathbb{R}$ *une fonction continue. Soit* γ *strictement entre* $f(a)$ *et* $f(b)$. *Alors* $\exists c \in]a, b[\text{ t. q. } f(x) = \gamma$.

Démonstration. Supposons $f(a) < \gamma < f(b)$, le cas $f(b) < \gamma f(a)$ se montre de la même manière.

Soit $S \coloneqq \{x \in [a, b] \text{ t. q. } f(x) < \gamma\} \subseteq \operatorname{dom} f$. Par définition de S, on sait que $a \in S$ et que S est majoré par b. Il existe une suite (x_n) dans S telle que $x_n \to \sup S$. Donc $\forall n \in \mathbb{N} : x_n \in S$ donc $f(x_n) < \gamma$. De plus, $\lim_{n \to +\infty} f(x_n) = f(\sup S) \leqslant \gamma$. Notons $c \coloneqq \sup S$. On sait donc que $f(c) \leqslant \gamma$.

Supposons par l'absurde $f(c) < \gamma$. Soit $\epsilon \coloneqq \gamma - f(x) > 0$. Par la continuité de f en c, on sait qu'$\exists \delta > 0 \text{ t. q.} |x - c| < \delta \Rightarrow \left|f(x) - f(c)\right| < \epsilon = \gamma - f(c)$. On sait que $a < c < b$.

Prenons $x \coloneqq c + \frac{\min\{\delta, b-c\}}{2}$. Alors $x \in [a, b], |x - c| < \delta$ et $x > c$. Donc $\left|f(x) - f(c)\right| < \epsilon$, ou encore :

$$f(c) - \gamma < f(x) - f(c) < \gamma - f(c).$$

On sait donc $f(x) - f(c) < \gamma - f(c)$ ce qui implique $f(x) < \gamma$. Ce qui veut dire que $x \in S$. Or, par construction de x, $x > c = \sup S$. Il y a une contradiction. Donc l'hypothèse $f(c) < \gamma$ est fausse. Dès lors, $f(c) = \gamma$. □

Exemple On peut, par ce théorème, montrer que tout polynôme de degré n impair admet au moins une racine réelle : soit $P(x) \coloneqq \sum_{i=1}^{n} a_i x^i$. Prenons $Q(x) \coloneqq \frac{P(x)}{a_n}$. Dès lors, les limites en les infinis de Q sont ces mêmes infinis : $\lim_{x \to \pm\infty} Q(x) = \pm\infty$. On sait dès lors qu'il existe deux réels a et b tels que $Q(a) < 0$ et $Q(b) > 0$. Par le théorème, on sait qu'il existe un $c \in]a, b[$ tel que $Q(c) = 0$. Si $Q(c) = 0$, alors $P(c) = 0$ par définition de Q.

4.5 Théorème de l'intervalle et de la réciproque

Proposition 4.26. *Soit* $I \subseteq \mathbb{R}$. *Alors* I *est un intervalle si et seulement si* $\forall \alpha, \beta \in I : [\alpha, \beta] \subset I$.

Théorème 4.27 (de l'intervalle)**.** *Si* $f : I \to \mathbb{R}$ *est continue et* I *est un intervalle, alors* $f(I)$ *est un intervalle. De plus, si* I *est fermé borné,* $f(I)$ *l'est aussi.*

Définition 4.28. Soit $f : U \subseteq \mathbb{R} \to \mathbb{R}$ une fonction définie sur un ensemble U f est strictement croissante (respectivement strictement décroissante) si pour tout $x, y \in U : x < y \Rightarrow f(x) < f(y)$ (respectivement $x < y \Rightarrow f(x) > f(y)$). f est simplement croissante (respectivement simplement décroissante) si pour tout $x, y \in U : x < y \Rightarrow f(x) \leqslant f(y)$ (respectivement $x < y \Rightarrow f(x) \geqslant f(y)$). De plus, f est strictement monotone si elle est strictement croissante ou strictement décroissante et est simplement monotone si elle est simplement croissante ou simplement décroissante.

Théorème 4.29 (de la réciproque)**.** *Soit* $f : I \subseteq \mathbb{R} \to \mathbb{R}$ *une fonction continue et strictement monotone définie sur un intervalle* I. *Alors* f *est injective et sa réciproque* $f^{-1} : f(I) \to I$ *est aussi continue.*

Démonstration. Soient $x, y \in I$ tels que $x \neq y$. Soit $f^{-1} : f(I) \to \mathbb{R}$ la réciproque de f. Supposons f strictement croissante (le cas strictement décroissant est similaire). Soit $p \in \operatorname{int} f(I)$ (on sait que $f(I)$ est un intervalle). Montrons que f^{-1} est continue en p. Prenons $q \in I$ tel que $f(q) = p$. On sait que q n'est pas une extrémité de I car les extrémités de I correspondent aux extrémités de $f(I)$ (par la monotonie stricte de f) et que p est dans l'intérieur de $f(I)$. Prenons $\epsilon > 0$ tel que $]q - \epsilon, q + \epsilon[\subset I$. Posons $c \coloneqq f(q - \epsilon)$ et $d \coloneqq f(q + \epsilon)$. On sait $c < p < d$ par la monotonie stricte de f. Soit $\delta \coloneqq \min(pc, d - p)$. Dès lors, $\delta > 0$ et $[p - \delta, p + \delta] \subset f([q - \epsilon, q + \epsilon])$, ou encore $f^{-1}([p - \delta, p + \delta]) \subset [q - \epsilon, q + \epsilon]$. Nous avons donc δ tel que $|x - p| < \delta \Rightarrow \left|f^{-1}(x) - f^{-1}(p)\right| < \epsilon$. Ce qui montre que f^{-1}, la réciproque de f est continue en p. La continuité à droite et à gauche de f^{-1} aux extrémités de $f(I)$ se montrent de la même manière. $\square$

4.6 Continuité uniforme

La notion de continuité de f en a est une notion locale. Il en existe une version globale : la continuité uniforme.

Définition 4.30. Soit $f : U \subseteq \mathbb{R} \to \mathbb{R}$ une fonction définie sur un ensemble réel U. f est uniformément continue sur U si :

$$\forall \epsilon > 0 : \exists \delta > 0 \text{ t. q. } \forall x, y \in U : |x - y| < \delta \Rightarrow \left|f(x) - f(y)\right| < \epsilon.$$

Cette définition englobe la continuité simple mais requiert en plus que la valeur de δ trouvée ne dépende pas du point a choisi. Comme cette définition englobe l'autre, une fonction uniformément continue est également continue.

Définition 4.31. Deux suites $(x_n), (y_n) \subset \mathbb{R}$ sont équivalentes si $|x_n - y_n| \to 0$ pour $n \to +\infty$. Donc si l'une de des suites converge en L, l'autre converge également en L.

Proposition 4.32. *Soit* $f : U \subseteq \mathbb{R} \to \mathbb{R}$. *La fonction* f *est uniformément continue si et seulement si* $\forall (x_n), (y_n) \subset U$ *équivalentes, les suites* $(f(x_n)), (f(y_n))$ *sont équivalentes.*

Démonstration. Soit $f : U \to \mathbb{R}$ uniformément continue. Soient $(x_n), (y_n) \subset \mathbb{R}$ deux suites équivalentes. Soit $\epsilon > 0$. On sait qu'il existe δ tel que $\forall x, y \in U : |x - y| < \delta \Rightarrow \left|f(x) - f(y)\right| < \epsilon$ par l'uniforme continuité de f. De plus, par l'équivalence des suites, on sait qu'il existe $N \in \mathbb{N}$ tel que $|x_n - y_n| < \delta$. On a donc $|x_n - y_n| < \delta \Rightarrow \left|f(x_n) - f(y_n)\right| < \epsilon$. Les deux suites sont bien équivalentes. Montrons maintenant que l'équivalence de $(f(x_n))$ et $(f(y_n))$ implique l'uniforme continuité de f.

Supposons par l'absurde qu'il existe $\epsilon_0 > 0$ tel qu'il n'existe pas de $\delta_0 > 0$. En particulier, $\delta_0 = \frac{1}{n}$ ne convient pas, quel que soit $n > 0$. Prenons donc $n > 0$. Alors il existe $x_n, y_n \in U$ tels que $|x_n - y_n| < \frac{1}{n}$ et $\left|f(x_n) - f(y_n)\right| \geqslant \epsilon_0$. On a donc les suites $(x_n), (y_n)$ équivalentes mais pas les suites $(f(x_n))$ et $(f(y_n))$ ce qui est une contradiction avec l'hypothèse. $\square$

Théorème 4.33. *Si* $f : [a, b] \to \mathbb{R}$ *est une fonction continue, alors* f *est uniformément continue.*

version 1. Reprenons le format de démonstration précédente : supposons par l'absurde que f n'est pas continue. Dès lors, on sait qu'il existe $\epsilon_0 > 0$ tel qu'il n'existe pas de δ satisfaisant la définition. Plus précisément, $\forall n > 0 : \exists x_n, y_n \in [a, b]$ tels que $|x_n - y_n| < \frac{1}{n}$ et $\left|f(x_n) - f(y_n)\right| \geqslant \epsilon_0$. Maintenant, prenons deux sous-suites (x_{n_k}) et (y_{n_k}). Puisque (x_n) et (y_n) sont des suites bornées, les sous-suites (x_{n_k}) et (y_{n_k}) sont convergentes par le théorème de Bolzano-Weierstrass. Comme $|x_n - y_n| \to 0$, on sait que les sous-suites convergentes convergent vers la même valeur : $\lim_{k\to+\infty} x_{n_k} = \lim_{k\to+\infty} y_{n_k}$. Posons $p \coloneqq \lim_{k\to+\infty} x_{n_k}$. Comme les suites considérées sont dans $[a, b]$, on sait que $p \in [a, b]$ également. Or par la continuité de f, on sait que $f(x_{n_k}) \to f(p)$ et $f(y_{n_k}) \to f(p)$. Dès lors, $\left|f(x_{n_k}) - f(y_{n_k})\right| \to 0$. Or $\left|f(x_n) - f(y_n)\right| \geqslant \epsilon_0 \forall n$. Il y a donc contradiction. □

Lemme 4.34. *Soit* (x_n) *une suite bornée. Si* $\forall (x_{n_k})$ *sous-suite de* $(x_n) : x_{n_k} \to L$, *alors* $x_n \to L$.

Démonstration. Par Bolzano-Weierstrass, il existe une sous-suite (x_{n_k}) telle que $x_{n_k} \to \liminf(x_n) = L$ et une sous-suite (x_{m_k}) telle que $x_{m_k} \to \limsup(x_n) = L$. Si $\liminf(x_n) = \limsup(x_n) = L$, alors $x_n \to L$. □

version 2. Montrons que si $x_n - y_n \to 0$, alors $f(x_n) - f(y_n) \to 0$. Comme $f(x_n) - f(y_n)$ est bornée, prenons $f(x_{n_k}) - f(y_{n_k})$, une sous-suite convergente et nommons la limite L. Par Bolzano-Weierstrass, on sait qu'il existe deux sous-suites $(x_{n_{k_l}})$ et $(y_{n_{k_l}})$ qui convergent (respectivement en u et v). Alors on peut dire que $f(x_{n_{k_l}}) - f(x_{n_{k_l}}) \to f(u) - f(v) = L$. Or, comme $x_{n_{k_l}} - y_{n_{k_l}} \to 0$, on sait que $u = v$. Dès lors, on peut dire que $L = 0$. □

Théorème 4.35. *Soit* $f :]a, b[\to \mathbb{R}$ *une fonction uniformément continue. Alors il existe un prolongement continu* $\bar{f} : [a, b] \to \mathbb{R}$ *de* f.

Démonstration. Prenons la fonction

$$\bar{f} : [a, b] \to \mathbb{R} : x \mapsto \begin{cases} \lim\limits_{x\to a^+} f(x) & \text{si } x = a \\ f(x) & \text{si } x \in]a, b[\\ \lim\limits_{x\to b^-} f(x) & \text{si } x = b \end{cases}$$

Dès lors, la fonction $\bar{f}$ est uniformément continue. □

Remarque. Soit $f : A \to \mathbb{R}$ une fonction uniformément continue. $\forall E \subset A$, $f_{|_E}$ est également uniformément continue.

4.7 Fonctions à valeur vectorielle

Définition 4.36. Soient $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n$. On définit la distance entre x et y par :

$$\|x - y\| = \sum_{i=1}^{n} (x_i - y_i)^2.$$

De plus, la *norme* d'un vecteur x est la distance entre lui-même et l'origine. Donc $\|x\| = \sum_{i=1}^{n} x_i^2$.

Définition 4.37. Soient $x, y \in \mathbb{R}^n$. On définit le produit scalaire de x et y par :

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i.$$

Remarque. $\sqrt{\|x\|} = \langle x, x\rangle$.

Proposition 4.38. *Soient* $x, y, z \in \mathbb{R}^n, \lambda, \mu \in \mathbb{R}$. *Alors :*

1. $\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle$;
2. $\langle x, y\rangle = \langle y, x\rangle$;
3. $\langle x, x\rangle \geqslant 0$ *et* $\langle x, x\rangle = 0 \Rightarrow x = 0$.

Théorème 4.39 (Inégalité de Cauchy-Schwartz)**.** *Soient* $x, y \in \mathbb{R}^n$. *Alors* $\left|\langle x, y\rangle\right| \leqslant \|x\|\|y\|$ *avec égalité si et seulement si* x *et* y *sont colinéaires.*

Théorème 4.40 (Inégalité triangulaire)**.** *Soient* $x, y \in \mathbb{R}^n$. *Alors :*

$$\|x + y\| \leqslant \|x\| + \|y\|.$$

Démonstration.

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

□

Définition 4.41. Soient $r > 0$ et $a \in \mathbb{R}^n$. On définit la boule ouverte en a de rayon r par :

$$B(a, r) \coloneqq \{x \in \mathbb{R}^n : \|x - a\| < r\}.$$

Définition 4.42. Soient $A \subset \mathbb{R}^n$ et $a \in \mathbb{R}^n$. $a \in \operatorname{adh} A$ si et seulement si $\forall \epsilon > 0 : B(a, \epsilon) \cap A \neq \emptyset$.

Définition 4.43. Soient $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$, $a \in \operatorname{adh} A$ et $L \in \mathbb{R}^n$. On définit la limite de f pour $x \to a$ par :

$$\forall \epsilon > 0 : \exists \delta > 0 \text{ t. q. } \forall x \in A : \|x - a\| < \delta \Rightarrow \left\|f(x) - L\right\| < \epsilon.$$

Remarque.

$$\lim_{x \to a} f(x) = l \in \mathbb{R}^n \iff \forall \epsilon > 0 : \exists \delta > 0 \text{ t. q. } f(A \cap B(a, \delta)) \subseteq B(L, \epsilon).$$

Définition 4.44. Soient $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$, $a \in A$. f est continue en a si $\lim_{x \to a} f(x) = f(a)$.

Définition 4.45. Soit $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$. Les composantes de f sont les fonctions $f_i : A \subset \mathbb{R}^m \to \mathbb{R}^n$ telles que $f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$.

Lemme 4.46. *Soit* $f : A \subset \mathbb{R} \to \mathbb{R}^n$ *une fonction définie sur un voisinage de* $a \in \mathbb{R}$. *La fonction* f *possède une limite en* $x \to a$ *si et seulement si toutes les composantes* $f_i : A \to \mathbb{R}^n$ *possèdent une limite en* $x \to a$. *Alors :*

$$\lim_{x \to a} f(x) = \left(\lim_{x \to a} f_1(x), \lim_{x \to a} f_2(x), \ldots, \lim_{x \to a} f_n(x)\right).$$

Démonstration. Montrons d'abord que $f(x) \to L \Rightarrow \forall i : f_i(x) \to L_i$. Montrons ensuite l'autre sens de l'implication.

Soit $f : A \subset \mathbb{R} \to \mathbb{R}^n$ une fonction convergent en $L = (L_1, \ldots, L_n) \in \mathbb{R}^n$ en $x \to a$. Soit $\epsilon > 0$. On sait qu'il existe $\delta > 0$ tel que pour tout $x \in A, |x - a| < \delta \Rightarrow \left\|f(x) - L\right\| < \epsilon$. Dès lors :

$$\left|f_i(x) - L_i\right| = \sqrt{(f_i(x) - L_i)^2} \leqslant \left\|f(x) - L\right\| < \epsilon.$$

Pour montrer l'implication dans l'autre sens, prenons $\epsilon > 0$. $\forall i \in \{1, \dots, n\} \;:\; \exists \delta_i$ t. q.$|x - a| < \delta_i \Rightarrow \left|f_i(x) - L_i\right| < \frac{\epsilon}{\sqrt{n}}$. On définit $\delta \coloneqq \min\{\{\delta_i\}_i\}$. Dès lors, si$|x - a| < \delta$, on a :

$$\left\|f(x) - L\right\| = \sqrt{\sum_{i=1}^{n} (f_i(x) - L_i)^2} < \sqrt{\sum_{i=1}^{n} \left(\frac{\epsilon}{\sqrt{n}}\right)^2} = \sqrt{n \frac{\epsilon^2}{n}} = \epsilon.$$

□

Lemme 4.47. *La fonction* $f : A \subseteq \mathbb{R} \to \mathbb{R}^n$ *est continue en* a *si et seulement si toutes ses composantes sont continues en* a.

Démonstration. Par le lemme précédent, la démonstration est triviale. □

Lemme 4.48. *Soient* $f, g : A \subseteq \mathbb{R} \to \mathbb{R}^n$ *deux fonctions continues en* $a \in A$. *Alors les fonctions* $\|f\|$, $f + g$ *et* fg *sont également continues en* a.

Remarque. Étant donné que la continuité d'une fonction à valeur vectorielle est équivalente à la continuité de ses composantes, les théorèmes sur les fonctions les fonctions réelles s'appliquent facilement aux fonctions à valeur vectorielle.

Proposition 4.49. *Soit* $f : [a, b] \to \mathbb{R}^n$ *continue. Alors il existe* $p, q \in [a, b]$ *tels que* $\forall x \in [a, b] \;:\; \left\|f(q)\right\| \leqslant \left\|f(x)\right\| \leqslant \left\|f(p)\right\|$.

Démonstration. Par le le lemme précédent, on sait que $\|f\|$ est continue. En appliquant le théorème des bornes atteintes sur chaque composante, la proposition est démontrée. □

5 Fonctions dérivables

5.1 Définitions

Définition 5.1. Soit $f : I \to \mathbb{R}$ une fonction définie sur un intervalle ouvert I contenant a. La fonction f est dérivable en a si la limite suivante existe :

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

Si la limite existe, on la note $f'(a)$. Si la fonction f est dérivable sur tout point a de son domaine, f est dérivable. On définit la fonction dérivée de f par $f' : I \to \mathbb{R} : x \mapsto f'(x)$.

Définition 5.2. Il existe des classes de dérivabilité notées C^k pour $k \in \mathbb{N}$. Si f est continue, alors $f \in C^0$. De plus, si f est dérivable et f' est C^k, alors $f \in C^{k+1}$. Et si $\forall k \in \mathbb{N} : f \in C^k$, alors on note $f \in C^\infty$.

Proposition 5.3. *Soit* $f : A \subseteq \mathbb{R} \to \mathbb{R}$ *dérivable en* a. *Alors* f *est continue en* a.

Démonstration.

$$\lim_{x \to a} f(x) = \lim_{x \to a} f(a) + \lim_{x \to a} \frac{f(x) - f(a)}{x - a}(x - a) = f(a) + f'(a) \cdot 0 = f(a).$$

□

Théorème 5.4 (Règles de calcul). *Soient* $f, g : I \subseteq \mathbb{R} \to \mathbb{R}$ *deux fonctions définies sur un intervalle ouvert* I *et dérivables en* $a \in I$. *Alors :*

1. $(f+g)'(a) = f'(a) + g'(a)$;
2. $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$;
3. *si* $g(a) \neq 0$, *alors* $\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$.

Démonstration. Le premier point découle directement des règles de calcul sur la somme de limites. Le second point se montre en réécrivant la limite :

$$\lim_{x \to a} \frac{(fg)(x) - (fg)(a)}{x - a} = \lim_{x \to a} f(x) \frac{g(x) - g(a)}{x - a} + \lim_{x \to a} g(x) \frac{f(x) - f(a)}{x - a} = f'(a)g(a) + f(a)g'(a).$$

Le dernier se montre d'abord par $\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g(a)^2}$:

$$\lim_{x \to a} \frac{\frac{1}{g}(x) - \frac{1}{g}(a)}{x - a} = \lim_{x \to a} \frac{g(a) - g(x)}{(x - a)g(a)g(x)} = -\frac{g'(a)}{g(a)^2}.$$

En utilisant le point 2 et cette propriété, le quotient est démontré. □

Théorème 5.5. *Soient* f, g *deux fonctions telles que* f *est dérivable en* a *et* g *est dérivable en* $f(a)$. *Alors la composée* $g \circ f$ *est dérivable en* a *et vaut :*

$$g'(f(a))f'(a).$$

Démonstration. Soient les fonctions suivantes :

$$F(x) \coloneqq \left\{ \begin{array}{ll} \dfrac{f(x) - f(a)}{x - a} & \text{si } x \neq 0 \\ f\prime(a) & \text{si } x = a \end{array} \right\},$$

et :

$$G(x) \coloneqq \left\{ \begin{array}{ll} \dfrac{g(x) - g(f(a))}{x - f(a)} & \text{si } x \neq f(a) \\ g\prime(f(a)) & \text{si } x = f(a) \end{array} \right.$$

Les fonctions F et G sont respectivement continues en a et $f(a)$. De plus, $\forall x \in \operatorname{dom} f : f(x) = f(a) + (x - a)F(x)$ et $\forall x \in \operatorname{dom} g : g(x) = g(f(a)) + (x - f(a))G(x)$. Dès lors, on peut calculer $g \circ f$:

$$\begin{aligned} (g \circ f)(x) = g(f(x)) = g(f(a) + (x - a)F(x)) &= g(f(a)) + (f(a) + (x - a)F(x) - f(a))G(f(a) + (x - a)F(x)) \\ &= (g \circ f)(a) + (x - a)F(x)G(f(x)). \end{aligned}$$

On peut dès lors calculer la dérivée en faisant :

$$(g \circ f)'(a) = \lim_{x \to a} \frac{(g \circ f)(x) - (g \circ f)(a)}{x - a} = \lim_{x \to a} \frac{(x - a)F(x)G(f(x))}{x - a} = \lim_{x \to a} F(x)G(f(x)).$$

Et puisque le produit de fonctions continues est toujours une fonction continue, par la continuité, cette valeur vaut $F(a)G(f(a)) = f'(a)g'(f(a))$. □

Théorème 5.6 (de la réciproque)**.** *Soit* $f : I \to J$ *une bijection continue réelle entre deux intervalles ouverts. Si* f *est dérivable en* $a \in I$ *telle que* $f'(a) \neq 0$*, alors la réciproque* f^{-1} *est dérivable en* $f(a)$ *et vaut :*

$$\left(f^{-1}\right)'(f(a)) = \frac{1}{f'(a)}.$$

Démonstration. Puisque f est une bijection continue, elle est strictement monotone. Donc par un théorème précédent, on sait que $f^{-1} : J \to I$ est également continue. Posons :

$$G(y) \coloneqq \frac{f^{-1}(y) - a}{y - f(a)}.$$

Dès lors :

$$\lim_{y \to f(a)} G(y) = \lim_{y \to f(a))} G(f(f^{-1}(y))) = \lim_{x \to \lim_{y \to f(a)} f^{-1}(y)} \frac{x - a}{f(x) - f(a)} = \frac{1}{f'(a)}.$$

□

5.2 Extrema

Définition 5.7. Soient $f : U \subseteq \mathbb{R} \to \mathbb{R}$ et $a \in U$. Le point a est un minimum local de f si $\exists \epsilon > 0$ t. q. $\forall x \in U : |x - a| < \epsilon \Rightarrow f(a) \leqslant f(x)$. De même, a et un maximum local si $\exists \epsilon > 0$ t. q. $\forall x \in U : |x - a| < \epsilon \Rightarrow f(a) \geqslant f(x)$. Si a est un minimum local ou un maximum local, alors a est un extremum local.

Proposition 5.8. *Soit* $f : U \subseteq \mathbb{R} \to \mathbb{R}$ *une fonction dérivable. Soit* $a \in \operatorname{int} U$ *un extremum de* f*. Alors* $f'(a) = 0$*.*

Démonstration. Montrons le cas où a est un minimum local (le cas du maximum est identique). Par la définition du minimum, on sait qu'il existe ϵ tel que $\forall x \in U : |x-a| < \epsilon \Rightarrow f(a) \leqslant f(x)$. Dès lors :

$$\forall x \in U \text{ t. q. } x < a : \frac{f(x)-f(a)}{x-a} \leqslant 0,$$
$$\forall x \in U \text{ t. q. } x > a : \frac{f(x)-f(a)}{x-a} \geqslant 0.$$

Or, comme par hypothèse f est dérivable en a, la limite pour $x \to a$ existe. Il faut donc $f'(a) = 0$. □

Définition 5.9. Soit f une fonction dérivable. Un point $a \in \operatorname{dom} f$ tel que $f'(a) = 0$ est appelé point critique.

5.3 Théorème de la moyenne

Lemme 5.10 (Théorème de Rolle)**.** *Soit* $f : [a,b] \subseteq \mathbb{R} \to \mathbb{R}$ *continue et dérivable sur* $]a,b[$. *Si* $f(a) = f(b)$, *alors* $\exists c \in]a,b[\text{ t. q. } f'(c) = 0$.

Démonstration. La fonction f est définie sur un intervalle fermé borné. Donc par le théorème des bornes atteintes, on sait qu'il existe $m, M \in \mathbb{R}$ tels que $f(m)$ est le minimum de f et $f(M)$ est le maximum de f. Si $f(m) = f(M)$, alors la fonction est constante. Alors prenons $c = \frac{a+b}{2}$. Sinon, si $m \neq a$ et $m \neq b$, prenons $c = m$ car m est un extremum. Par la proposition précédente, $f'(c) = 0$. □

Théorème 5.11 (de la moyenne/eds accroissements finis)**.** *Soit* $f : [a,b] \subseteq \mathbb{R} \to \mathbb{R}$ *continue sur* $[a,b]$ *et dérivable sur* $]a,b[$. *Alors il existe* $c \in]a,b[$ *tel que :*

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

Démonstration. Soit $G(x)$ une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$ définie par :

$$G(x) \coloneqq f(x) - \frac{f(b)-f(a)}{b-a}x.$$

Dès lors, $G(b) - G(a) = 0$. Donc par le théorème de Rolle, on sait qu'il existe c tel que $G'(c) = 0$. Or

$$G'(x) = f'(x) - \frac{f(b)-f(a)}{b-a} = 0.$$

On a donc bien $f'(c) = \frac{f(b)-f(a)}{b-a}$. □

Proposition 5.12. *Soit* $f : [a,b] \to \mathbb{R}$ *continue sur* $[a,b]$ *et dérivable sur* $]a,b[$.

(i) *Si* $\forall x \in]a,b[: f'(x) > 0$, *alors* f *est strictement croissante sur* $]a,b[$ *;*

(ii) *Si* $\forall x \in]a,b[: f'(x) < 0$, *alors* f *est strictement décroissante sur* $]a,b[$ *;*

(iii) *Si* $\forall x \in]a,b[: f'(x) = 0$, *alors* f *est constante sur* $]a,b[$.

Démonstration. Puisque f est définie et continue sur un intervalle borné fermé, pour tout $x_1 < x_2 \in]a,b[$, on sait que :

$$\exists c \in]a,b[\text{ t. q. } f'(c) = \frac{f(x_1)-f(x_2)}{x_1-x_2}.$$

Donc si $f'(c) > 0$ (premier cas), il faut $f(x_2) > f(x_1)$, si $f'(c) < 0$ (second cas), il faut $f(x_2) < f(x_1)$ et si $f'(c) = 0$ (dernier cas), il faut $f(x_1 = f(x_2)$. On a donc f soit strictement croissante, soit strictement décroissante soit constante. □

Théorème 5.13 (Comparatif de la moyenne)**.** *Soient* $f, g : [a, b] \to \mathbb{R}$ *continues sur* $[a, b]$ *et dérivables sur* $]a, b[$. *Si pour tout* $x \in]a, b[$, *on a* $g(x) \neq 0$, *alors il existe* $c \in]a, b[$ *tel que :*

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

5.4 Règle de l'Hospital

Théorème 5.14. *Soient* $f, g : [c, d] \to \mathbb{R}$ *deux fonctions continues sur* $[c, d]$ *et dérivables sur* $]c, d[$. *Soit* $a \in]c, d[$. *Si* $f(a) = g(a) = 0$, *alors :*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

si cette limite existe.

Démonstration. Calculons d'abord la limite à droite, puis la limite à gauche. Soit $x \in]a, d]$. Dès lors, par le théorème comparatif de la moyenne sur des fonctions f et g réduites au domaine $[a, x]$, on sait qu'il existe $\xi(x) \in]a, x[$ tel que :

$$\frac{f'(\xi(x))}{g'(\xi(x))} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x) - 0}{g(x) - 0} = \frac{f(x)}{g(x)}.$$

De plus, quand $x \to a^+$, il faut $\xi(x) \to a^+$ car $a < \xi(x) < x$. On sait donc que :

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = \lim_{\xi(x) \to a^+} \frac{f'(\xi(x))}{g'(\xi(x))}.$$

De manière similaire, en prenant $x \in [c, a[$, on trouve $\xi(x) \in]x, a[$ tel que :

$$\frac{f'(\xi(x))}{g'(\xi(x))} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x)}{g(x)}.$$

À nouveau, quand $x \to a^-$, il faut $\xi(x) \to a^-$. Dès lors :

$$\lim_{x \to a^-} \frac{f(x)}{g(x)} = \lim_{\xi(x) \to a^-} \frac{f'(\xi(x))}{g'(\xi(x))}.$$

Donc si les deux limites existent et sont égales, on a :

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

□

Proposition 5.15. *Soient* $f, g : [c, d] \to \mathbb{R}$ *continues sur* $[c, d]$ *et dérivables sur* $]c, d[$. *Si* $\lim_{x \to a^+} g(x) = \pm\infty$, $\forall x : x \in]a, d[\Rightarrow g'(x) \neq 0$, *et si* $\lim_{x \to a^+} \frac{f'(x)}{g'(x)} = L \in \overline{\mathbb{R}}$ *alors :*

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = L.$$

Démonstration. Soient $x < y \in]a, d[$. Il existe $\xi \in]x, y[$ tel que :

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(\xi)}{g'(\xi)}.$$

Comme $g(x) \to \pm\infty$ quand $x \to a$, on sait qu'il existe un voisinage de a où $g(x) \neq 0$. De plus, en réécrivant :

$$f(x) = f(y) + (f(x) - f(y)) = f(y) + (g(x) - g(y))\frac{f(x) - f(y)}{g(x) - g(y)} = f(y) + (g(x) - g(y))\frac{f'(\xi)}{g'(\xi)},$$

on peut trouver :

$$\frac{f(x)}{g(x)} = \frac{1}{g(x)}\left[f(y) + (g(x) - g(y))\frac{f'(\xi)}{g'(\xi)}\right] = \frac{f(y)}{g(x)} + \left(1 - \frac{g(y)}{g(x)}\right)\frac{f'(\xi)}{g'(\xi)}.$$

Soit $\epsilon > 0$. Prenons $\delta_1 > 0$ tel que $\forall \chi \in]a, a + \delta_1[: \left|\frac{f'(\chi)}{g'(\chi)} - L\right| < \epsilon_1$. Fixons $y = a + \delta_1$. On a $\frac{g(y)}{g(x)} \to 0$ quand $x \to a$ car $g(y)$ est fixé et $g(x) \to +\infty$ (pareil pour $\frac{f(y)}{g(x)} \to 0$). Dès lors, il existe δ_2 tel que $\forall x : x \in]a, a + \delta_2[\Rightarrow \left|\frac{g(y)}{g(x)}\right| < \epsilon_2$ et $\left|\frac{f(y)}{g(x)}\right| < \epsilon_3$. Prenons $(\epsilon_1, \epsilon_2, \epsilon_3) = \left(\frac{\epsilon}{6}, \min\left\{\frac{\epsilon}{3|L|+1}, \frac{1}{2}\right\}, \frac{\epsilon}{3}\right)$. Prenons $\delta := \min\{\delta_1, \delta_2\}$. On a donc :

$$\begin{aligned}
\forall x : x \in]a, a + \delta[: \left|\frac{f(x)}{g(x)} - L\right| &= \left|\frac{f(y)}{g(x)} + \left(1 - \frac{g(y)}{g(x)}\right)\left(\frac{f'(\chi)}{g'(\chi)} - L + L\right) - L\right| \\
&= \left|\frac{f(y)}{g(x)} + \left(\frac{f'(\chi)}{g'(\chi)} - L\right) - \frac{g(y)}{g(x)}\left(\frac{f'(\chi)}{g'(\chi)} - L\right) + L - \frac{g(y)}{g(x)}L - L\right| \\
&< \epsilon_3 + \epsilon_1 + \epsilon_2\epsilon_1 + |L|\,\epsilon_2 \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{6} + \frac{\epsilon}{6} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

□

5.5 Dérivées de fonctions à valeur dans $\mathbb{R}^n$

Définition 5.16. Soit $f : A \subseteq \mathbb{R} \to \mathbb{R}^n$. Soit $a \in \operatorname{int} A$. f est dérivable en a si :

$$\lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

existe dans $\mathbb{R}^n$ cet élément.

Lemme 5.17. *La fonction* $r : A \subseteq \mathbb{R} \to \mathbb{R}^n$ *est dérivable en* $a \in \operatorname{int} A$ *si et seulement si toutes ses composantes sont dérivables. En ce cas,* $f'(a) = (f_1'(a), \dots, f_n'(a))$.

Théorème 5.18. *Soient* $f : \mathbb{R} \to \mathbb{R}$ *dérivable en* $a \in \mathbb{R}$ *et* $g : \mathbb{R} \to \mathbb{R}^n$ *dérivable en* $f(a)$. *Alors* $(g \circ f)'(a) = f'(a)g'(f(a))$.

Démonstration. Par le lemme précédent et la règle de dérivation de composée pour les fonctions réelles. □

Théorème 5.19. *Soient* $f : \mathbb{R} \to \mathbb{R}, g : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R} \to \mathbb{R}^n$ *et* $a \in \mathbb{R}$ *tels que* f, g, h *sont dérivables en* a*. Alors :*

1. $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$;
2. $(\langle g, h\rangle)'(a) = \left\langle h'(a), g(a)\right\rangle + \left\langle h(a), g'(a)\right\rangle$;
3. $\left(\frac{g}{f}\right)'(a) = \frac{g'(a)f(a)-g(a)f'(a)}{f(a)^2}$;
4. $(g+h)'(a) = g'(a) + h'(a)$.

5.6 Dérivées de fonctions vectorielles

Définition 5.20. Le graphe d'une fonction $f : A \subseteq \mathbb{R}^m \to \mathbb{R}$ est l'ensemble des couples :

$$\Gamma_f \coloneqq \{((x, f(x)) \in \mathbb{R}^n \times \mathbb{R} \cong \mathbb{R}^{n+1} \text{ t. q. } x \in A\}.$$

Définition 5.21. Soit $f : A \subseteq \mathbb{R}^m \to \mathbb{R}$. Soit $a \in \operatorname{int} A$. La jème dérivée partielle de f (pour $1 \leqslant j < m$) est donnée par :

$$\lim_{h\to 0} \frac{f(a_1, \ldots, a_j + h, \ldots, a_n) - f(a)}{h} = \lim_{h\to 0} \frac{f(a + he_j) - f(a)}{h}.$$

si cette limite existe et se note $\frac{\partial f}{\partial x_j}(a)$ ou $(\partial_j f)(a)$. Si la limite n'existe pas, alors f n'est pas dérivable en a.

Définition 5.22. On définit le gradient de $f : \mathbb{R}^m \to \mathbb{R}$ en a par le vecteur des dérivées partielles :

$$(\nabla f)(a) = ((\partial_1 f)(a), \ldots, (\partial_m f)(a)) = ((\partial_i f)(a))_i \in \mathbb{R}^m.$$

Définition 5.23. e_j est le jème vecteur de la base canonique de $\mathbb{R}^n$.

Définition 5.24. Soit $f : \mathbb{R}^m \to \mathbb{R}$, $a \in \mathbb{R}^m$ et $v \in \mathbb{R}^m$. f est dérivable en a dans la direction v si la limite suivante existe dans $\mathbb{R}$:

$$\lim_{h\to 0} \frac{f(a + hv) - f(a)}{h}.$$

On note $(\partial_v f)(a)$ ce réel et on l'appelle dérivée directionnelle de la fonction f dans la direction v au point a.

Remarque. Comme e_j est un vecteur de la base de $\mathbb{R}^m$, $(\partial_j f)(a) = (\partial_{e_j} f)(a)$ est une dérivée directionnelle.

Définition 5.25. Soient $f : \mathbb{R}^m \to \mathbb{R}$, $a \in \mathbb{R}^m$. La fonction f est différentiable s'il existe $u \in \mathbb{R}^m$ tel que :

$$\lim_{x\to a} \frac{f(x) - f(a) - \langle u, x - a\rangle}{\|x - a\|} = 0 \in \mathbb{R}.$$

Proposition 5.26. *Soient* $f : \mathbb{R}^m \to \mathbb{R}$, $a \in \mathbb{R}^m$*. Si* f *est différentiable en* a*, alors la fonction* $\partial_. f : \mathbb{R}^m \to \mathbb{R} : v \mapsto (\partial_v f)(a)$ *est définie pour tout vecteur* v *et est linéaire. De plus,* $\forall v \in \mathbb{R}^m : (\partial_v f)(a) = \left\langle(\nabla f)(a), v\right\rangle$.

Démonstration. Soit $x(t) \coloneqq a + tv$. Donc, $x(t) \to a$ si $t \to 0$. Dès lors :

$$\begin{aligned}
\lim_{t\to 0} \frac{f(a+tv) - (f(a) + \langle u, tv\rangle)}{|t|} &= 0 \\
\lim_{t\to 0} \frac{f(a+tv) - f(a)}{|t|} - \lim_{t\to 0} \frac{\langle u, tv\rangle}{|t|} &= 0 \\
(\partial_v f)(a) - \langle u, v\rangle \lim_{t\to 0} \frac{t}{t} &= 0 \\
(\partial_v f)(a) &= \langle u, v\rangle.
\end{aligned}$$

Dès lors, pour $v = e_j$, on a $(\partial_{e_j} f)(a) = (\partial_j f)(a) = \langle u, v\rangle = \left\langle u, e_j \right\rangle = u_j$. Donc $u = ((\partial_1 f)(a), \ldots, (\partial_m f)(a)) = (\nabla f)(a)$. $\square$

6 Intégrales de Riemann

6.1 Définitions

Définition 6.1. une partition de $[a, b]$ est la donnée $\{x_i \text{ t. q. } 0 \leqslant i \leqslant n\} \subset [a, b]$ telle que $a = x_0 < \ldots < x_n = b$.

Définition 6.2. Soit $f : [a, b] \to \mathbb{R}$ bornée. Soit $P = \{x_i\}_{i \in [n]}$ une partition de $[a, b]$. Pour tout $1 \leqslant i \leqslant n$, on définit :

$$\begin{aligned} m_i &\coloneqq \inf\{f(x) \text{ t. q. } x \in [x_{i-1}, x_i]\}, \\ M_i &\coloneqq \sup\{f(x) \text{ t. q. } x \in [x_{i-1}, x_i]\}. \end{aligned}$$

On définit ensuite :

$$\begin{aligned} \mathcal{L}(f, P) &\coloneqq \sum_{i=1}^{n} (x_i - x_{i-1}) m_i, \\ \mathcal{U}(f, P) &\coloneqq \sum_{i=1}^{n} (x_i - x_{i-1}) M_i. \end{aligned}$$

Qui sont respectivement l'aire signée de la somme des rectangles inférieurs et supérieurs. À partir de cela, on définit :

$$\begin{aligned} \mathcal{L}(f) &\coloneqq \sup\{\mathcal{L}(f, P) \text{ t. q. } P \text{ est une partition de } [a, b]\}, \\ \mathcal{U}(f) &\coloneqq \inf\{\mathcal{U}(f, P) \text{ t. q. } P \text{ est une partition de } [a, b]\}. \end{aligned}$$

Définition 6.3. une fonction $f : [a, b] \to \mathbb{R}$ bornée est intégrale si $\mathcal{U}(f) = \mathcal{L}(f)$. On note cette valeur $\int_a^b f(x)\, dx$ ou encore $\int_a^b f$.

Lemme 6.4. *Si* $f : [a, b] \to \mathbb{R}$ *est bornée et* P *est une partition de* $[a, b]$, $y \in [a, b]$. *On définit* $P' \coloneqq P \cup \{y\}$. *Alors :*

$$\mathcal{L}(f, P) \leqslant \mathcal{L}(f, P) \leqslant \mathcal{U}(f, P) \leqslant \mathcal{U}(f, P').$$

Démonstration. Soit $r \in \{1, \ldots, n\}$ tel que $x_{r-1} < y < x_r$. On sait que $\mathcal{L}(f, P) = m_1(x_1 - x_0) + \ldots + m_n(x_n - x_{n-1})$ et on sait que $\mathcal{L}(f, P') = m_1(x_1 - x_0) + \ldots + \alpha(y - x_{r-1}) + \beta(x_r - y) + \ldots + m_n(x_n - x_{n-1})$ où $\alpha = \inf\{f(x) \text{ t. q. } x \in [x_{r-1}, y]\}$ et $\beta = \inf\{f(x) \text{ t. q. } x \in [y, x_r]\}$. Dès lors :

$$\begin{aligned} \mathcal{L}(f, P') - \mathcal{L}(f, P) &= \alpha(y - x_{r-1}) + \beta(x_r - y) - m_r(x_r - x_{r-1}) = x_{r-1}(-\alpha + m_r) + x_r(\beta - m_r) + \alpha y - \beta y + m_r y - m_r y \\ &= (\alpha - m_r)(y - x_{r-1}) + (\beta - m_r)(x_r - y). \end{aligned}$$

Étant donné que $x_{r-1} < y < x_r$, on sait que les secondes parenthèses sont positives. De plus, comme les intervalles dont α et β sont les minima sont inclus dans l'intervalle dont m_r est le minimum, il est nécessaire que $\alpha \geqslant m_r$ et $\beta \geqslant m_r$. Les premières parenthèses sont dès lors également positives. Si $\mathcal{L}(f, P') - \mathcal{L}(f, P) \geqslant 0$, alors $\mathcal{L}(f, P') \geqslant \mathcal{L}(f, P)$. La partie pour les aires supérieures est identique. $\square$

Corollaire 6.5. *Soient* P, P' *deux partitions de* $[a, b]$ *telles que* $P \subset P'$*, alors* $\mathcal{L}(f, P') \leqslant \mathcal{L}(f, P) \leqslant \mathcal{U}(f, P) \leqslant \mathcal{U}(f, P')$.

Démonstration. En écrivant $P' = P \cup \{y_1, \dots, y_k\}$ et en appliquant k fois le lemme précédent. □

Lemme 6.6. *Soient* P *et* P' *deux partitions quelconques du même intervalle* $[a, b]$*. Alors :*

$$\mathcal{L}(f, P) \leqslant \mathcal{U}(f, P').$$

Démonstration. Soit $P'' = P \cup P'$. Dès lors, on sait $P'' \subset P'$ et $P'' \subset P$. Donc :

$$\mathcal{L}(f, P) \leqslant \mathcal{L}(f, P'') \leqslant \mathcal{U}(f, P'') \leqslant \mathcal{U}(f, P').$$

□

Proposition 6.7. *Soit* $f : [a, b] \to \mathbb{R}$ *bornée. Alors* $\mathcal{L}(f) \leqslant \mathcal{U}(f)$.

Démonstration. On sait que pour tout P, P' partitions de $[a, b]$, $\mathcal{L}(f, P) \leqslant \mathcal{U}(f, P')$. En particulier, en prenant le sup à gauche et l'inf à droite, l'inégalité reste vraie. Donc :

$$\begin{gathered}\sup\{\mathcal{L}(f, P) \text{ t. q. } P \text{ est une partition de } [a, b]\} \leqslant \inf\{\mathcal{U}(f, P) \text{ t. q. } P \text{ est une partition de } [a, b]\} \\ \mathcal{L}(f) \leqslant \mathcal{U}(f)\end{gathered}$$

□

6.2 Fonctions intégrables

Proposition 6.8 (Critère de Riemann)**.** *Soit* $f : [a, b] \to \mathbb{R}$ *bornée. Alors* f *est intégrable si et seulement* $\forall \epsilon > 0$: $\exists P$ *une partition de* $[a, b]$ *telle que* $\mathcal{U}(f, P) - \mathcal{L}(f, P) < \epsilon$.

Démonstration. Montrons d'abord l'implication $\Rightarrow$. On sait par hypothèse que $\mathcal{L}(f) = \mathcal{U}(f)$. Soit $\epsilon > 0$. On sait qu'il existe P_1, P_2 partitions de $[a, b]$ tels que $\mathcal{U}(f, P_1) < \mathcal{U}(f) + \frac{\epsilon}{2}$ et $\mathcal{L}(f, P_2) > \mathcal{L}(f) + \frac{\epsilon}{2}$. Posons $P \coloneqq P_1 \cup P_2$. Dès lors :

$$\mathcal{L}(f) - \frac{\epsilon}{2} \leqslant \mathcal{L}(f, P_2) \leqslant \mathcal{L}(f, P) \leqslant \mathcal{L}(f) \overset{\text{par hypothèse}}{=} \mathcal{U}(f) \leqslant \mathcal{U}(f, P) \leqslant \mathcal{U}(f, P_1) \leqslant \mathcal{U}(f) + \frac{\epsilon}{2}.$$

On sait donc $\mathcal{U}(f, P) < \mathcal{U}(f, P_1) < \mathcal{U}(f) + \frac{\epsilon}{2}$ et $-\mathcal{L}(f, P) < -\mathcal{L}(f, P_2) < -\mathcal{L}(f) + \frac{\epsilon}{2}$. Donc :

$$\mathcal{U}(f, P) - \mathcal{L}(f, P) < \mathcal{U}(f) + \frac{\epsilon}{2} - \mathcal{L}(f) + \frac{\epsilon}{2} = \epsilon.$$

Montrons maintenant l'autre sens de l'implication $\Leftarrow$. Soit $\epsilon > 0$. On sait qu'il existe P une partition de $[a, b]$ telle que $\mathcal{U}(f, P) - \mathcal{L}(f, P) < \epsilon$. De plus, on sait $0 \leqslant \mathcal{U}(f) - \mathcal{L}(f) \leqslant \mathcal{U}(f, P) - \mathcal{L}(f, P) < \epsilon$. On a donc une quantité $\mathcal{U}(f) - \mathcal{L}(f)$ ne dépendant pas de ϵ mais étant plus petite qu'ϵ pour tout $\epsilon > 0$. Il faut dès lors $\mathcal{U}(f) - \mathcal{L}(f) = 0$. □

Théorème 6.9. *Si* $f : [a, b] \to \mathbb{R}$ *est continue, alors* f *est intégrable.*

Démonstration. Soit $\epsilon > 0$. Si f est continue, alors elle est uniformément continue. Donc il existe δ tel que :

$$\forall x, y : |x - y| < \delta \Rightarrow \left|f(x) - f(y)\right| < \frac{\epsilon}{b - a}.$$

On construit P telle que $\forall 1 \leqslant 1 \leqslant n : x_i - x_{i-1} < \delta$. Pour tout i, on sait qu'il existe x_* et x^* tels que $M_i = f(x^*)$ et $m_i = f(x_*)$ par le théorème des bornes atteintes. Donc :

$$\mathcal{U}(f, P) - \mathcal{L}(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} (f(x^*) - f(x_*))(x_i - x_{i-1}) \leqslant \frac{\epsilon}{b - a} \sum_{i=1}^{n} (x_i - x_{i-1}) = \frac{\epsilon}{b - a}(b - a) = \epsilon.$$

□

6.3 Propriétés des intégrales

Définition 6.10. Soit $f : [a, b] \to \mathbb{R}$ intégrable. On définit :

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

Proposition 6.11. *Soient $f, g : [a, b] \to \mathbb{R}$ intégrables. Soient $\alpha, \beta \in \mathbb{R}$ et $c \in [a, b]$. Alors :*

1. $\int_a^b \big(\alpha f(x) + \beta g(x)\big)\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx$;
2. $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$;
3. *Si* $\forall x \in [a, b] : f(x) \leqslant g(x)$, *alors* $\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx$;
4. $|f|$ *est intégrable et* $\left|\int_a^b f(x)\,dx\right| \leqslant \left|\int_a^b \left|f(x)\right|\,dx\right|$.

6.4 Théorème fondamental de l'analyse

Théorème 6.12 (Théorème fondamental du calcul et intégral)**.** *Soit $f : [a, b] \to \mathbb{R}$ continue (donc intégrable). Alors la fonction $F : [a, b] \to \mathbb{R} : x \mapsto \int_a^x f$ est l'unique primitive de f sur $[a, b]$ qui s'annule en a.*

Démonstration. Soit $c \in [a, b]$. Soit $\epsilon > 0$. Par la continuité(uniforme) de f en c, il existe δ tel que $|x - c| < \delta \Rightarrow \left|f(x) - f(c)\right| < \epsilon$. De plus, notons que :

$$\left|\frac{F(x) - F(c)}{x - c} - f(c)\right| = \left|\frac{\int_a^x f - \int_a^c f}{x - c} - f(x)\right| = \left|\frac{\int_c^x f - f(c)(x - c)}{x - c}\right| = \left|\frac{\left|\int_c^x (f(t) - f(c))\,dt\right|}{x - c}\right| \leqslant \frac{1}{|x - c|}\left|\int_c^x \left|f(t) - f(c)\right|\,dt\right|$$

Et si $|x - c| < \delta$, alors $\forall t \in [c, x] : |t - c| < \delta$. Et donc $\left|f(t) - f(c)\right| < \epsilon$. Donc :

$$\frac{1}{|x - c|}\left|\int_c^x \left|f(t) - f(c)\right|\,dt\right| < \frac{1}{|x - c|}\left|\int_c^x \epsilon\,dt\right| = \epsilon.$$

On a donc montré que pour $x \to c$, on a $\frac{F(x) - F(c)}{x - c} = F'(c) \to f(c)$.

Montrons maintenant que F est l'**unique** primitive s'annulant en a. Soit $G : [a, b] \to \mathbb{R}$ telle que $G' = f$ et $G(a) = 0$. Montrons que $G = F$:

$$(G - F)' = (f - f)' = 0.$$

On sait donc que $G - F$ est une fonction constante. Et comme $G(a) = F(a) = 0$, on sait que $\forall x \in [a, b]$: $(G - F)(x) = 0$. $\square$

Corollaire 6.13. *Soit* $f : [a, b] \to \mathbb{R}$ *intégrable. Sois* $F := \int f$. *On a :*

$$\int_a^b f(x)\,dx = F(b) - F(a) = \left[F(x)\right]_{x=a}^{x=b}.$$

Démonstration. Les fonctions suivantes :

$$x \mapsto \int_a^x f(t)\,dt,$$
$$x \mapsto F(x) - F(a),$$

sont deux primitives de f s'annulant en a et donc sont égales. $\square$

Proposition 6.14. *Soient* $f, g : [a, b] \to \mathbb{R} \in C^1$. *Alors :*

$$\int_a^b f(x)g'(x)\,dx = \left[(fg)(x)\right]_{x=a}^{x=b} - \int_a^b f'(x)g(x)\,dx.$$

Démonstration. Soit $h(x) := (fg)'(x) = f'(x)g(x) + f(x)g'(x)$. Par le théorème fondamental, on sait :

$$\int_a^b h'(x)\,dx = \int_a^b (f'(x)g(x) + f(x)g'(x))\,dx = \left[h(x)\right]_{x=a}^{x=b}.$$

Ou encore :

$$\int_a^b f'(x)g(x)\,dx = \left[h(x)\right]_{x=a}^{x=b} - \int_a^b f(x)g'(x)\,dx.$$

$\square$

Proposition 6.15. *Soient* $f : [a, b] \to \mathbb{R}$ *continue et* $g : [a, b] \to \mathbb{R} \in C^1$. *Posons* $\alpha := g(a), \beta := g(b)$. *Alors :*

$$\int_a^b f(x)\,dx = \int_\alpha^\beta (f \circ g)(t)g'(t)\,dt.$$

Démonstration. Soient $F(x) := \int_a^x f(t)\,dt$ et $G(t) := (F \circ g)(t)$. Dès lors, on sait que $G'(t) = (f \circ g)(t)g'(t)$. De plus, l'intégration bornée donne :

$$\int_\alpha^\beta (f \circ g)(t)g'(t)\,dt = \int_\alpha^\beta G'(t)\,dt = \left[G(t)\right]_{t=\alpha}^{t=\beta} = (F \circ g)(\beta) - (F \circ g)(\alpha) = \int_a^b f(x)\,dx.$$

$\square$

6.5 Les intégrales impropres

Définition 6.16. Soit $f : [a, +\infty[\to \mathbb{R}$ une fonction bornée et intégrable sur tout $[a, b]$ pour $b > a$. Alors si la limite suivante existe :

$$\lim_{b \to +\infty} \int_a^b f(x)\, dx,$$

on dit que $\int_a^{+\infty} f(x)\, dx$ converge. On appelle une telle limite une intégrale impropre. Si la limite n'existe pas, on dit que $\int_a^{\infty}$ diverge. De manière similaire, soit $f :]-\infty, b] \to \mathbb{R}$ une fonction bornée et intégrable sur tout $[a, b]$ pour $a < b$. Alors si la limite suivante existe :

$$\lim_{a \to -\infty} \int_a^b f(x)\, dx,$$

on dit que $\int_{-\infty}^b f(x)\, dx$ converge. Pour une fonction $f : \mathbb{R} \to \mathbb{R}$, bornée, si les deux intégrales impropres suivantes existent :

$$\int_{-\infty}^0 f(x)\, dx, \int_0^{+\infty} f(x)\, dx,$$

alors on définit l'intégrale suivante :

$$\int_{-\infty}^{+\infty} f(x)\, dx = \int_{-\infty}^0 f(x)\, dx + \int_0^{+\infty} f(x)\, dx.$$

Définition 6.17. De manière similaire, pour des fonctions non bornées, on a les définitions suivantes. Soit $f :]a, b] \to \mathbb{R}$. Si f est intégrable sur tout $[c, b]$ avec $c \in]a, b]$, alors on définit :

$$\int_a^b f(x)\, dx = \lim_{c \to a} \int_c^b f(x)\, dx.$$

De même, soit $f : [a, b[\to \mathbb{R}$ intégrable sur tout $[a, c]$ pour $c \in [a, b[$. On définit alors :

$$\int_a^b f(x)\, dx = \lim_{c \to b} \int_a^c f(x)\, dx.$$

Et finalement soit $f :]a, b[\to \mathbb{R}$. Si les deux intégrales impropres existent, on définit :

$$\int_a^b f(x)\, dx = \int_a^{\frac{a+b}{2}} f(x)\, dx + \int_{\frac{a+b}{2}}^b f(x)\, dx.$$

Proposition 6.18 (Critère de comparaison)**.** *Soient* $f, g : [a, +\infty[\to \mathbb{R}$ *intégrables sur* $[a, b]$ *pour tout* $b > a$*. Si* $\forall x > a : 0 \leqslant f(x) \leqslant g(x)$ *et* $\int_a^{+\infty} g(x)\, dx$ *converge, alors* $\int_a^{+\infty} f(x)\, dx$ *converge également telle que :*

$$\int_a^{+\infty} f(x)\, dx \leqslant \int_a^{+\infty} g(x)\, dx.$$

Démonstration. Comme $f(x) \leqslant g(x)$ pour tout $x > a$, on sait que pour tout $b > a$, on a :

$$\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx.$$

De plus, les fonctions suivantes sont croissantes car f et g sont toujours positives :

$$F : b \mapsto \int_a^b f(x)\,dx,$$
$$G : b \mapsto \int_a^b g(x)\,dx.$$

Dès lors, F est majorée et bornée par $\int_a^{+\infty} g(x)\,dx$. La limite de $F(b)$ pour $b \to +\infty$ existe. □

6.6 Longueur de courbes

Définition 6.19. Une courbe différentiable est une application $\gamma[a,b] \subseteq \mathbb{R} \to \mathbb{R}^n$ de classe C^1 telle que $\forall t : t \in [a,b] \Rightarrow \gamma'(t) \neq 0 \in \mathbb{R}^n$.

Définition 6.20. L'ensemble $C \coloneqq \operatorname{Im} \gamma$ est la courbe différentiable associée à γ.

Définition 6.21. Soit $\gamma : [a,b] \to \mathbb{R}^n$ une courbe différentiable. Si $\gamma(a) = \gamma(b)$, alors γ est un lacet. Et si γ est un lacet tel que $\gamma'(a) = \gamma'(b)$ (dérivées respectivement « à droite » et « à gauche »), alors γ est un lacet différentiable.

Définition 6.22. Si γ est une courbe injective sur $]a,b[$, alors γ est dite simple.

Définition 6.23. Soit γ une courbe paramétrée simple et différentiable. On définit sa longueur par :

$$L(\gamma) \coloneqq \int_a^b \left\|\gamma'(t)\right\| dt.$$

Proposition 6.24. *Soient* $f : [a,b] \to [c,d] \in C^1$ *bijective,* $\gamma : [a,b] \to \mathbb{R}^n$ *et* $\eta : [c,d] \to \mathbb{R}^n$ *deux courbes paramétrées différentiables. Supposons que* $\gamma = (\eta \circ f)$ *(que* γ *est une reparamétrisation de* η*). Alors* $L(\gamma) = L(\eta)$.

Démonstration. Supposons f croissante. Dès lors (en posant $u \coloneqq f(t)$) :

$$\begin{aligned} L(\gamma) = L(\eta \circ f) &= \int_a^b \left\|(\eta \circ f)'(t)\right\| dt = \int_a^b \left\|(\eta' \circ f)(t) f'(t)\right\| dt = \int_a^b \left\|(\eta' \circ f)(t)\right\| \left|f'(t)\right| dt \\ &= \int_a^b \left\|(\eta' \circ f)(t)\right\| f'(t)\, dt = \int_{f(a)}^{f(b)} \left\|\eta'(u)\right\| du = \int_c^d \left\|\eta'(u)\right\| du = L(\eta). \end{aligned}$$

Si f était décroissant, il aurait fallu mettre un $-$ en sortant $f'(t)$ de la valeur absolue mais les bornes d'intégration $f(a)$ et $f(b)$ auraient respectivement donné d et c. Donc en rentrant le moins dans les bornes d'intégration, on obtient le même résultat. □

Définition 6.25. Soit $\gamma : [a,b] \to \mathbb{R}^n$ une courbe paramétrée différentiable telle que $\forall t : \left\|\gamma'(t)\right\| = 1$. Alors γ est une paramétrisation par longueur.

Remarque. La longueur d'une telle courbe est $b-a$ et la distance entre deux points p, q quelconques est $|p-q|$.

Lemme 6.26. *Toute courbe paramétrée différentiable possède une paramétrisation par longueur.*

Démonstration. Soit $\gamma : [a,b] \to \mathbb{R}^n$ une courbe paramétrée différentiable. Définissons $\ell : [a,b] \to \mathbb{R}^+ : t \mapsto \int_a^t \gamma'(u)\,du$. ℓ est une fonction strictement croissante et une bijection $\in C^1$. Soit $\eta = (\gamma \circ \ell^{-1})$. On a :

$$\left\|\eta'(t)\right\| = \left\|\gamma((\ell^{-1})(t))(\ell^{-1})'(t)\right\| = \left\|\gamma'((\ell^{-1})(t))\right\| \frac{1}{\left|\ell'((\ell^{-1})(t))\right|}.$$

Et comme, par définition, $\ell'(t) = \left\|\gamma'(t)\right\|$, on a $\left\|\eta'(t)\right\| = 1$. □

6.7 Intégrales curvilignes

Définition 6.27. Soient $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$ et $\gamma : [a,b] \to \mathbb{R}^n$ avec $\operatorname{Im}\gamma \subseteq E$. On définit l'intégrale de f le long de la courbe γ par :

$$\int_\gamma f\,ds \coloneqq \int_a^b f(\gamma(t))\left\|f'(t)\right\|dt.$$

Définition 6.28. Un champ de vecteurs est une application $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$.

Définition 6.29. Soient $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ un champ de vecteurs continu et $\gamma : [a,b] \to \mathbb{R}^n$ une courbe paramétrée différentiable. On définit le travail de f le long de γ par :

$$\int_\gamma \langle f, ds\rangle \coloneqq \int_a^b \left\langle f(\gamma(t)), \gamma'(t)\right\rangle dt.$$

Proposition 6.30. *Soient γ et η deux paramétrisations d'une même courbe. Alors $\int_\gamma \langle f, ds\rangle = \pm \int_\eta \langle f, ds\rangle$. Le signe moins peut apparaitre si γ et η dont d'orientation opposée.*

6.8 Champs conservatifs

Définition 6.31. Soit $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ un champ de vecteurs. f est conservatif si il existe $F : E \to \mathbb{R}$ tel que $f = \nabla F$.

Proposition 6.32. *Soit $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ un champ conservatif. Alors $\int_\gamma \langle f, ds\rangle$ ne dépend que des extrémités de γ.*

Démonstration. Soit $F = \nabla f$.

$$\int_\gamma \langle f, ds\rangle = \int_a^b \left\langle \nabla F(\gamma(t)), \gamma'(t)\right\rangle dt = \int_a^b (F(\gamma(t)))'(t)\,dt = F(\gamma(b)) - F(\gamma(a)).$$

□

7 Fonctions continues à plusieurs variables

7.1 Introduction

Remarque. C'est d'abord l'espace euclidien $\mathbb{R}^n$ qui sera étudié ici.

Définition 7.1. Soit $f : \mathbb{R}^m \to \mathbb{R}^n$ une fonction. On dit que f est une fonction à m variables et à valeur vectorielle.

Rappel Soit $U \subseteq \mathbb{R}^m$. On dit que U est un voisinage de $a \in \mathbb{R}^m$ s'il existe $\delta > 0$ tel que pour tout $x \in \mathbb{R}^m$, si $\|x - a\| < \delta$, alors $x \in U$.

7.2 Norme et distance dans $\mathbb{R}^n$

Le choix de la norme euclidienne pour exprimer la distance dans $\mathbb{R}^n$ est arbitraire. Mais une fonction de distance $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ doit respecter certaines propriétés.

Définition 7.2. Soit X un ensemble. Une **distance** sur X est une application $d : X \times X \to \mathbb{R}$ qui respecte les propriétés suivantes $\forall x, y \in \mathbb{R}^n$:

1. $d(x, y) \geqslant 0$;
2. $d(x, y) = 0 \iff x = y$;
3. $d(x, y) = d(y, x)$;
4. $\forall z \in \mathbb{R}^n : d(x, z) \leqslant d(x, y) + d(y, z)$.

Définition 7.3. Une **norme** sur $\mathbb{R}^n$ est une application $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ respectant les propriétés suivantes $\forall x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$:

1. $\|x\| \geqslant 0$;
2. $\|\lambda x\| = |\lambda| \|x\|$;
3. $\|x\| = 0 \iff x = 0$;
4. $\|x + y\| \leqslant \|x\| + \|y\|$.

Proposition 7.4. *La définition d'une norme sur* $\mathbb{R}^n$ *implique l'existence d'une distance* $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} : (x, y) \mapsto \|x - y\|$.

Démonstration. Soient $x, y, z \in \mathbb{R}^n$. Montrons les propriétés nécessaires pour une distance :

1. $d(x, y) = \|x - y\| \geqslant 0$;
2. $0 = d(x, y) = \|x - y\| \iff x - y = 0 \iff x = y$;
3. $d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1| \|y - x\| = \|y - x\| = d(y, x)$;
4. $d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leqslant \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$.

$\square$

Lemme 7.5. *Soient* $x, y \in \mathbb{R}^n$. *Alors* $|\langle x, y \rangle| \leqslant \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}$.

Démonstration. Tout d'abord, supposons $y = 0$. On a alors :

$$\left|\langle x, 0\rangle\right| = 0 = \sqrt{\langle x, x\rangle}\sqrt{\langle 0, 0\rangle}.$$

L'inégalité large est donc vérifiée puisque les deux membres sont égaux.

On sait que $\forall \lambda \in \mathbb{R} : 0 \leqslant \langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - \lambda\langle y, x\rangle - \lambda(\langle x, y\rangle - \lambda\langle y, y\rangle)$. Supposons maintenant $y \neq 0$. Prenons λ tel que $\langle x, y\rangle - \lambda\langle y, y\rangle = 0$ (c.-à-d. $\lambda = \frac{\langle x,y\rangle}{\langle y,y\rangle}$). On a donc :

$$0 \leqslant \langle x, x\rangle - \lambda y x = \langle x, x\rangle - \frac{\langle x, y\rangle}{\langle y, y\rangle}\langle y, x\rangle = \langle x, x\rangle - \frac{\langle x, y\rangle^2}{\langle y, y\rangle}.$$

□

7.3 Convergence des suites dans $\mathbb{R}^n$

Définition 7.6 (Convergence de suite dans $\mathbb{R}^n$)**.** Soit $(x_n)_n \subset \mathbb{R}^n$. $(x_n)_n$ converge en $x \in \mathbb{R}^n$ si :

$$\forall \epsilon > 0 : \exists K(\epsilon) \in \mathbb{N}\, \text{t. q.}\, \forall k > K(\epsilon) : \|x_k - x\| < \epsilon,$$

ce que l'on écrit :

$$\lim_{k \to +\infty} \|x_k - x\| = 0.$$

Définition 7.7. Soit $(x_k) \subset \mathbb{R}^n$ une suite. $(x_{k\,i})_k$ désigne la suite de la ième composante de (x_k).

Proposition 7.8. *La suite $(x_n)_n \subset \mathbb{R}^n$ converge en $x \in \mathbb{R}^n$ si et seulement si chaque suite de composante $(x_{k\,i})_k \subset \mathbb{R}$ converge en x_i.*

Démonstration. Dans le sens direct, on suppose $x_k \to x \in \mathbb{R}^n$. Prenons $1 \leqslant i \leqslant n$, alors :

$$|x_{k\,i} - x_i| \leqslant \|x_k - x\| \to 0.$$

Dans le sens indirect, on suppose que $\forall i \in \{1, \ldots, n\} : x_{k\,i}$. Alors :

$$\|x_k - x\|^2 \coloneqq \sum_{i=1}^{n} (x_{k\,i} - x_i)^2 \to 0.$$

□

Corollaire 7.9. *Soit $(x_k)_k \subset \mathbb{R}^n$. Pour tout $1 \leqslant p, q \leqslant n - 1$, la suite $(x_k)_k \coloneqq (y_k, z_k)_k \subset \mathbb{R}^p \times \mathbb{R}^q$ converge si et seulement si (y_k) converge dans $\mathbb{R}^p$ et $qz_k)$ converge dans $\mathbb{R}^q$.*

Démonstration. Trivial par la proposition 7.8. □

Définition 7.10. Une suite $(x_k)_k \subset \mathbb{R}^n$ est dite de Cauchy si $\forall 1 \leqslant i \leqslant n$, la suite $(x_{k\,i})_k \subset \mathbb{R}$ est de Cauchy.

Théorème 7.11. *La suite $(x_k)_k \subset \mathbb{R}^n$ converge si et seulement si elle est de Cauchy.*

Lemme 7.12 (Inégalité de Young)**.**

$$\forall a, b \in \mathbb{R}^+, p, q \in (0, +\infty)\, \text{t. q.}\, p + q = pq : \frac{a^p}{p} + \frac{b^q}{q} \geqslant ab.$$

Démonstration. Par les propriétés de la fonction log (concavité), on a :

$$\log(ab) = \log\left(a^{\frac{p}{p}}\right) + \log\left(b^{\frac{q}{q}}\right) = \frac{1}{p}\log\left(a^p\right) + \frac{1}{q}\log\left(b^q\right) \leqslant \log\left(\frac{a^p}{p} + \frac{b^q}{q}\right).$$

La fonction exp est croissante, donc on peut composer les deux membres de l'inégalité avec exp, et on obtient bien l'inégalité de Young. □

Définition 7.13. Soit $x \in \mathbb{R}^n$ un vecteur. On définit la norme $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$ telle que :

$$\|x\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}.$$

Théorème 7.14 (Inégalité de Minkowski[3])**.** *Soient* $p, q \in [1, +\infty)$ *tels que* $p + q = pq$. *Alors* $\forall x, y \in \mathbb{R}^n$:

$$\left|\sum_{i=1}^n x_i y_i\right| \leqslant \sum_{i=1}^n |x_i y_i| \leqslant \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q\right)^{\frac{1}{q}}.$$

Démonstration. Par Young, on a le cas où $\sum_{k=1}^n |x_k|^p = \sum_{k=1}^n |y_k|^q = 1$:

$$|x_k||y_k| \leqslant \frac{|x_k|^p}{p} + \frac{|y_k|^q}{q}.$$

En sommant sur k de 1 à n, on obtient :

$$\sum_{k=1}^n |x_k||y_k| \leqslant \frac{1}{p}\sum_{k=1}^n |x_k|^p + \frac{1}{q}\sum_{k=1}^n |y_k|^q = \frac{1}{p} + \frac{1}{q} = 1 = \left(\sum_{k=1}^n |x_k|^p\right)\left(\sum_{k=1}^n |y_k|^q\right).$$

Si $\|x\|_p \neq 1$ ou $\|y\|_q \neq 1$, alors on renormalise les vecteurs avec $x' := \frac{x}{\|x\|_p}$, $y' := \frac{y}{\|y\|_q}$. On a alors :

$$\frac{1}{\|x\|_p \|y\|_q}\sum_{k=1}^n |x_k y_k| = \sum_{k=1}^n \left|x_k' y_k'\right| \leqslant 1,$$

ou encore :

$$\sum_{k=1}^n |x_k y_k| \leqslant \|x\|_p \|y\|_q .$$

□

Définition 7.15. On définit $\|\cdot\|_\infty := \lim_{p\to+\infty} \|\cdot\|_p$. Ce qui donne, pour $x \in \mathbb{R}^n$:

$$\|x\|_\infty = \max_{k=1,\dots,n} |x_k| .$$

Définition 7.16. Les normes $\|\cdot\|_a$ et $\|\cdot\|_b$ sont dites équivalentes s'il existe deux constantes $C_1, C_2 \in \mathbb{R}_0^+$ telles que :

$$\forall x \in \mathbb{R}^n : C_1\|x\|_b \leqslant \|x\|_a \leqslant C_2\|x\|_b .$$

[3]Généralisation de l'inégalité de Cauchy-Schwartz en dimension $n \geqslant 1$

Remarque. On observe alors que :

$$(\|x\|_p)^p = \sum_{k=1}^{n} |x_k|^p \leqslant n(\max_k |x_k|)^p \leqslant n \sum_{k=1}^{n} |x_k|^p \,.$$

En prenant la racine pème des inégalités, on obtient :

$$\|x\|_p \leqslant \sqrt[p]{n}\|x\|_\infty \leqslant \sqrt[p]{n}\|x\|_p \,.$$

Selon la définition 7.16, on remarque que pour tout p, les normes $\|\cdot\|_p$ et $\|\cdot\|_\infty$ sont équivalentes. On remarque donc que la notion de convergence (définie par $\|\cdot\|_2$, la norme euclidienne) aurait pu être définie selon une norme $\|\cdot\|_p$ quelconque.[4]

7.4 Limite d'une fonction en un point

Définition 7.17. Soit $f : \mathbb{R}^m \to \mathbb{R}^n$. Si $m > 1$, on dit que f est une fonction à plusieurs variables réelles, et si $n > 1$, on dit que f est à valeurs vectorielles.

Définition 7.18 (Adéhrence en dimension $n > 1$)**.** Soit $A \subseteq \mathbb{R}^n$. Un point $a \in \mathbb{R}^n$ est dit **adhérent** à A si $\forall \delta > 0 : \exists x \in A \text{ t. q.} \|x - a\| < \delta$.

L'ensemble des points adhérents à A s'appelle l'*adhérence* de A et se note $\operatorname{adh} A$.

Lemme 7.19. *Soient* $A \subseteq \mathbb{R}^n$ *et* $a \in \mathbb{R}^n$. $a \in \operatorname{adh} A$ *si et seulement si* $\exists (x_k) \subset A \text{ t. q. } x_k \to a$.

Démonstration. Pour le sens indirect, soit $(x_k) \subset A$ une suite convergente en $a \in \mathbb{R}^n$. Soit $\delta > 0$. On sait qu'il existe $N(\delta) \in \mathbb{N}^*$ tel que $\forall n \geqslant N(\delta) : \|x_n - a\| < \delta$. Donc $a \in \operatorname{adh} A$.

Pour le sens direct, on prend $a \in \operatorname{adh} A$. On sait donc que pour tout $\delta > 0$, il existe $x(\delta) \in A$ t. q. $\left\|x(\delta) - a\right\| < \delta$. On construit alors une suite $(x_k) \subset A$ telle que $x_k \coloneqq x(1/k)$ pour tout $k \geqslant 1$. Dès lors, soit $\delta > 0$. On sait qu'il existe $N = \left\lceil \frac{1}{\delta} \right\rceil$ tel que $\forall n > N : \|x_n - a\| < \delta$. Donc $(x_k) \subset A$ par définition, et on vient de montrer que $x_k \to a$. □

Définition 7.20. La boule ouverte a été définie au 4.41 ; on définit également la boule fermée de centre a et de rayon r par :

$$\bar{B}(a, r) \coloneqq \{x \in \mathbb{R}^n \text{ t. q.} \|x - a\| \leqslant r\} = B(a, r) \cup \{x \in \mathbb{R}^n \text{ t. q.} \|x - a\| = r\}.$$

Remarque. L'adhérence d'un ensemble A peut également se définir à l'aide de la boule ouverte :

$$\operatorname{adh} A = \{x \in \mathbb{R}^n \text{ t. q. } \forall r > 0 : B(x, r) \cup A \neq \emptyset\}.$$

Remarque. La boule fermée $\bar{B}(a, r)$ est l'adhérence de la boule ouverte $B(a, r)$.

Définition 7.21 (Intérieur en dimension $n > 1$)**.** Soit $A \subset \mathbb{R}^n$. Un élément $a \in \mathbb{R}^n$ est dit *intérieur* à A si $\exists \delta > 0 \text{ t. q. } \forall x \in \mathbb{R}^n : \|x - a\| < \delta \Rightarrow x \in A$.

Remarque. Géométriquement, on dit que l'intérieur $\operatorname{int} A$ est l'ensemble :

$$\operatorname{int} A \coloneqq \{x \in \mathbb{R}^n \text{ t. q. } \exists r > 0 \text{ t. q. } B(x, r) \subseteq A\}.$$

[4] Pour être complet, toutes les normes sur $\mathbb{R}^n$ sont équivalentes, mais ce résultat dépasse le contenu de ce cours.

Définition 7.22. La frontière de A est l'ensemble $\operatorname{adh} A \setminus \operatorname{int} A$. On le note ∂A.

Définition 7.23. Un ensemble $A \subseteq \mathbb{R}^n$ est :

- *ouvert* si et seulement si $A = \operatorname{int} A$;
- *fermé* si et seulement si $A = \operatorname{adh} A$.

Définition 7.24 (Limite d'une fonction multivariée à valeur vectorielle)**.** Soient $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$, $B \subset \mathbb{R}^m$, et $a \in \operatorname{adh}(U \cap B)$. La limite de f dans B pour x tendant vers a existe et vaut $L \in \mathbb{R}^n$ si :

$$\forall \epsilon > 0 : \exists \delta > 0 \text{ t. q. } \forall x \in \operatorname{adh}(U \cap B) : \|x - a\| < \delta \Rightarrow \left\| f(x) - L \right\| < \epsilon.$$

On écrit cela :

$$\lim_{x \xrightarrow{B} a} f(x) = L \qquad \text{ou} \qquad \lim_{\substack{x \to a \\ x \in B}} f(x) = L.$$

Remarque. Dans la définition 7.24, ϵ sert à déterminer une boule ouverte $B(L, \epsilon)$ et δ sert à déterminer une boule ouverte $B(a, \delta)$ tels que :

$$f\left(B(a, \delta) \cap (B \cap U)\right) \subset B(L, \epsilon).$$

Définition 7.25. La limite de $f : \mathbb{R}^m \to \mathbb{R}^n$ en x tendant vers $a \in \mathbb{R}^m$ existe et vaut $L \in \mathbb{R}^n$ si et seulement si la limite de f dans $B = \mathbb{R}^m$ existe et vaut L pour x tendant vers a. On note cela :

$$\lim_{x \to a} f(x) = L.$$

Définition 7.26. Soient $f : \mathbb{R}^m \to \mathbb{R}^n$, et $a \in \operatorname{adh}\left(\operatorname{dom} f \setminus \{a\}\right)$. Alors la limite pointée :

$$\lim_{\substack{x \to a \\ x \neq a}} f(x) = L$$

est vérifiée si la limite de f dans $B = \mathbb{R}^m \setminus \{a\}$ existe et vaut L.

Proposition 7.27. *Soient* $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$, $B \subset \mathbb{R}^m$, *et* $a \in \operatorname{adh}(B \cap U)$. *La limite de* f *dans* B *pour* $x \to a$ *existe et vaut* L *si et seulement si* $\forall A \subseteq B$: *la limite de* f *dans* A *de* $x \to a$ *existe et vaut* L.

Proposition 7.28. *Soient* $f : \mathbb{R}^m \to \mathbb{R}^n$, *et* $a \in \operatorname{adh}(\operatorname{dom} f)$. *Alors :*

$$\lim_{x \to a} f(x) = L \iff \forall (x_n)_n \subset \operatorname{dom} f : x_n \to a \Rightarrow (f(x_n))_n \to L.$$

Démonstration. Supposons d'abord que $f(x) \to L$ quand $x \to a$.

Soit $\epsilon > 0$. On sait qu'il existe $\delta > 0$ tel que :

$$\forall x \in \mathbb{R}^m : \|x - a\| < \delta \Rightarrow \left\| f(x) - L \right\| < \epsilon.$$

Soit $(x_k) \subset \mathbb{R}^m$ une suite convergente en a. On sait qu'il existe $N > 0$ tel que :

$$\forall n > N : \|x_n - a\| < \delta.$$

Dès lors :

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

Supposons ensuite que pour toute suite $(x_k) \subset \mathbb{R}^m$ convergente en a, on ait $(f(x_k))_k \subseteq \mathbb{R}^n$ convergente en L. Supposons par l'absurde que $f(x) \not\to L$ pour $x \to a$. On sait alors qu'il existe ϵ_0 tel que :

$$\forall \delta_0 > 0 : \exists x \in \mathbb{R}^m \text{ t. q.} \|x_n - a\| < \delta \wedge \left\| f(x_n) - L \right\| \geqslant \epsilon_0.$$

En particulier, pour $\delta_0 = \frac{1}{n}$, on trouve :

$$\forall n \geqslant 1 : \exists x_n \in \mathbb{R}^m \text{ t. q.} \|x_n - a\| < \delta_0 = \frac{1}{n} \wedge \left\|f(x_n) - L\right\| \geqslant \epsilon_0.$$

Or, par hypothèse, o na supposé que pour toute suite $(x_k) \subset \mathbb{R}^m$ t. q. $x_k \to a$, la suite $(f(x_k))_k$ convergeait en L. Ce qui est une contradiction, et donc $f(x) \to L$. □

7.5 Continuité d'une fonction

Définition 7.29. Soient $f : \mathbb{R}^m \to \mathbb{R}^n$, et $a \in \operatorname{dom} f$. On dit que f est continue en a si :

$$\lim_{x \to a} f(x) = f(a),$$

c'est-à-dire :

$$\forall \epsilon > 0 : \exists \delta > 0 \text{ t. q. } \forall x \in \operatorname{dom} f : \|x - a\| < \delta \Rightarrow \left\|f(x) - f(a)\right\| < \epsilon.$$

Définition 7.30. Une fonction $f : \mathbb{R}^m \to \mathbb{R}^n$ est dite continue si elle est continue en tout point de son domaine.

Proposition 7.31. *Soient* $f : \mathbb{R}^m \to \mathbb{R}^p$ *et* $g : \mathbb{R}^p \to \mathbb{R}^n$ *deux fonctions continues. Alors* $(g \circ f)$ *est également continue.*

Démonstration. Soit $\epsilon > 0$. On sait (par la continuité de g) qu'il existe $\delta_g > 0$ tel que :

$$\forall y \in \mathbb{R}^p : \left\|y - f(a)\right\| < \delta_g \Rightarrow \left\|g(y) - g(f(a))\right\| < \epsilon.$$

On sait également, par continuité de f qu'il existe $\delta_f > 0$ tel que :

$$\forall x \in \mathbb{R}^m : \|x - a\| < \delta_f \Rightarrow \left\|f(x) - f(a)\right\| < \delta_g.$$

Dès lors, on trouve :

$$\forall x \in \mathbb{R}^m : \|x - a\| < \delta_f \Rightarrow \left\|f(x) - f(a)\right\| < \delta_g \Rightarrow \left\|g(f(x)) - g(f(a))\right\| < \epsilon.$$

□

7.5.1 Fonctions à valeur réelle

Définition 7.32. On appelle la fonction $\pi_j : \mathbb{R}^m \to \mathbb{R}$ définie par :

$$\pi_j((x_1, \dots, x_m)) = x_j$$

la jème fonction de projection.

Lemme 7.33. *La fonction de projection* π_j *est une fonction continue.*

Démonstration. Soient $\epsilon > 0$, et $a \in \mathbb{R}^m$. Prenons $\delta = \epsilon$. En supposant $\|x - a\| < \delta$, on trouve :

$$\epsilon = \delta > \|x - a\| = \sqrt{(x_j - a_j)^2 + \sum_{j \neq i} (x_i - a_i)^2} \geqslant \sqrt{(x_j - a_j)^2} = \left|\pi_j(x) - \pi_j(a)\right|.$$

On a bien $\left|\pi_j(x) - \pi_j(a)\right| < \epsilon$. □

Proposition 7.34. *Soient* $f : \mathbb{R}^m \to \mathbb{R}$, *et* $a \in \mathbb{R}^m$. *Alors* f *est continue en* a *si et seulement si :*

$$\forall (x_k)_k \subset \mathbb{R} : x_k \to a \Rightarrow (f(x_k))_k \to f(a).$$

Démonstration. Corollaire direct de la proposition 7.28. □

Remarque. Ce résultat nous permet d'écrire :

$$\lim_{k\to+\infty} f(x_k) = f\left(\lim_{k\to+\infty} x_k\right)$$

pour f continue, ce qui est assez pratique pour les règles de calcul.

Corollaire 7.35. *Soit* $f : \mathbb{R}^m \to \mathbb{R}$. *S'il existe deux suites* $(x_k), (y_k) \subset \mathbb{R}^m$ *convergentes en* a *telles que :*

$$\lim_{k\to+\infty} f(x_k) \neq \lim_{k\to+\infty} f(y_k),$$

alors f *n'est pas continue.*

Remarque. Ce corollaire est en réalité la contraposée de la proposition 7.34.

7.5.2 Fonctions à valeur vectorielle

Théorème 7.36. *Une fonction* $f = (f_1, \dots, f_n) : \mathbb{R}^m \to \mathbb{R}^n$ *est continue en* $a \in \mathbb{R}^m$ *si et seulement si les* n *fonctions* f_i *sont continues en* a.

Démonstration. Si f est continue, alors $\forall i \in \{1, \dots, n\} : f_i \coloneqq (\pi_i \circ f)$ est continue par composée de fonctions continues.

Si f_i est continue pour tout i, on calcule :

$$\lim_{x\to a} \left\| f(x) - f(a) \right\| \coloneqq \lim_{x\to a} \sqrt{\sum_{i=1}^{n} \left| f_i(x) - f_i(a) \right|^2} = \sqrt{\sum_{i=1}^{n} \lim_{x\to a} \left| f_i(x) - f_i(a) \right|^2} = 0,$$

ce qui est équivalent à $\lim_{x\to a} f(x) = f(a)$. □

Remarque. Pour déterminer la continuité de f, on peut déterminer la continuité de ses composantes. On ne peut cependant pas déterminer la continuité des restrictions de la fonction selon les différents arguments !

7.6 Valeur intermédiaire et ensembles connexes par arcs

Définition 7.37. Soit $E \subset \mathbb{R}^n$. On dit que l'ensemble E est *connexe par arcs* s'il existe $\gamma : [0,1] \to E$ une application continue telle que $\gamma(0) = p$ et $\gamma(1) = q$ pour tout $p, q \in E$.

γ est appelé un *chemine de* p *à* q.

Proposition 7.38. *L'ensemble des ensembles connexes par arcs de* $\mathbb{R}$ *sont les intervalles.*

Démonstration. Montrons que les intervalles sont des ensembles connexes. Soit $I = [p, q]$ un intervalle. Alors $\gamma : t \mapsto tq + (1-t)p$ est un chemin de p à q.

Soit I un ensemble connexe par arcs. S'il y a un chemin de p à q dans I, alors $\forall r \in \mathbb{R}$ t. q. $p < r < q$, il faut $r \ni I$. □

Théorème 7.39 (Image continue d'un ensemble connexe par arcs). *Soit* $f : E \subset \mathbb{R}^m \to \mathbb{R}^n$ *où* E *est un ensemble connexe par arcs. Si* f *est continue sur* E*, alors* $f(E)$ *est connexe par arcs.*

Démonstration. Prenons $p, q \in f(E)$. Par définition, on sait qu'il existe $a, b \in E$ tels que $f(a) = p$ et $f(b) = q$. Soit γ un chemin de a à b dans E. Par continuité de γ on sait que $(f \circ \gamma)$ est continue et représente un chemin de p à q. ☐

Corollaire 7.40. *Soit* $f : E \subset \mathbb{R}^m \to \mathbb{R}$ *une fonction continue où* E *est un ensemble connexe par arcs. Si* $\alpha < \beta$ *sont deux valeurs prises par* f*, alors* $\forall c \in (\alpha, \beta) : \exists m \in E \text{ t.q. } f(m) = c$.

Démonstration. Corollaire direct du théorème 7.39 : $f(E)$ est un intervalle. ☐

Proposition 7.41. *Il n'existe pas de bijection continue* $f : \mathbb{R}^m \to \mathbb{R}$ *pour* $m > 1$.

Démonstration. Supposons par l'absurde qu'il existe $f : \mathbb{R}^n \to \mathbb{R}$ continue et bijective. Prenons $g : \mathbb{R}^m \setminus \{0\} \to \mathbb{R} \setminus \{x\}$, la restriction de f à $\mathbb{R} \setminus \{x\}$ où $x = f(0) \in \mathbb{R}^m$. On observe que $\mathbb{R}^m \setminus \{0\}$ est un ensemble connexe par arcs alors que $\mathbb{R} \setminus \{x\}$ n'est pas un intervalle, ce qui contredit le théorème 7.39. ☐

Remarque. Pour être complet, il est possible de déterminer une bijection de $\mathbb{R}^n$ dans $\mathbb{R}$ mais elle ne peut être continue. Ces ensembles ont donc la même cardinalité. Il existe un résultat similaire disant qu'il n'existe pas de bijection continue $\mathbb{R}^m \to \mathbb{R}^n$ pour $n \neq m$, mais qui dépasse le contenu de ce cours.

7.7 Bornes atteintes et ensembles compacts

Théorème 7.42 (Théorème de Bolzano-Weierstrass dans $\mathbb{R}^n$). *Dans* $\mathbb{R}^m$*, toute suite bornée* $(x_k)_k$ *contient une sous-suite convergente.*

Démonstration. On note $x_{k\,i}$ la ième composante du vecteur x_k. Comme x_k est bornée par hypothèse, on sait que pour tout i, la suite $(x_{k_i})_k$ est également bornée. Par le théorème 3.2 de Bolzano-Weierstrass, on sait que chaque composante de la suite contient une sous-suite convergente.

On nomme $(x_{k_{l(1)}\,1})_k$ une sous-suite convergente de $(x_{k\,1})_k$. On s'intéresse à la suite $(x_{k_{l(1)}})_k$. On prend également une sous-suite convergente de $(x_{k\,2})_k$ que l'on appelle $(x_{k_{m(2)}\,2})_k$. On prend alors les éléments communs entre $(x_{k_{l(1)}})_k$ et $(x_{m(2)})_k$ et on appelle cette suite $(x_{k_{l(2)}})$. On répète cet argument au total m fois afin d'obtenir une suite $(x_{k_{l(m)}})$ qui est l'intersection entre tous les $(x_{k_{m(i)}})$. Cette suite est une sous-suite de $(x_k)_k$ est est convergente en toutes ses composantes par construction. ☐

Proposition 7.43. *Un ensemble* $E \subset \mathbb{R}^m$ *est fermé et borné si et seulement si :*

$$\forall (x_k)_k \subset E : \exists (x_{k_l})_l \text{ convergente en } e \in E.$$

Démonstration. Supposons d'abord que E est fermé et borné. Par Bolzano-Weierstrass, on sait que toute suite admet une sous-suite convergente en une valeur e adhérente à E. Or puisque E est fermé, on sait que $e \in E$.

Supposons maintenant que toute suite dans E admet une sous-suite convergente. Prenons $x \in \operatorname{adh} E$. Soit $(x_k)_k \subset E$, une suite convergente en x. On sait donc qu'il existe une sous-suite de (x_k) qui converge dans E par hypothèse, et par unicité de la limite, on sait que la sous-suite converge en x. Donc $x \in E$, et donc $\operatorname{adh} E \subset E$, ou encore, E est fermé.

On montre que E est borné par l'absurde en trouvant $y_i \in E \setminus B(y_{i-1}, R_i)$, qui montre que la suite $(y_i)_i$ n'admet pas de sous-suite convergente, ce qui est une contradiction. ☐

Définition 7.44. Un ensemble $E \subset \mathbb{R}^m$ est dit *compact* si toute suite $(x_k) \subset E$ admet une sous-suite convergente dans E.

Théorème 7.45 (Théorème des bornes atteintes)**.** *Soit $f : E \subseteq \mathbb{R}^m \to \mathbb{R}$ une fonction continue. Si E est un ensemble fermé borné, alors $f(E) \subseteq \mathbb{R}$ est un ensemble fermé borné. En particulier, f est bornée et atteint ses bornes.*

Démonstration. Notons m, M respectivement les quantités $\inf_{x\in E} f(x)$ et $\sup_{x\in E} f(x)$. Montrons que f est bornée, et donc supposons par l'absurde que $m = -\infty$ ou $M = +\infty$. On trouve donc une suite $(x_k) \subset E$ telle que $\left|f(x_k)\right| \to +\infty$ (ou encore $\forall k \in \mathbb{N} : \left|f(x_k)\right| \geqslant k$). Puisque E est un ensemble fermé borné, on peut extraire (x_{k_n}) une sous-suite de (x_k) convergente en $x \in E$. Par continuité de f, on sait que $f(x_{k_n})$ est également une suite convergente en $f(x)$. La fonction f est donc bornée, ce qui contredit l'hypothèse.

Montrons ensuite que $m \in f(E)$. (L'argument pour montrer $M \in f(E)$ est identique.) Par définition de l'infimum, il existe $(x_k) \subset E$ une suite telle que $f(x_k)$ est convergente en m. Puisque E est fermé borné, on peut extraire une sous-suite (x_{k_n}) telle que $x_{k_n} \to x \in E$. On sait alors $f(x_{k_n}) \to f(x)$ par continuité de f, et donc $m \in f(E)$.

Montrons maintenant que $f(E)$ est fermé. Prenons $c \in \operatorname{adh} f(E)$. On sait qu'il existe une suite $(x_k) \subset f(E)$ convergente en c. Si x_k est dans E, alors il existe $z_k \in E$ tel que $x_k = f(z_k)$. La suite $(z_k) \subset E$ possède une sous-suite (z_{k_n}) convergente en $x \in E$. Par continuité de f, on sait que $f(z_k) \to f(x) \in f(E)$. Or, par unicité de la limite, on a $c = f(x) \in f(E)$. □

7.8 continuité uniforme

Définition 7.46. Soit $f : E \subseteq \mathbb{R}^m \to \mathbb{R}$ une fonction. Si :

$$\forall \epsilon > 0 : \exists \delta > 0 \text{ t. q. } \forall x, y \in E : \|x - y\| < \delta \Rightarrow \left|f(x) - f(y)\right| < \epsilon,$$

alors on dit que que f est *uniformément continue*.

Définition 7.47. Une fonction $f : E \subseteq \mathbb{R}^m \to \mathbb{R}$ est dite *Lipschitzienne* si :

$$\exists \alpha > 0 \text{ t. q. } \forall x, y \in E : \left\|f(x) - f(y)\right\| \leqslant \alpha \|x - y\|.$$

Lemme 7.48. *Une fonction Lipschitzienne est uniformément continue.*

Démonstration. Soit $f : E \subseteq \mathbb{R}^m \to \mathbb{R}$ une fonction Lipschitzienne. Soit $\epsilon > 0$. Prenons $\delta = \frac{\epsilon}{\alpha}$ où α est le paramètre lipschitzien de f. Supposons que $\|x - y\| < \delta$, on trouve alors :

$$\left|f(x) - f(y)\right| \leqslant \alpha \|x - y\| < \alpha\delta = \epsilon.$$

□

Remarque. On remarque également que l'uniforme continuité implique la continuité.

Proposition 7.49. *Soit $f : E \subseteq \mathbb{R}^m \to \mathbb{R}$ une fonction continue définie sur E fermé borné. Alors f est uniformément continue.*

Démonstration. Par Bolzano-Weierstrass. □

8 Fonctions différentiables à plusieurs variables

8.1 Rappels

On appelle la *différentielle* en a l'application $(\partial_v f)(a)$ qui envoie v sur $\langle (\nabla f)(a), v\rangle$, où $(\nabla f)(a)$ est le gradient de f au point a, donné par le vecteur $\big((\partial_i f)(a)\big)_i$.

Définition 8.1. Soit $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ et soient $a \in \operatorname{int} E, v \in \mathbb{R}^m$. On dit que f *est dérivable au point* a *dans la direction* v si :

$$\forall i \in \{1, \cdots, n\} : (\partial_i f)(a) \coloneqq \lim_{h\to 0} \frac{f_i(a+hv) - f_i(a)}{h} \qquad \text{existe.}$$

Dans ce cas, on note :

$$(\partial_v f)(a) \coloneqq \lim_{h\to 0} \frac{f(a+hv) - f(a)}{h} = \big((\partial_v f_i)(a)\big)_i .$$

La définition de différentiabilité des fonctions réelles a été vue à la section 5. $f : E \subseteq \mathbb{R}^m \to \mathbb{R}$ est différentiable en $x \in \mathbb{R}^m$ s'il existe $u \in \mathbb{R}^m$ tel que :

$$\lim_{x\to a} \frac{f(x) - f(a) - \langle u, x-a\rangle}{|x-a|} = 0.$$

Si un tel vecteur u existe, il faut que que ce vecteur u soit le gradient $(\nabla f)(a)$ de f. On peut dès lors déterminer les différentes dérivées directionnelles d'une fonction $f : E \subseteq \mathbb{R}^m \to \mathbb{R}$ en connaissant son gradient :

$$(\partial_v f)(a) = \big\langle (\nabla f)(a), v\big\rangle.$$

On extrapole cette formule dans le cas de fonctions à valeur vectorielle :

$$(\partial_v f)(a) = \big((\partial_v f_i)(a)\big)_i = \Big(\big\langle (\nabla f_i)(a), v\big\rangle\Big)_i .$$

Cette formule peut s'écrire de manière matricielle :

$$(\partial_v f_i)(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_m} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}.$$

Définition 8.2. La matrice $J_f(a) \coloneqq \left[\frac{\partial f_i}{\partial x_j}\right]$ est appelée la *matrice Jacobienne* de $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ au point a.

Remarque. La matrice jacobienne (ou le jacobien) est une généralisation du gradient : $J_f(a)$ est une matrice à n lignes si f est à valeurs dans $\mathbb{R}^n$. Donc si $n = 1$ (cas de fonction à valeur réelle), alors la matrice ne contient qu'une seule ligne et est donc le vecteur gradient.

8.2 Différentiabilité de fonctions à valeur vectorielle

Définition 8.3. Soient $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ et $a \in \operatorname{int} E$. Ont dit que $L : \mathbb{R}^m \to \mathbb{R}^n$ est une *approximation linéaire de* f *en* a si :

$$\lim_{x\to a} \frac{\big\|f(x) - f(a) - L(x-a)\big\|}{\|x-a\|} = 0.$$

Définition 8.4. Soient $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ et $a \in \operatorname{int} E$. On dit que f est *différentiable au point* a s'il existe une approximation linéaire de f en a. Cette application linéaire est appelé la *différentielle* de f en a et se note $(Df)(a)$ ou $(df)(a)$.

Lemme 8.5. *Soient* $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ *et* $a \in \operatorname{int} E$. *Soient* L_1, L_2 *deux approximations linéaires de* f *autour de* a. *Alors* $L_1 = L_2$.

Démonstration. On observe que la somme de deux applications linéaires reste une application linéaire. De plus, on trouve :

$$\begin{aligned}\frac{\left\|(L_1 - L_2)(x-a)\right\|}{\|x-a\|} &= \frac{\left\|(L_1 - L_2)(x-a) + f(a) - f(x) - f(a) + f(x)\right\|}{\|x-a\|} \\ &\leqslant \frac{\left\|-L_1(x-a) - f(a) + f(x)\right\|}{\|x-a\|} + \frac{\left\|f(x) - f(a) - L_2(x-a)\right\|}{\|x-a\|}.\end{aligned}$$

Par définition, en faisant tendre x vers a dans ces deux termes, on obtient 0. On a :

$$0 \leqslant \frac{\left\|(L_1 - L_2)(x-a)\right\|}{\|x-a\|} \leqslant \frac{\left\|-L_1(x-a) - f(a) + f(x)\right\|}{\|x-a\|} + \frac{\left\|f(x) - f(a) - L_2(x-a)\right\|}{\|x-a\|}.$$

Par le théorème du sandwich, on trouve que :

$$\lim_{x \to a} \frac{\left\|(L_1 - L_2)(x-a)\right\|}{\|x-a\|} = 0.$$

Soit $v \in \mathbb{R}^m$ tel que $v \neq 0$. On considère la suite $(x_k) \subset \mathbb{R}^m$ telle que $x_k = a + \frac{v}{k}$. Par hypothèse, $a \in \operatorname{int} E$. Dès lors, il existe un voisinage de a entièrement inclus dans E. On peut donc trouver K tel que $\forall k > K : x_k \in E$. En faisant tendre k vers $+\infty$, on retrouve la même limite que juste au-dessus :

$$\lim_{k \to +\infty} \frac{\left\|(L_1 - L_2)(x_k - a)\right\|}{\|x_k - a\|} = 0.$$

Or, par linéarité, on trouve :

$$0 = \lim_{k \to +\infty} \frac{\left\|(L_1 - L_2)(a + \frac{v}{k} - a)\right\|}{\left\|a + \frac{v}{k} - a\right\|} = \lim_{k \to +\infty} \frac{\frac{1}{k}}{\frac{1}{k}} \frac{\left\|(L_1 - L_2)(v)\right\|}{\|v\|}.$$

On trouve effectivement $0 = (L_1 - L_2)(v)$ pour tout v. ☐

Définition 8.6. Soient $f : \mathbb{R}^m \to \mathbb{R}^n$. On dit que f est un petit o de $\|h\|$ si :

$$\lim_{h \to 0} \frac{\left\|f(h)\right\|}{\|h\|} = 0.$$

Ce la se note $f(h) = o(\|h\|)$.

Théorème 8.7. *Soient* $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ *et* $a \in \operatorname{int} E$. *La fonction* f *est différentiable en* a *si et seulement si toutes les composantes* f_i *pour* $1 \leqslant i \leqslant n$ *sont différentiables en* a. *De plus, on détermine* $(df)(a) = J_f(a)$ *où* $J_f(a)$ *est la matrice jacobienne.*[5]

[5] Pour être entièrement rigoureux, on dit que $(df)(a)$ est *déterminée* par $J_f(a)$ car $(df)(a)$ est une application linéaire et toute application linéaire est sous la forme $v \mapsto av$ où a est une matrice. Dans un tel cas, on dit que la matrice a détermine (ou définit) l'*application linéaire associée*.

Démonstration. Supposons d'abord que f est différentiable en a. On sait alors qu'il existe L, une application linéaire telle que :

$$f(x) = f(a) + L(x-a) + o(\|x-a\|).$$

On note $U = [U_{ij}]$ la matrice associée à l'application linéaire L. Soit $1 \leqslant k \leqslant n$. Alors :

$$\begin{aligned}
\lim_{x\to a} \frac{\left|f_k(x) - f_k(a) - \langle U_{k\,\cdot}, x-a\rangle\right|}{\|x-a\|} &\leqslant \lim_{x\to a} \frac{1}{\|x-a\|}\sqrt{\sum_{i=1}^{n}\left|f_i(x) - f_i(a) - \langle U_{i\,\cdot}, x-a\rangle\right|^2} \\
&= \lim_{x\to a} \frac{1}{\|x-a\|}\left\|\left[f_i(x) - f_i(a) - \sum_{j=1}^{m}(U_{ij}(x_j - a_j))\right]_i\right\| \\
&= \lim_{x\to a} \frac{1}{\|x-a\|}\left\|[f_i(x)]_i - [f_i(a)]_i - \left[\sum_{j=1}^{m} U_{ij}(x_j - a_j)\right]_i\right\| \\
&= \lim_{x\to a} \frac{1}{\|x-a\|}\left\|f(x) - f(a) - L(x-a)\right\|.
\end{aligned}$$

Or ce dernier membre tend vers 0 par définition de la différentiabilité. Par le théorème du sandwich, en sachant que :

$$0 \leqslant \frac{\left|f_k(x) - f_k(a) - \langle U_{k\,\cdot}, x-a\rangle\right|}{\|x-a\|},$$

on trouve que :

$$\lim_{x\to a} \frac{\left|f_k(x) - f_k(a) - \langle U_{k\,\cdot}, x-a\rangle\right|}{\|x-a\|} = 0,$$

et donc la composante f_k est différentiable en a.

Supposons maintenant que f_k est différentiable pour tout $1 \leqslant k \leqslant n$ et montrons que f est différentiable. On sait que :

$$\forall 1 \leqslant i \leqslant n : \lim_{x\to a} \frac{f(x) - f(a) - \langle U_i, x-a\rangle}{\|x-a\|} = 0.$$

On définit ensuite $U = [U_{ij}]$ où $U_{ij} \coloneqq (U_i)_j$. Si L est l'application linéaire telle que $L(v) = Uv$, on trouve :

$$\begin{aligned}
\lim_{x\to a} \frac{\left\|f(x) - f(a) - L(x-a)\right\|^2}{\|x-a\|^2} &= \lim_{x\to a} \frac{1}{\|x-a\|^2}\sum_{i=1}^{n}\left|f_i(x) - f_i(a) - \langle U_{i\,\cdot}, x-a\rangle\right|^2 \\
&= \sum_{i=1}^{n}\lim_{x\to a}\left\|\frac{\left|f_i(x) - f_i(a) - \langle U_{i\,\cdot}, x-a\rangle\right|}{\|x-a\|}\right\|^2 \\
&= \sum_{i=1}^{n} 0^2 = 0.
\end{aligned}$$

Ce qui veut dire que f est différentiable.

La seconde partie du théorème est de dire que la différentielle de f est l'application associée à la matrice jacobienne $J_f(a)$. Afin de montrer cela, on a vu dans le cas $n = 1$ qu'il était obligatoire d'avoir $U_i = (\nabla f_i)(a)$. Par définition de U (mettre les vecteurs U_i en lignes pour en faire une matrice), on a bien :

$$U_{ij} = (U_i)_j = \big((\nabla f_i)(a)\big)_j = \frac{\partial f_i}{\partial x_j}(a).$$

□

Proposition 8.8. *Soient* $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ *et* $a \in \operatorname{int} E$. *Si* f *est différentiable en* a, *alors il existe* $C, R \in \mathbb{R}_0^+$ *tels que :*

$$\forall x \in B(a, r) : \left\|f(x) - f(a)\right\| \leqslant \|x - a\| .$$

En particulier, f *est continue en* a.

8.3 Condition suffisante de différentiabilité

Définition 8.9. Soit $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ où E est ouvert. Si les dérivées partielles de f existent en tout point $a \in E$ et qu'elles définissent des fonctions continues, on dit que f est de classe C^1.

Théorème 8.10. *Soient* $f : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$ *et* $a \in \operatorname{int} E$. *S'il existe* $\delta > 0$ *tel que :*

$$\forall x \in B(a, \delta), j \in \{1, \ldots, m\} : \frac{\partial f}{\partial x_j} : \operatorname{int} E \to \mathbb{R} \qquad \textit{existe et est continue en } a,$$

alors f *est différentiable en* a.

Démonstration. Considérons le cas où $n = 1$. On sait alors revenir au cas $n \gneqq 1$ par le théorème qui dit qu'une fonction f est différentiable si et seulement si ses composantes le sont.

Soient $a \in \operatorname{int} E, \epsilon > 0$. Par hypothèse, toutes les dérivées partielles sont continues en a. Donc, pour tout $1 \leqslant j \leqslant m$, il existe $\delta_j > 0$ (et $\delta_j < r$[6]) tel que pour tout $x \in B(a, \delta_j)$, on ait :

$$\left|\tfrac{\partial f}{\partial x_j}(x) - \tfrac{\partial f}{\partial x_j}(a)\right| \leqslant \frac{\epsilon}{m}.$$

On pose $\delta \coloneqq \min_{j=1,2,\ldots,m} \delta_j$. Ainsi, $\forall x \in B(a, \delta)$, l'inéquation précédente est vérifiée.

Soit $\beta \coloneqq \{e_1, \cdots, e_m\}$ la base canonique de $\mathbb{R}^m$. On pose la fonction :

$$h_1 : [0,1] \to \mathbb{R} : t \mapsto f(a + t(x-a)_1 e_1).$$

L'évaluation $h_1'(s)$ au point $s \in [0,1]$ représente la dérivée partielle de f par rapport à x_1 au point $a + s(x - a)_1 e_1$. Par le théorème de la moyenne, on sait qu'il existe $t_1 \in (0,1)$ tel que :

$$f(a + (x-a)_1 e_1) - f(a) = h_1(1) - h_1(0) = h_1'(t_1)(x-a)_1 = \frac{\partial}{\partial x_1} f(a + t(x-a)_1 e_1)(x-a)_1.$$

Par choix de δ, on sait que :

$$\left|\tfrac{\partial}{\partial x_1} f(a + t_1(x-a)_1 e_1) - \tfrac{\partial}{\partial x_1} f(a)\right| \leqslant \frac{\epsilon}{m}.$$

En multipliant de part et d'autre par $\left|(x-a)_1\right|$, on trouve :

$$\left|f(a + t(x-a)e_1) - f(a) - \tfrac{\partial}{\partial x_j} f(a)(x-a)_1\right| \leqslant \frac{\epsilon}{m}\left|(x-a)_1\right| \leqslant \frac{\epsilon}{m}|x-a| .$$

En reproduisant l'argument pour $h_2, \ldots, h_m$ et en appliquant à nouveau le théorème de la moyenne, on trouve les inégalités suivantes :

$$\forall i \in \{1, \ldots, m\} : \left| f\left(a + \sum_{j=1}^{i} (x-a)_j e_j\right) - f\left(a + \sum_{j=1}^{i-1} (x-a)_j e_j\right) - \tfrac{\partial f}{\partial x_i}(a) \right| \leqslant \frac{\epsilon}{m} \|x - a\|$$

[6] Cette hypothèse nous permet que les dérivées partielles soient définies sur $B(a, \delta_j) \subset B(a, r)$ sans souci.

En les sommant, on obtient :

$$\left|f(x)-f(a)-\sum_{j=1}^{m}\tfrac{\partial f}{\partial x_j}(a)(x-a)_j\right|=\left|f\left(a+\sum_{k=1}^{m}(x-a)_k e_k\right)-f(a)-\sum_{j=1}^{m}\tfrac{\partial f}{\partial x_k}(a)(x-a)_j\right|\leqslant\epsilon\|x-a\|\,,$$

ce qui prouve que la limite tend vers 0. La fonction f est donc différentiable en a. □

Remarque. Il est habituellement plus simple d'établir les dérivées partielles et leur continuité que de définir la différentiabilité sur base de la définition *brute*. C'est pour cela que le théorème 8.10 s'avère très utile dans beaucoup de situations.

8.4 Différentiation de fonction composées

On a vu que pour $f:\mathbb{R}\to\mathbb{R}^n$ et $g:\mathbb{R}^n\to\mathbb{R}$, la dérivée de la composée était donnée par :

$$(g\circ f)(a)=\big\langle(\nabla g)(f(a)),f'(a)\big\rangle.$$

Soient maintenant deux fonctions $f:\mathbb{R}^m\to\mathbb{R}^n$ et $g:\mathbb{R}^m\to\mathbb{R}^p$. Leur composée $(g\circ f):\mathbb{R}^m\to\mathbb{R}^p$ est bien définie. Supposons que f est différentiable en a et a $(df)(a)$ pour différentielle et que g est différentiable en $f(a)$ et a $(dg)(f(a))$ pour différentielle.

Théorème 8.11. *Soient* $f:A\subset\mathbb{R}^m\to\mathbb{R}^n$ *et* $g:B\subset\mathbb{R}^n\to\mathbb{R}^p$.*Si* $f(A)\subseteq B$, $a\in\operatorname{int}A$, $f(a)\in\operatorname{int}B$, *et si* f *et* g *sont différentiables respectivement en* a *et* $f(a)$, *alors* $(g\circ f)$ *est différentiable en* a *et sa différentielle est donnée par :*

$$(d(g\circ f))(a):\mathbb{R}^m\to\mathbb{R}^p:v\mapsto\big[(dg)(f(a))\circ(df)(a)\big]\,v=J_g(f(a))J_f(a)v,$$

où J_g *et* J_f *représentent respectivement la matrice jacobienne de* g *et de* f.

Démonstration. Étant donné que l'on peut réduire la différentiabilité d'une fonction vectorielle à la différentiabilité de ses composantes, concentrons-nous sur le cas $p=1$. On pose :

$$Q(x)\coloneqq\frac{(dg)(f(a))\big(f(x)-f(a)-(df)(a)(x-a)\big)}{\|x-a\|}=\sum_{k=1}^{n}\frac{\partial g}{\partial x_k}(f(a))\frac{f_k(x)-f_k(a)-(df_k)(a)(x-a)}{\|x-a\|}.$$

Par différentiabilité de f, on peut dire que f_k est différentiable en a pour tout k. Donc en passant à la limite, on obtient :

$$\lim_{x\to a}Q(x)=\sum_{k=1}^{n}\frac{\partial g}{\partial x_k}(f(a))\lim_{x\to a}\frac{f_k(x)-f_k(a)-(df_k)(a)(x-a)}{\|x-a\|}=0.$$

Si $f(x)=f(a)$, alors $P(x)=0$. Sinon, on pose :

$$R(x)\coloneqq\frac{g(f(x))-g(f(a))-(df)(a)(f(x)-f(a))}{\big\|f(x)-f(a)\big\|}$$
$$S(x)\coloneqq\frac{\big\|f(x)-f(a)\big\|}{\|x-a\|}.$$

On peut alors réécrire $P(x)=(RS)(x)$. Par la proposition 8.8, on sait que $f(x)\to f(a)$ et par différentiabilité de g, on sait que $R(x)\to 0$ pour $x\to a$. On sait donc également que $P(x)\to 0$ pour $x\to a$. De là, on trouve :

$$0=\lim_{x\to a}P(x)-Q(x)=\lim_{x\to a}\frac{g(f(x))-g(f(a))-(dg)(f(a))((df)(a)(x-a)}{\|x-a\|},$$

ou encore $(g\circ f)$ est différentiable en a et a $(dg)(f(a))\circ(df)(a)$ pour différentielle. □

8.5 Changement de variables

Théorème 8.12. *Soient* $U, V \subseteq \mathbb{R}^n$*. Soient* $f : U \to V$ *et* $g : V \to \mathbb{R}$ *deux fonction admettant des dérivées partielles continues. Alors la fonction définie par* $h : U \to \mathbb{R} : x \mapsto (g \circ f)(x)$ *admet des dérivées partielles définies par :*

$$\frac{\partial h}{\partial y_j}(a) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(f(a)) \frac{\partial f_i}{\partial y_j}(a).$$

Démonstration. Par le théorème 8.10, on sait que les fonctions f et g sont différentiables. On sait également que sa différentielle est donnée par :

$$(dh)(a) = (dg)(f(a)) \circ (df)(a).$$

On peut exprimer ses dérivées partielles (qui existent puisque h est différentiable) comme :

$$\frac{\partial h}{\partial y_j} = (dh)(a)(e_j),$$

où e_j est un vecteur de la base canonique de $\mathbb{R}^n$. On peut alors exprimer la différentielle de f par la différentielle de ses composantes :

$$(df)(a)(e_j) = \left[\tfrac{\partial f_i}{\partial x_j}(a)\right]_i \in \mathbb{R}^n.$$

Également, on peut écrire pour tout $u, b \in \mathbb{R}^n$:

$$(dg)(b)(u) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(b) u_i.$$

En posant $u \coloneqq (df)(a)(e_j)$ et $b \coloneqq f(a)$, cela revient à écrire :

$$(dg)(f(a))((df)(a)(e_j)) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(f(a)) \frac{\partial f_i}{\partial y_j}(a).$$

On remarque que :

$$(dg)(f(a))((df)(a)(e_j)) = \frac{\partial h}{\partial y_j}.$$

□

Remarque. En changeant ainsi les variables, on passe de x_k qui sont des variables indépendantes à une collection de fonctions f_k qui sont liées aux variables x_k.

9 Intégration de fonctions à plusieurs variables

9.1 Introduction

Remarque. Il existe deux théories principales de l'intégration : l'intégration de Riemann et l'intégration de Lebesgue. L'intégration de Lebesgue est intuitivement *meilleure* que l'intégration de Riemann au sens que si f est une fonction intégrable au sens de Riemann, alors elle l'est également au sens de Lebesgue et la valeur de l'intégrale concorde alors qu'il existe des fonctions[7] qui sont intégrables au sens de Lebesgue et qui ne le sont pas au sens de Riemann.

L'intégration de Riemann est cependant plus simple à définir et à comprendre, Lebesgue dépasse donc le cadre de ce cours.

9.2 Intégrale de Darboux sur un rectangle

Définition 9.1. On définit un *rectangle* R dans $\mathbb{R}^n$ par :

$$R = \prod_{i=1}^{n} [a_i, b_i] = \{x \in \mathbb{R}^n \text{ t. q. } \forall 1 \leqslant i \leqslant n : a_i \leqslant x_i \leqslant b_i\}.$$

Définition 9.2. Une partition d'un intervalle $I = [a, b] \subset \mathbb{R}$ est un choix de k nombres $t_i \in I$ tels que :

$$a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b.$$

Une partition rectangulaire d'un rectangle $R = \prod_k [a_k, b_k] \subset \mathbb{R}^n$ est une partition de chaque intervalle de R. Cela revient à choisir $t_{i\,j}$ tels que :

$$\begin{aligned}
a_1 &= t_{1\,0} < t_{1\,1} < \cdots < t_{1\,k_1-1} < t_{1\,k_1} = b_1 \\
a_2 &= t_{2\,0} < t_{2\,1} < \cdots < t_{2\,k_2-1} < t_{2\,k_2} = b_2 \\
&\vdots \\
a_n &= t_{n\,0} < t_{n\,1} < \cdots < t_{n\,k_n-1} < t_{n\,k_n} = b_n.
\end{aligned}$$

Définition 9.3. On définit R_j pour $1 \leqslant j \leqslant p$ (où p est le nombre de *sous-rectangles* de la partition $\mathcal{P}$) comme étant le jème rectangle. On définit ensuite es valeurs maximales et minimales atteintes par rectangle par :

$$\begin{aligned}
m_j &\coloneqq \inf\{f(x) \text{ t. q. } x \in R_j\} \\
M_j &\coloneqq \sup\{f(x) \text{ t. q. } x \in R_j\}.
\end{aligned}$$

Définition 9.4. Soit $f : R \subset \mathbb{R}^n \to \mathbb{R}$ une fonction définie sur un rectangle de $\mathbb{R}^n$, et $\mathcal{P}$ une partition rectangulaire de R. Les sommes :

$$\begin{aligned}
\underline{S}(f, \mathcal{P}) &\coloneqq \sum_j m_j \operatorname{vol}(R_j), \\
\overline{S}(f, \mathcal{P}) &\coloneqq \sum_j M_j \operatorname{vol}(R_j)
\end{aligned}$$

s'appellent respectivement la *somme inférieure* et la *somme supérieure* de Darboux de f par rapport à $\mathcal{P}$.

[7] Dont la fonction de Dirichlet définie par $f(x) = 1$ si $x \in \mathbb{Q}$ et $f(x) = 0$ si $x \notin \mathbb{Q}$.

Définition 9.5. On définit les intégrales respectivement *inférieure* et *supérieure* de Darboux de f par :

$$\underline{\int}_R f(x)\,dx \coloneqq \sup_{\mathcal{P}} \underline{S}(f,\mathcal{P}),$$
$$\overline{\int}_R f(x)\,dx \coloneqq \inf_{\mathcal{P}} \overline{S}(f,\mathcal{P}).$$

Remarque. Il suit directement de la définition que pour toute partition $\mathcal{P}$, on a :

$$\underline{S}(f,\mathcal{P}) \leqslant \overline{S}(f,\mathcal{P}),$$

et donc, en particulier :

$$\underline{\int}_R f(x)\,dx \leqslant \overline{\int}_R f(x)\,dx.$$

Définition 9.6. On dit que la fonction $f : R \subset \mathbb{R}^m \to \mathbb{R}$ est intégrable au sens de Riemann si les intégrales inférieure et supérieure de Darboux sont égales :

$$\underline{\int}_R f(x)\,dx = \overline{\int}_R f(x)\,dx.$$

Dans ce cas, on définit l'*intégrale de Riemann* de la fonction f par :

$$\int_R f(x)\,dx,$$

qui vaut cette valeur commune.

Définition 9.7. Soit $f : R \subset \mathbb{R}^m \to \mathbb{R}^n$. f est dite intégrable sur R si toutes ses composantes sont intégrables, et on définit son intégrale par le vecteur correspondant aux intégrales des composantes :

$$\int_R f(x)\,dx = \left[\int_R f_i(x)\,dx\right]_i \in \mathbb{R}^n.$$

Lemme 9.8. *Soit $f : R \to \mathbb{R}$ bornée. Si pour tout $\epsilon > 0$, il existe $\mathcal{P}(\epsilon)$ une partition rectangulaire de R telle que :*

$$\overline{S}(f,\mathcal{P}(\epsilon)) - \underline{S}(f,\mathcal{P}(\epsilon)) \leqslant \epsilon,$$

alors f est intégrable au sens de Riemann.

Remarque. À nouveau, et pour la plupart des résultats qui suivront, le fait que la définition d'une intégrale d'une fonction à valeur vectorielle est définie par le vecteur dont les composantes sont les intégrales des composantes implique que les résultats basés sur une fonction $f : R \subset \mathbb{R}^m \to \mathbb{R}$ peuvent être étendus aux fonctions $f : R \subset \mathbb{R}^m \to \mathbb{R}^n$.

Démonstration. Supposons par l'absurde que le lemme soit faux et donc que sous ces hypothèses, il existe f non intégrable au sens de Riemann[8]. On peut alors prendre :

$$\epsilon = \inf_{\mathcal{P}} \overline{S}(f,\mathcal{P}) - \sup_{\mathcal{P}} S(f,\mathcal{P}) \gneqq 0.$$

On sait par les hypothèses qu'il existe $\mathcal{P}(\epsilon)$ une partition de R telle que :

$$\overline{S}(f,\mathcal{P}(\epsilon)) \leqslant \underline{S}(f,\mathcal{P}(\epsilon)) + \epsilon = \underline{S}(f,\mathcal{P}(\epsilon)) + \inf_{\mathcal{P}} \overline{S}(f,\mathcal{P}) - \sup_{\mathcal{P}} \underline{S}(f,\mathcal{P}).$$

[8] À compter d'ici, le terme *intégrable* fera toujours référence à la notion d'*intégrable au sens de Riemann*.

De plus, par définition, on sait que :

$$\inf_{\mathcal{P}} \overline{S}(f, \mathcal{P}) \leqslant \overline{S}(f, \mathcal{P}(\epsilon)) \lneqq \underline{S}(f, \mathcal{P}(\epsilon)) + \inf_{\mathcal{P}} \overline{S}(f, \mathcal{P}) - \sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}).$$

En soustrayant $\inf_{\mathcal{P}}$ de part et d'autre, on obtient :

$$\sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) < \underline{S}(f, \mathcal{P}(\epsilon)),$$

ce qui est une contradiction, par définition du supremum. □

Proposition 9.9. *Si* $R \subset \mathbb{R}^m$ *est un rectangle et* $f : E \to \mathbb{R}$ *est continue, alors* f *est intégrable.*

Démonstration. Puisque E est un rectangle, il est fermé borné, et donc f est bornée également (par sa continuité). f est également uniformément continue (puisque continue sur un ensemble fermé borné). Dès lors, pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que :

$$\forall x, x' \in R : \left\|x - x'\right\| < \delta \Rightarrow \left|f(x) - f(x')\right| < \frac{\epsilon}{\text{vol}(R)}.$$

En choisissant $\mathcal{P}$ de manière à ce que chaque R_j soit contenu dans une boule ouverte de diamètre δ. On trouve donc, pour tout R_j dans $\mathcal{P}$:

$$M_j(f) - m_j(f) = \sup_{R_j} f(x) - \inf_{R_j} f(x) < \frac{\epsilon}{\text{vol}(R)}.$$

On a donc construit une partition $\mathcal{P}$ telle que :

$$\overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) = \sum_j \big(M_j(f) - m_j(f)\big) \,\text{vol}(R) < \frac{\epsilon}{\text{vol}(R)} \,\text{vol}(R) = \epsilon.$$

Dès lors, par le lemme 9.8, sait que f est intégrable. □

Proposition 9.10 (Propriétés des intégrales)**.** 1. *Soient* $f, g : R \subset \mathbb{R}^m \to \mathbb{R}$ *intégrables et* $\alpha, \eta \in \mathbb{R}$. *Alors* $(\alpha f + \beta g)$ *est intégrable telle que :*

$$\int_R (\alpha f + \beta g)(x)\, dx = \alpha \int_R f(x)\, dx + \beta \int_R g(x)\, dx.$$

2. *Soient* $f, g : R \subset \mathbb{R}^m \to \mathbb{R}$ *intégrables telles que pour tout* $x \in R$, *on a* $f(x) \geqslant g(x)$, *alors :*

$$\int_R f(x)\, dx \geqslant \int_R g(x)\, dx.$$

3. *Soit* $f : R \subset \mathbb{R}^m \to \mathbb{R}$ *intégrable. Alors* $\left|f(x)\right|$ *est également intégrable telle que :*

$$\int_R \left|f(x)\right| dx \geqslant \left|\int_R f(x)\, dx\right|.$$

4. *Soit* $R \subset \mathbb{R}^m$ *un rectangle et soient* P, S *tels que* $R = P \cup S$ *où* P *et* S *sont également des rectangles n'ayant pour seule intersection éventuelle leurs bords. Soit* $f : R \to \mathbb{R}$. *Alors* f *est intégrable, si et seulement si* f *réduit à* P *et* f *réduit à* S *sont intégrables. De plus :*

$$\int_R f(x)\, dx = \int_P f(x)\, dx + \int_S f(x)\, dx.$$

Démonstration. Les démonstrations sont similaires à celles concernant les fonctions univariées. $\square$

Définition 9.11. Soit E un ensemble quelconque (discret, infini dénombrable, infini indénombrable, etc.) Alors on définit la *fonction caractéristique* de E[9] par :

$$1_E : x \mapsto \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{sinon} \end{cases}.$$

Remarque. Cette fonction indicatrice permet, entre autres, de définir une intégrale sur un ensemble qui n'est pas un rectangle. En effet, soit S un sous-ensemble de $\mathbb{R}^m$. Si S n'est pas un rectangle, o peut déterminer un rectangle R tel que $S \subset R$. Dans ce cas, on intègre la fonction $(1_S f)(x)$ sur le rectangle S. La fonction s'annulant sur tous les points de $R \setminus S$ c'est bien l'aire de f sur S qui est obtenue.

Attention tout de même : la fonction $(1_S f)(x)$ n'est (à priori) pas continue.

9.3 Critère de Lebesgue

Définition 9.12. Soit $N \subset \mathbb{R}^m$. On dit que N est de *mesure nulle* si pour tout $\epsilon > 0$, il existe une collection dénombrable de rectangles R_j dans $\mathbb{R}^m$ telle que :

(i) $N \subset \displaystyle\bigcup_{j=1}^{+\infty} R_j$,

(ii) $\displaystyle\lim_{k \to +\infty} \sum_{j=0}^{k} \text{vol}(R_j) \leqslant \epsilon$.

Remarque. Pour le point (ii), c'est une série qui doit être inférieure à ϵ. Les séries seront détaillées dans la section 10. L'idée derrière est le fait qu'une somme d'un nombre fini de termes est bien définie, mais une somme infinie ne l'est pas directement. Dès lors, on passe à la limite sur $k \to +\infty$.
Remarque. Un ensemble fini ou dénombrable est de mesure nulle.

Lemme 9.13. *Soit $f : E \subset \mathbb{R}^{m-1} \to \mathbb{R}$ une fonction continue où E est un ensemble compact. Alors le graphe de f est de mesure nulle.*

Démonstration. Soit R un rectangle tel que $\text{dom } f \subseteq R$. Soit $\epsilon > 0$. Par continuité uniforme (continuité d'une fonction bornée sur un ensemble compact), on sait qu'il existe $\delta > 0$ tel que :

$$\forall x, x' \in \text{dom } f : \|x - x'\| < \delta \Rightarrow |f(x) - f(x')| < \frac{\epsilon}{\text{vol}(R)}.$$

Soit ensuite $\mathcal{P}$ une partition rectangulaire de R telle que chaque rectangle R_j a un diamètre[10] $\lneqq \delta$. Pour chaque rectangle R_j, on construit $R'_j \coloneqq R_j \times [m_j, M_j]$. Chacun de ces sous-rectangles a un volume donné par :

$$\text{vol}(R'_j) = \text{vol}(R_j)(M_j - m_j) \lneqq \text{vol}(R_j)\frac{\epsilon}{\text{vol}(R)}.$$

En sommant le volume de tous ces rectangles, on trouve un volume inférieur à ϵ. $\square$

[9] Également appelée *fonction* indicatrice de l'ensemble E.
[10] Le diamètre d'un rectangle correspond à la longueur de ses diagonales.

Théorème 9.14 (Critère de Lebesgue)**.** *Soit* $f : R \subset \mathbb{R}^m \to \mathbb{R}$ *une fonction définie sur un rectangle. Soit :*

$$D := \{p \in R \text{ t. q. } f \textit{ n'est pas continue en } p\}.$$

Alors f *est intégrable si et seulement si* f *est bornée et* D *est de mesure nulle.*

Remarque. Ce théorème n'est pas démontré ici.

Corollaire 9.15. *Soit* $f : R \to \mathbb{R}_0^+$ *intégrable et positive définie sur un rectangle . Si* $\int_R f(x)\,dx = 0$*, alors l'ensemble :*

$$N := \{x \in R \text{ t. q. } f(x) \gneqq 0\}$$

est de mesure nulle. De plus, si f *est continue, alors* $N = \emptyset$ *et* f *est la fonctions constante nulle.*

Démonstration. Soit D l'ensemble des points de R où f n'est pas continue. Par Lebesgue, on sait que D est de mesure nulle. Montrons que $N \subseteq D$. Supposons par l'absurde qu'il existe $p \in N$ tel que f est continue en p. Par définition de N, on sait que $f(p) > 0$. Par la continuité de f en p, on sait qu'il existe un rectangle $R' \subset R$ tel que $p \in R'$ et pour tout $x \in R'$, on a :

$$f(x) > \frac{f(p)}{2}.$$

Puisque $f(x)$ est définie positive, on sait que pour tout x : $(f)(x) \geqslant (1_{R'}f)(x)$. Dès lors :

$$\int_R f(x)\,dx \geqslant \int_R (1_{R'}f)(x)\,dx = \int_{R'} f(x)\,dx \geqslant \operatorname{vol}(R')\frac{f(p)}{2} \gneqq 0.$$

Ce qui est une contradiction. Dès lors, on sait $N \subseteq D$. Or un sous-ensemble d'un ensemble dénombrable est dénombrable et donc N est de mesure nulle. Également, si f est continue, il est évident que $D = \emptyset$, et donc forcément $N = D$. □

9.4 Théorème de Fubini

Théorème 9.16 (Théorème de Fubini simplifié dans $\mathbb{R}^2$)**.** *Soit* $f : [a,b] \times [c,d] \to \mathbb{R}$ *une fonction continue définie sur un rectangle de* $\mathbb{R}^2$*. Les fonctions* $F : [a,b] \to \mathbb{R}$ *et* $G : [c,d] \to \mathbb{R}$ *définies par :*

$$\begin{cases} F(s) &= \int_c^d f(s,t)\,dt \\ G(t) &= \int_a^b f(s,t)\,ds \end{cases}$$

sont continues. De plus, on a :

$$\int_c^d F(s)\,ds = \int_a^b G(t)\,dt = \int_R f(x)\,dx.$$

Démonstration. On remarque que les fonctions F et G sont bien définies de par la continuité de f. De plus, comme R est fermé borné, la fonction f est uniformément continue sur R. Dès lors, en posant $\epsilon > 0$, on sait qu'il existe $\delta > 0$ tel que pour tout $(s_1,t_1), s_2, t_2$, si $\left|(s_1,t_1) - (s_2,t_2)\right| < \delta$, alors $\left|f(s_1,t_1) - f(s_2,t_2)\right| < \frac{\epsilon}{d-c}$. Dès lors, soient s_1, s_2, en supposant$|s_1 - s_2| < \delta$, alors on trouve :

$$\left|F(s_1) - F(s_2)\right| = \left|\int_c^d \big(f(s_1,t) - f(s_2,t)\big)\,dt\right| \leqslant \int_c^d \left|f(s_1,t) - f(s_2,t)\right|dt = \int_c^d \frac{\epsilon}{d-c} = \epsilon.$$

On observe donc que F est également uniformément continue. On détermine similairement que G est uniformément continue.

Soit $\mathcal{P}$ une partition rectangulaire de R. Par définition, on peut écrire :

$$\begin{cases} \mathcal{P}_1 & : a = s_0 < s_1 < \cdots < s_{p-1} < s_p = b \\ \mathcal{P}_2 & : c = t_0 < t_1 < \cdots < t_{q-1} < t_q = d. \end{cases}$$

On note alors $\mathcal{P}_{ij}$ le rectangle $[s_{i-1}, s_i] \times [t_{j-1}, t_j]$. Les sommes de Darboux peuvent s'exprimer ainsi :

$$\begin{aligned} \underline{S}(f, \mathcal{P}) &= \sum_{i=1}^{p} \sum_{j=1}^{q} \left|[s_{i-1}, s_i]\right|\left|[t_{j-1}, t_j]\right| \inf_{(s,t)\in\mathcal{P}_{ij}} f(s,t) = \sum_{i=1}^{p} \sum_{j=1}^{q} (s_i - s_{i-1})(t_j - t_{j-1}) \inf_{s\in[s_{i-1},s_i]} \left(\inf_{t\in[t_{j-1},t_j]} f(s,t) \right) \\ &= \sum_{i=1}^{p} \left((s_i - s_{i-1}) \inf_{s\in[s_{i-1},s_i]} \left[\sum_{j=1}^{q} (t_j - t_{j-1}) \inf_{t\in[t_{j-1},t_j]} f(s,t) \right] \right). \end{aligned}$$

On remarque cependant que :

$$\sum_{j=1}^{q} (t_j - t_{j-1}) \inf_{t\in[t_{j-1},t_j]} f(s,t) = \underline{S}(t \mapsto f(s,t), \mathcal{P}_2) \leqslant \int_c^d f(s,t)\,dt = F(s).$$

On peut alors réécrire :

$$\underline{S}(f, \mathcal{P}) \leqslant \sum_{i=1}^{p} (s_i - s_{i-1}) \inf_{s\in[s_{i-1},s_i]} F(s) = \underline{S}(F, \mathcal{P}_1) \leqslant \int_a^b F(s)\,ds.$$

De manière sensiblement équivalente, on montre que :

$$\overline{S}(f, \mathcal{P}) \geqslant \int_a^b F(s)\,ds.$$

On sait cependant que f est intégrable et donc que :

$$\int_R f(x)\,dx = \int_a^b F(s)\,ds.$$

On prouve de la même manière que $\int_R f(x)\,dx = \int_c^d G(t)\,dt$.Ce qui conclut la preuve du théorème. □

Théorème 9.17 (Théorème de Fubini)**.** *Soient* $R_1 \subset \mathbb{R}^{m_1}, R_2 \subset \mathbb{R}^{m_2}$ *et* $f : R_1 \times R_2 \to \mathbb{R}$ *une fonction intégrable. Alors les fonctions suivantes sont intégrables également :*

$$\begin{aligned} R_1 \ni x &\mapsto \underline{\int}_{R_2} f(x,y)\,dy, \\ R_1 \ni x &\mapsto \overline{\int}_{R_2} f(x,y)\,dy, \\ R_2 \ni y &\mapsto \underline{\int}_{R_1} f(x,y)\,dx, \\ R_2 \ni y &\mapsto \overline{\int}_{R_1} f(x,y)\,dx. \end{aligned}$$

De plus :

$$\begin{aligned}\int_{R_1\times R_2} f(x)\,dx &= \int_{R_1}\left(\underline{\int}_{R_2} f(x,y)\,dy\right)dx\\ &= \int_{R_1}\left(\overline{\int}_{R_2} f(x,y)\,dy\right)dx\\ &= \int_{R_2}\left(\underline{\int}_{R_1} f(x,y)\,dx\right)dy\\ &= \int_{R_2}\left(\overline{\int}_{R_1} f(x,y)\,dx\right)dy.\end{aligned}$$

Remarque. Lorsque la fonction $x \mapsto f(x,y)$ est intégrable pour tout $y \in R_2$, alors :

$$\int_{R_1} f(x,y)\,dx = \underline{\int}_{R_1} f(x,y)\,dx = \overline{\int}_{R_1} f(x,y)\,dx.$$

Dans ce cas, le théorème de Fubini dit que :

$$\int_{R_1\times R_2} f(x,y)\,d(x,y) = \int_{R_1}\left(\int_{R_2} f(x,y)\,dy\right)dx.$$

De même, si la fonction $y \mapsto f(x,y)$ est intégrable pour tout $x \in R_1$, alors Fubini dit que :

$$\int_{R_1\times R_2} f(x,y)\,d(x,y) = \int_{R_2}\left(\int_{R_1} f(x,y)\,dx\right)dy.$$

Corollaire 9.18. *Soit* $f : R_1 \times R_2 \to \mathbb{R}$ *intégrable. Alors l'ensemble :*

$$N := \{y \in R_2 \text{ t. q. } x \mapsto f(x,y) \textit{ n'est pas intégrable }\}$$

est de mesure nulle.

Démonstration. Par définition, pour tout $y \in R_2$, n peut définir :

$$\Delta(y) := \overline{\int}_{R_1} f(x,y)\,dx - \underline{\int}_{R_1} f(x,y)\,dx \geqslant 0.$$

De plus, $y \in N$ si et seulement si $\Delta(y) \gneqq 0$. Or Fubini dit que $\Delta(y)$ est intégrable et que $\int_{R_2} \Delta(y)\,dy = 0$. On peut donc dire que $\Delta(y) = 0$ par le corollaire 9.15 sauf quand $y \in E$ tel que E est de mesure nulle. □

9.5 Changements de variables

Définition 9.19. Soit M une matrice de dimension $n \times m$. On note, pour tout $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$, $M(i,j)$ pour faire référence à la matrice M à laquelle la ième ligne et la jème colonne ont été retirées.

Remarque. Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une fonction de classe C^1. On prend $a, \epsilon \in \mathbb{R}^n$, et on définit R_ϵ le rectangle :

$$R_\epsilon \coloneqq \prod_{k=1}^{n} [a_k, a_k + \epsilon_k].$$

Rien ne garantit que $f(R_\epsilon)$ soit toujours un rectangle, cependant, pour ϵ *tout petit*, on peut utiliser $(df)(a)$, la différentielle de f en a pour approximer (linéairement) $f(R_\epsilon)$. La différentielle envoie R_ϵ, un rectangle sur $f(R_\epsilon)$, un parallélogramme. On peut définir ce parallélogramme par P engendré par les vecteurs :

$$f(a) + \frac{\partial f}{\partial x_1}(a)\epsilon_1,\ f(a) + \frac{\partial f}{\partial x_2}(a)\epsilon_2,\ \ldots,\ f(a) + \frac{\partial f}{\partial x_n}(a)\epsilon_n.$$

Il s'ensuit que le volume de ce parallélogramme est donné par le déterminant de la matrice dont les vecteurs de P sont les vecteurs de P :

$$\operatorname{vol}(f(R_\epsilon)) = \operatorname{vol}(P) = \left|\det J_f(a)\right| \prod_{k=1}^{n} \epsilon_k = \left|\det J_f(a)\right| \operatorname{vol}(R_\epsilon).$$

On en déduit que $\left|\det J_f(a)\right|$ définit la variation infinitésimale de volume en a quand les variables sont changées par la fonction f.

Définition 9.20. Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ de classe C^1 définie sur U un ouvert. On dit que f est un changement de variables si $\det J_f(a) \neq 0$ pour tout $a \in U$.

Théorème 9.21 (Formule du changement de variables)**.** *Soit* $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *définie sur* U *un ouvert. Si* $g : f(U) \to \mathbb{R}^n$ *est une fonction intégrable, alors la fonction* $(g \circ f)\left|\det J_f(a)\right| : U \to \mathbb{R}^n$ *est intégrable et :*

$$\int_V g(x)\, dx = \int_U g(f(x))\left|\det J_f(x)\right| dx.$$

Remarque. Ce théorème n'est pas démontré.

9.6 Surfaces dans $\mathbb{R}^3$

Définition 9.22. Soit $f : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ injective de classe C^1 où U est un ouvert. On dit que $f(U)$ est une *surface paramétrisable* si pour tout $p \in U$: $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont linéairement indépendants. On dit que l'application f est une *paramétrisation* de S.

Remarque. Si $V \subset \mathbb{R}^2$ est un ouvert et si $\phi : V \to U$ est une bijection de classe C^1 telle que pour tout $q \in V$: $(d\phi)(q) : U \to V$ est un isomorphisme, alors $\hat{f} : V \to \mathbb{R}^3 : x \mapsto (f \circ \phi)(x)$ a la même image que f et de plus, on sait :

$$(d\hat{f})(q) = (df)(\phi(q))(d\phi)(q).$$

De plus, puisque $(d\phi)(q)$ est un isomorphisme et $(df)(\phi(q))$ est injective, $(d\hat{f})(q)$ est également injective. Dès lors, $\hat{f}$ est également une paramétrisation de S.

Définition 9.23. Une *surface paramétrisable* est un sous-ensemble $S \subset \mathbb{R}^3$ qui est localement paramétrisable. C'est-à-dire que pour tout $p \in S$, il existe $V \subset \mathbb{R}^3$ un ouvert contenant p tel que $V \cap S$ est paramétrisable.

S'il existe un ensemble de paramétrisations telles que toutes les paramétrisations sont localement injectives, alors la surface S est dite *plongée*.

9.6.1 Surfaces orientables

Définition 9.24. Soient $u, v \in \mathbb{R}^3$. On définit le produit vectoriel :

$$u \times v := \big(u_2v_3 - u_3v_2, -(u_1v_3 - u_3v_1), u_1v_2 - u_2v_1\big) .$$

Remarque. On peut définir le produit vectoriel par le déterminant de la matrice :

$$\begin{pmatrix} e_x & e_y & e_z \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix}.$$

Lemme 9.25. *La norme du produit vectoriel de* $u, v \in \mathbb{R}^3$ *respecte l'égalité suivante :*

$$\|u \times v\| = \|u\|\|v\| \sin\theta,$$

où θ *est l'angle formé par les deux vecteurs.*

Remarque. En supposant $u \neq 0 \neq v$, on trouve que $u \times v = 0$ si et seulement si u et v sont colinéaires. Le produit vectoriel donne donc un moyen de tester l'indépendance linéaire entre deux vecteurs.

Définition 9.26. Soit S une surface paramétrée par $f : U \to \mathbb{R}^3$. On définit le *vecteur normal unitaire positif à* S par :

$$n : U \to \mathbb{R}^3 : p \mapsto \frac{\frac{\partial f}{\partial x}(p) \times \frac{\partial f}{\partial y}(p)}{\left\|\frac{\partial f}{\partial x}(p) \times \frac{\partial f}{\partial y}(p)\right\|}.$$

Remarque. Par définition d'une surface paramétrée, les vecteurs $\frac{\partial f}{\partial x}(p)$ et $\frac{\partial f}{\partial y}(p)$ sont linéairement indépendants. Dès lors diviser la norme du produit scalaire a du sens car cette norme est strictement positive.

Le produit scalaire donne un vecteur orthogonal aux vecteurs u et v. Donc par définition, $n(p)$ donne un vecteur (unitaire) orthogonal aux dérivées partielles $\frac{\partial f}{\partial x}(p)$ et $\frac{\partial f}{\partial y}(p)$, et donc par extension, $n(p)$ est également orthogonal à l'image de $(df)(p) : \mathbb{R}^2 \to \mathbb{R}^3$.

De plus, il existe deux vecteurs unitaires orthogonaux au plan tangent à $f(p)$, à savoir $n(p)$ et $-n(p)$. En choisissant $\hat{f}$, une autre paramétrisation de S, on aura toujours un de ces deux vecteurs.

Définition 9.27. Soient $f, \hat{f}$ deux paramétrisations d'une même surface S paramétrisable. On note respectivement n et $\hat{n}$ les vecteurs unitaires normaux. On dit que f et $\hat{f}$ sont de même orientation si $n = \hat{n}$, et on dit que f et $\hat{f}$ sont d'orientation opposée si $n = \hat{n}$.

Remarque. Le nombre d'orientations possibles étant réduit à 2, on peut subdiviser toutes le paramétrisations f d'une même surface paramétrisable S en 2 classes (selon le vecteur normal unitaire).

Définition 9.28. Une surface S est dite orientable si on peut choisir toutes les paramétrisations locales telles qu'elles induisent toutes le même vecteur normal (donc de même orientation) en tout point où leur image se rencontre.

Remarque. Sur une surface orientable, on peut dès lors définir le champ normal $S \to \mathbb{R}^3$ de manière globale sur la surface. Une *orientation* est alors un choix explicite d'une des deux possibilités du champ.

Il existe des surfaces non-orientables (c.f. le ruban de Möbius ou la bouteille de Klein).

9.7 Aire d'une surface et intégrale sur une surface

Définition 9.29. Soit S une surface paramétrée par $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}^3$. L'aire de S est définie par :

$$A(S) := \int_S \left\| \frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y} \right\| d(x, y).$$

Remarque. L'idée derrière cette définition est de définir l'aire d'une surface par la valeur numérique d'un volume de hauteur 1 et de base S.

S'il faut plus d'une paramétrisation locale de la courbe pour définir l'aire, on prend une paramétrisation qui couvre la surface S en y enlevant un nombre fini de courbes $C_1, \dots, C_n$. On appelle cette surface $\hat{S}$. L'aire d'une courbe étant nulle il est intuitivement vrai que S et $\hat{S}$ ont la même aire.

Définition 9.30. Soit $g : \mathbb{R}^3 \to \mathbb{R}$ une fonction continue et soit $f : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ une paramétrisation de la surface S. L'intégrale de g sur S est définie par :

$$\int_S g \, dA := \int_U g(f(x, y)) \left\| \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \right\| d(x, y).$$

Remarque. À nouveau, s'il faut plus d'une paramétrisation locale, on définit $\widehat{S} = S \setminus (\bigcup_{i-1}^{n} C_i)$ qui est la surface S à laquelle un nombre fini de courbes a été retiré. On définit alors :

$$\int_S g \, dA = \int_{\hat{S}} g \, dA.$$

9.7.1 Flux d'un champ de vecteurs au travers d'une surface

Si F est un champ de vecteur (à savoir si $F : \mathbb{R}^3 \to \mathbb{R}^3$), alors on peut définir la notion de *flux* de vecteur à travers une surface S.

Définition 9.31. Soient $f : U \subset \mathbb{R}^2$ une paramétrisation de la surface S définie sur un ouvert U et $F : \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteur continu. Le *flux de* F *au travers de* S est défini par :

$$\int_S \langle F, dn \rangle := \left\langle F, \frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y} \right\rangle d(x, y).$$

Dans ce contexte, dn représente l'élément normal à la surface S est *vaut*[11] $(\frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y}) \, dx \, dy$.

L'orientation de la paramétrisation a beaucoup d'importance : si $\hat{f}$ est d'orientation opposée à f, alors l'intégrale revient à *parcourir la surface dans l'autre sens*. Dès lors, la notion de flux n'a de sens que pour les surfaces orientées.

9.8 Théorèmes fondamentaux

9.8.1 Divergence et rotationnel d'un champ

Définition 9.32. Soit $F : U \subset \mathbb{R}^3 \to \mathbb{R}^3 \in C^1$ où U est un ouvert.

[11] Gardons en tête que cette notation n'a un sens - jusqu'à présent et dans ce cours - que dans une intégrale

On définit la *divergence de* F parla fonction :

$$\langle \nabla, F \rangle : U \to \mathbb{R} : (x,y,z) \mapsto \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \right)(x,y,z).$$

La divergence se note également parfois $\operatorname{div} F$.

On définit également le *rotationnel* de F par :

$$\nabla \times F : U \to \mathbb{R}^3 : (x,y,z) \mapsto \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right)(x,y,z).$$

Le rotationnel se note également parfois $\operatorname{curl} F$ ou $\operatorname{rot} F$.

On pose les notations suivantes pour F, G deux champs de vecteurs et f une fonction quelconques (tous les trois de classe C^1) :

$$\begin{aligned} \langle F, \nabla \rangle (f) &\coloneqq F_1 \frac{\partial f}{\partial x} + F_2 \frac{\partial f}{\partial y} + F_3 \frac{\partial f}{\partial z}, \\ \langle F, \nabla \rangle (G) &\coloneqq \left(\langle F, \nabla \rangle (G_i) \right)_i . \end{aligned}$$

Lemme 9.33. *Soient* ϕ, ψ, *deux fonctions et* F, G *deux champs de vecteurs, tous de classe* C^1. *Alors :*

1. $\nabla(\phi\psi) = \phi\nabla\psi + \psi\nabla\phi$;
2. $\langle \nabla, \phi F \rangle = \phi \langle \nabla, F \rangle + \langle \nabla\phi, F \rangle$;
3. $\nabla \times \phi F = \phi(\nabla \times F) - F \times (\nabla\phi)$;
4. $\nabla \langle F, G \rangle = F \times (\nabla \times G) + G \times (\nabla \times F) + \langle F, \nabla \rangle (G) + \langle G, \nabla \rangle (F)$;
5. $\langle \nabla, F \times G \rangle = \langle G, \nabla \times F \rangle - \langle F, \nabla \times G \rangle$;
6. $\nabla \times (F \times G) = F \langle \nabla, G \rangle - G \langle \nabla, F \rangle + \langle G, \nabla \rangle (F) - \langle F, \nabla \rangle (G)$.

Démonstration. Seuls les points 1, 2 , 3, et 5 sont démontrés ici.

Point 1 :

$$\begin{aligned} \nabla(\phi\psi) &= \left(\frac{\partial(\phi\psi)}{\partial x}, \frac{\partial(\phi\psi)}{\partial y}, \frac{\partial(\phi\psi)}{\partial z} \right) = \left(\frac{\partial\phi}{\partial x}\psi + \phi\frac{\partial\psi}{\partial x}, \frac{\partial\phi}{\partial y}\psi + \phi\frac{\partial\psi}{\partial y}, \frac{\partial\phi}{\partial z}\psi + \phi\frac{\partial\psi}{\partial z} \right) \\ &= \psi \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z} \right) + \phi \left(\frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y}, \frac{\partial\psi}{\partial z} \right) = \psi\nabla\phi + \phi\nabla\psi. \end{aligned}$$

Point 2 :

$$\begin{aligned} \langle \nabla, \phi F \rangle &= \frac{\partial(\phi F)_1}{\partial x} + \frac{\partial(\phi F)_2}{\partial y} + \frac{\partial(\phi F)_3}{\partial z} = \frac{\partial\phi}{\partial x}F_1 + \phi\frac{\partial F_1}{\partial x} + \frac{\partial\phi}{\partial y}F_2 + \phi\frac{\partial F_2}{\partial y} + \frac{\partial\phi}{\partial z} + \phi\frac{\partial F_3}{\partial z} \\ &= \langle F, \nabla \rangle(\phi) + \langle \nabla\phi, F \rangle. \end{aligned}$$

Point 3 :

$$\begin{aligned} \nabla \times \phi F &= \left(\frac{\partial(\phi F)_3}{\partial y} - \frac{\partial(\phi F)_2}{\partial y}, \frac{\partial(\phi F)_1}{\partial z} - \frac{\partial(\phi F)_3}{\partial x}, \frac{\partial(\phi F)_2}{\partial x} - \frac{\partial(\phi F)_1}{\partial y} \right) \\ &= \left(\frac{\partial\phi}{\partial y}F_3 + \phi\frac{\partial F_3}{\partial y} - \frac{\partial\phi}{\partial z}F_2 - \phi\frac{\partial F_2}{\partial z}, \frac{\partial\phi}{\partial z}F_1 + \phi\frac{\partial F_1}{\partial z} - \frac{\partial\phi}{\partial x}F_3 - \phi\frac{\partial F_3}{\partial x}, \frac{\partial\phi}{\partial x}F_2 + \phi\frac{\partial F_2}{\partial x} - \frac{\partial\phi}{\partial y}F_1 - \phi\frac{\partial F_1}{\partial y} \right) \\ &= (\nabla\phi \times F) + \phi(\nabla \times F). \end{aligned}$$

Point 5 :

$$\begin{aligned}
\langle \nabla, F \times G \rangle &= \tfrac{\partial (F\times G)_1}{\partial x} + \tfrac{\partial (f\times G)_2}{\partial y} + \tfrac{\partial (F\times G)_3}{\partial z} = \tfrac{\partial (F_2G_3 - F_3G_2)}{\partial x} + \tfrac{\partial (F_3G_1 - F_1G_3)}{\partial y} + \tfrac{\partial (F_1G_2 - F_2G_1)}{\partial z} \\
&= G_1(\tfrac{\partial F_3}{\partial y} - \tfrac{\partial F_2}{\partial z}) + G_2(\tfrac{\partial F_1}{\partial z} - \tfrac{\partial F_3}{\partial x}) + G_3(\tfrac{\partial F_2}{\partial x} - \tfrac{\partial F_1}{\partial y}) + F_1(\tfrac{\partial G_2}{\partial z} - \tfrac{\partial G_3}{\partial y}) + F_2(\tfrac{\partial G_3}{\partial x} - \tfrac{\partial G_1}{\partial z}) + F_3(\tfrac{\partial G_1}{\partial y} - \tfrac{\partial G_2}{\partial x}) \\
&= G_1(\nabla \times F)_1 + G_2(\nabla \times F)_2 + G_3(\nabla \times F)_3 - F_1(\nabla \times G)_1 - F_2(\nabla \times G)_2 - F_3(\nabla \times G)_3 \\
&= \langle G, \nabla \times F \rangle - \langle F, \nabla \times G \rangle.
\end{aligned}$$

□

9.8.2 Questions d'orientation

Définition 9.34. Soit $S \subset \mathbb{R}^3$ une surface et soit C un lacet simple de S. Supposons qu'il existe $S_i, S_e \subset S$ tels que $S_i \cap S_e = \emptyset$ et $S \setminus C = S_i \cup S_e$. Dans ce cas, on dit que C est le *bord* de S_i (ou de S_e), ce qui se note :

$$C = \partial S_e = \partial S_e.$$

Remarque. L'orientation de S et le choix de S_i confèrent une orientation à C.

Définition 9.35. Soit $\gamma : [a, b] \to S$ une paramétrisation de C.On dit que γ parcourt C dans le sens *positif* si pour tout $t \in [a, b]$, le vecteur donné par $n(\gamma(t)) \times \gamma'(t)$ se trouve sur le côté intérieur de S.

Remarque. Par continuité, il suffit de vérifier cette propriété sur un point de γ.
Remarque. On remarque également que si $\gamma : [a, b] \to S$ est une paramétrisation du lacet simple C, alors l'image de γ est le bord de de S. Plus formellement :

$$\operatorname{Im} \gamma = \partial S.$$

9.8.3 Théorème de Cauchy-Green-Riemann

Le théorème CGR (également appelé théorème de Green) est une généralisation du théorème fondamental du calcul différentiel et intégral.

Définition 9.36. Soit $F : U \subset \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteur et soit $\gamma : [a, b] \to \mathbb{R}^3$. On introduit la notation suivante :

$$F_k \, dx_k := (F_k \circ \gamma)(t)\gamma_k'(t) \, dt.$$

Lemme 9.37. *Toute surface* $D \subset \mathbb{R}^2$ *peut se* découper *en un ensemble de sous-surfaces de la forme :*

$$[x_i, x_{i+1}] \times [y_j, f(x)].$$

Remarque. La preuve de ce lemme est technique, on peut donc le supposer comme vrai dans ce contexte.

Lemme 9.38. *Soient* $F, G : \mathbb{R}^2 \to \mathbb{R}^2$ *tels que :*

$$F(x, y) = \int_{y_0}^{y} G(x, \gamma) \, d\gamma.$$

Alors :

$$\frac{\partial F}{\partial x}(x, y) = \int_{y_0}^{y} \frac{\partial G}{\partial x}(x, \gamma) \, d\gamma.$$

Remarque. L'idée derrière ce lemme est de montrer que lorsque la variable d'intégration (ici x_2) n'est pas la variable de dérivation (ici x_1), alors les opérateurs de dérivation et d'intégration peuvent être intervertis.

Théorème 9.39. *Soit* $D \subset \mathbb{R}^2$ *une surface dont le bord est une courbe simple de classe* C^1 *par morceaux. Si* P *et* Q *sont de classe* C^1 *sur un ouvert contenant* D*, alors :*

$$\iint_D \left(\tfrac{\partial Q}{\partial x} - \tfrac{\partial P}{\partial y} \right) d(x,y) = \oint_{\partial D} (P\,dx + Q\,dy),$$

où ∂D *est orienté dans le sens anti-horlogique.*

Démonstration. [12] On commence par décomposer la surface D de telle façon que toutes les sous-surfaces soient sous la forme $[x_0, x_1] \times [y, f(x)]$, ce qui est possible grâce au lemme 9.37.

De là, on intègre sur toutes les sous-surfaces et on somme toutes ces intégrales. On obtient alors l'intégrale sur le *vrai bord* du fait que les autres côtés (que le bord ∂D) se compensent dans la somme (en étant parcourus dans des sens opposés). Dès lors, il suffit d'étudier le comportement sur une telle sous-surface pour extrapoler (par la somme).

Supposons donc que D soit sous la forme $[x_0, x_1] \times [y_0, f(x)]$. On peut donc dire que ∂D est composé de quatre courbes C^1, respectivement C_1, C_2, C_3, C_4 où on peut *visualiser* :

$$\begin{aligned}
C_1 &= \{(x, y_0) \text{ t. q. } x_0 \leqslant x \leqslant x_1\} \\
C_2 &= \{(x_1, y) \text{ t. q. } y_0 \leqslant y \leqslant f(x_1)\} \\
C_3 &= \{(x, f(x)) \text{ t. q. } x_0 \leqslant x \leqslant x_1\} \\
C_4 &= \{(x_0, y) \text{ t. q. } y_0 \leqslant y \leqslant f(x_0)\}.
\end{aligned}$$

Comme on a supposé ∂D parcouru dans le sens anti-horlogique, on peut supposer que ∂D correspond au parcours des C_i dans l'ordre C_1 à C_4.

À l'aide du théorème de Fubini et du théorème fondamental, calculons :

$$\iint_D \frac{\partial P}{\partial y}\, d(x,y) = \int_{x_0}^{x_1} \left(\int_{y_0}^{f(x)} \tfrac{\partial P}{\partial y}\, dy \right) dx = \int_{x_0}^{x_1} \big(P(x, f(x)) - P(x, y_0) \big)\, dx.$$

De plus, par définition, on peut exprimer :

$$\int_{\partial D} P\, dx = \sum_{i=1}^{4} \int_{a_i}^{b_i} P \left(\gamma_{i1}(t), \gamma_{i2}(t) \cdot \gamma'_{i1}(t) \right) dt.$$

On observe que les côtés C_2 et C_4 correspondent aux « *barres verticales* » de D, et donc $\gamma'_{i1}(t)$ vaut 0 pour tout t dans $[a, b]$. Il reste alors à sommer les intégrales pour $i \in \{1, 3\}$. On trouve alors :

$$\int_{\partial D} P\, dx = \int_{x_0}^{x_1} P(x, y_0)\, dx + \int_{x_1}^{x_0} P\left(x, f(x)\right) dx = \int_{x_0}^{x_1} P(x, y_0) - P\left(x, f(x)\right) dx = -\iint_D \frac{\partial P}{\partial y}\, d(x,y).$$

Ensuite, on calcule :

$$\iint_D \frac{\partial Q}{\partial x}\, d(x,y) = \int_{x_0}^{x_1} \left(\int_{y_0}^{f(x)} \tfrac{\partial Q}{\partial x}(x,y)\, dy \right) dx.$$

On note alors :

$$U(x,y) := \int_{y_0}^{y} Q(x,s)\, ds.$$

[12] Ceci est une idée de preuve, la preuve donnée n'est pas rigoureuse.

Par le théorème fondamental, on sait :

$$\frac{\partial U}{\partial y}(x, y) = Q(x, y).$$

Le lemme 9.38 dit que l'on peut écrire l'égalité suivante :

$$\frac{\partial U}{\partial x}(x, y) = \int_{y_0}^{y} \frac{\partial Q}{\partial x}(x, s)\, ds.$$

On sait que ∂D est une courbe fermée, par définition. On sait également que l'intégrale du travail d'un champ de vecteur sur une telle courbe est nulle. Dès lors :

$$0 = \oint_{\partial D} \langle \nabla U, d\eta\rangle = \oint_{\partial D} \frac{\partial U}{\partial x}\, dx + \oint_{\partial D} \frac{\partial U}{\partial y}\, dy = \oint_{\partial D} \frac{\partial U}{\partial x}\, dx + \oint_{\partial D} Q\, dy.$$

On trouve donc :

$$\int_{\partial D} Q\, dy = -\int_{\partial D} \frac{\partial U}{\partial x}\, dx = +\iint_{D} \frac{\partial}{\partial x}\left(\tfrac{\partial U}{\partial y}(x, y)\right) d(x, y) = \iint_{D} \frac{\partial Q}{\partial x}\, d(x, y).$$

En sommant $\int_{\partial D} P\, dx$ et $\int_{\partial D} Q\, dy$, on obtient :

$$\int_{\partial D} (P\, dx + Q\, dy) = \iint_{D} \left(\tfrac{\partial Q}{\partial x} - \tfrac{\partial P}{\partial y}\right) d(x, y).$$

□

9.8.4 Théorème de Stokes

Définition 9.40. Une surface S différentiable par morceaux est l'union d'un nombre fini de surfaces paramétrisables qui satisfont les propriétés suivantes :

1. $\forall i \neq j : \int S_i \cap \int S_j = \emptyset$;
2. $\forall i \neq j : \partial S_i \cap \partial S_j$ est :
   - un point ;
   - l'ensemble vide ;
   - une courbe C^1 par morceaux ;
3. $\forall i \neq j, k \neq j, k \neq i : \partial S_i \cap \partial S_j \cap \partial S_k$ est :
   - l'ensemble vide ;
   - un point ;
4. $\forall p, q \in S : \exists \gamma$, un chemin dans S de p à q ;
5. l'ensemble de toutes les courbes ne se trouvant que sur le bord d'un unique S_i dorme un nombre fini de courbes simples, fermées, disjointes et C^1 par morceaux.

Théorème 9.41 (Théorème de Stokes)**.** *Soient $S \subset \mathbb{R}^3$ une surface orientée différentiable par morceaux et $\{C_i\}_{1\leqslant i\leqslant n}$ une collection de lacets simples sur S tels que $S \setminus \left(\bigcup_{i=1}^{n} C_i\right) = S_i \cup S_e$ où $S_i \cap S_e = \emptyset$ et $S_i^{\complement} \cap S_e^{\complement} = \bigcup_{i=1}^{n} C_i$. On suppose S_i compact. Si F est une champ de vecteurs C^1 sur un ouvert contenant $\partial S \cup S$, alors :*

$$\iint_{S_i} \langle \nabla \times F, dn\rangle = \int_{\partial S_i} \langle F, dl\rangle = \sum_{k=1}^{n} \int_{C_k} \big\langle F, (dx, dy, dz)\big\rangle.$$

Démonstration. La preuve donnée ici ne concerne que les surfaces différentiables de classe C^2, ce qui est une hypothèse de régularité trop fort. Soit $U \subset \mathbb{R}^2$. On prend $f : U \to \mathbb{R}^3$ une paramétrisation de $S_i : f(U)$ telle que $\partial S_i = f(\partial U)$. On suppose $f \in C^2$. On prend également $u : \mathbb{R} \to \mathbb{R}^2$ une paramétrisation de ∂U. On définit :

$$\gamma : [a, b] \subset \mathbb{R} \to \mathbb{R}^3 : t \mapsto (f \circ u)(t),$$

une paramétrisation de ∂S_i. On peut alors écrire :

$$\int_{\partial S_i} \left\langle F, (dx, dy, dz) \right\rangle = \int_a^b \left\langle F(\gamma(t)), \tfrac{d\gamma}{dt}(t) \right\rangle dt,$$

où :

$$\frac{d\gamma}{dt}(t) = \frac{d(f \circ u)}{dt}(t) = \frac{\partial f}{\partial u_1}(u(t))\frac{du_1}{dt}(t) + \frac{\partial f}{\partial u_2}(u(t))\frac{du_2}{dt}(t).$$

On peut alors réécrire :

$$\int_{\partial S_i} \left\langle F, (dx, dy, dz) \right\rangle = \int_a^b \left(\left\langle F(f(u(t))), \tfrac{\partial f}{\partial u_1}(u(t))\tfrac{du_1}{dt}(t) \right\rangle + \left\langle F(f(u(t))), \tfrac{\partial f}{\partial u_2}(u(t))\tfrac{du_2}{dt}(t) \right\rangle \right) dt.$$

En posant :

$$\begin{aligned} P(u) &\coloneqq \left\langle F(f(u)), \tfrac{\partial f}{\partial u_1}(u) \right\rangle, \\ Q(u) &\coloneqq \left\langle F(f(u)), \tfrac{\partial f}{\partial u_2}(u) \right\rangle, \end{aligned}$$

on peut réécrire :

$$\begin{aligned} \int_{\partial S_i} \left\langle F, (dx, dy, dz) \right\rangle &= \int_a^b \left(\left\langle F(f(u(t))), \tfrac{\partial f}{\partial u_1}(u(t))\tfrac{du_1}{dt}(t) \right\rangle + \left\langle F(f(u(t))), \tfrac{\partial f}{\partial u_2}(u(t))\tfrac{du_2}{dt}(t) \right\rangle \right) dt \\ &= \int_a^b \left\langle F(f(u(t))), \tfrac{\partial f}{\partial u_1}(u(t)) \right\rangle \tfrac{du_1}{dt}(t)\, dt + \left\langle F(f(u(t))), \tfrac{\partial f}{\partial u_2}(u(t)) \right\rangle \tfrac{du_2}{dt}(t)\, dt \\ &= \int_{\partial U} P\, dx + Q\, dy. \end{aligned}$$

Par le théorème CGR (théorème 9.39), on peut écrire :

$$\int_{\partial U} P\, dx + Q\, dy = \iint_U \left(\tfrac{\partial Q}{\partial u_1} - \tfrac{\partial P}{\partial u_2} \right) d(u_1, u_2).$$

On peut ensuite chercher à évaluer les dérivées partielles de P et Q, ce qui donne :

$$\begin{aligned} \tfrac{\partial P}{\partial u_2} &= \tfrac{\partial}{\partial u_2} \left(\left\langle F(f(u)), \tfrac{\partial f}{\partial u_1}(u) \right\rangle \right) = \tfrac{\partial}{\partial u_2} \sum_{j=1}^{3} F(f(u))_j \cdot \left(\tfrac{\partial f}{\partial u_1}(u) \right)_j = \tfrac{\partial}{\partial u_2} \sum_{j=1}^{3} F_j(f(u)) \tfrac{\partial f_j}{\partial u_1}(u) \\ &= \sum_{j=1}^{3} \tfrac{\partial}{\partial u_2} \left(F_j(f(u)) \tfrac{\partial f_j}{\partial u_1}(u) \right) = \sum_{j=1}^{3} \left(\tfrac{\partial}{\partial u_2}(F_j(f(u))) \tfrac{\partial f_j}{\partial u_2}(u) \tfrac{\partial f_j}{\partial u_1} + F_j(f(u)) \tfrac{\partial}{\partial u_2} \tfrac{\partial f_j}{\partial u_1}(u) \right) \\ &= \sum_{j=1}^{3} \left[\sum_{k=1}^{3} \left(\tfrac{\partial}{\partial x_k} F_j(f(u)) \tfrac{\partial f_k}{\partial u_2}(u) \tfrac{\partial f_j}{\partial u_1}(u) \right) + F_j(f(u)) \tfrac{\partial^2 f}{\partial u_2 \partial u_1}(u) \right]. \end{aligned}$$

De manière similaire, on trouve :

$$\frac{\partial Q}{\partial u_1} = \sum_{j=1}^{3}\left[\sum_{k=1}^{3}\left(\tfrac{\partial}{\partial x_k}F_j(f(u))\tfrac{\partial f_k}{\partial u_1}(u)\tfrac{\partial f_j}{\partial u_2}(u)\right) + F_j(f(u))\tfrac{\partial^2 f}{\partial u_1 \partial u_2}(u)\right].$$

Puisque f est de classe C^2, l'ordre de dérivation n'a as d'importance. Dès lors :

$$\frac{\partial^2 f}{\partial u_1 \partial u_2}(u) = \frac{\partial^2 f}{\partial u_2 \partial u_1}(u).$$

En évaluant $\frac{\partial Q}{\partial u_2} - \frac{\partial P}{\partial u_1}$, on peut supprimer les dérivées secondes de f qui se compensent. Il reste donc :

$$\frac{\partial Q}{\partial u_2} - \frac{\partial P}{\partial u_1} = \sum_{j=1}^{3}\sum_{k=1}^{3}\frac{\partial F_j}{\partial x_k}(f(u))\left(\tfrac{\partial f_k}{\partial u_1}(u)\tfrac{\partial f_j}{\partial u_2}(u) - \tfrac{\partial f_k}{\partial u_2}(u)\tfrac{\partial f_j}{\partial u_1}(u)\right).$$

On y voit apparaître $\frac{\partial f}{\partial u_1} \times \frac{\partial f}{\partial u_2}$. De plus les éléments se somment avec $\nabla \times F$, on retrouve bien :

$$\frac{\partial Q}{\partial u_1} - \frac{\partial P}{\partial u_1} = \left\langle \nabla \times F, \tfrac{\partial f}{\partial u_1} \times \tfrac{\partial f}{\partial u_2}\right\rangle.$$

En prenant l'intégrale double sur U, on retrouve bien $\iint_U \langle \nabla \times F, dn\rangle$ car dn est le vecteur normal, or par définition du produit vectoriel, le vecteur $\frac{\partial f}{\partial u_1} \times \frac{\partial f}{\partial u_2}$ est normal à $\frac{\partial f}{\partial u_1}$ **et** $\frac{\partial f}{\partial u_2}$. □

9.8.5 Théorème de la divergence

Le théorème de la divergence est une conséquence directe du théorème de Stokes qui permet de relier l'intégrale de la divergence d'un champ de vecteurs défini sur un ensemble borné fermé de $\mathbb{R}^3$ et le flux de ce champ de vecteur à travers le bord de cet ensemble.

Théorème 9.42 (Théorème de la divergence)**.** *Soit $\Omega \subset \mathbb{R}^3$ un ensemble borné, fermé et régulier dont le bord est une surface différentiable de classe C^1 par morceaux. Si F est un champ de vecteur de classe C^1 sur un ouvert de $\mathbb{R}^3$ contenant $\Omega \cup \partial\Omega$, alors :*

$$\iiint_{\Omega} \langle \nabla, F\rangle \, d(x, y, z) = \iint_{\partial\Omega} \langle F, dn\rangle.$$

Remarque. Le théorème n'est pas démontré ici.

10 Les séries

10.1 Suites réelles ou complexes

On a une notion de convergence dans $\mathbb{R}^n$, et donc en particulier dans $\mathbb{R}^2$. Cela induit une convergence dans $\mathbb{C}$.
Remarque. Dans cette section, on note $\mathbb{K}$ un corps désignant tant bien les réels que les complexes.
Remarque. Si $z \in \mathbb{K}$, la quantité $|z|$ représente sa valeur absolue s'il est réel et son module s'il est complexe.

Définition 10.1. La suite $(z_k) \subset \mathbb{K}$ converge vers z si :

$$\forall \epsilon > 0 : \exists N \in \mathbb{N} \text{ t. q. } \forall k \geqslant N : |z_k - z| < \epsilon.$$

Proposition 10.2. *La suite complexe (z_k) converge si et seulement si les suites réelles définies par $(\Re z_k)$ et $(\Im z_k)$ convergent. Dans ces conditions, on a :*

$$\lim_{k \to +\infty} z_k = \lim_{k \to +\infty} \Re z_k + i \lim_{k \to +\infty} \Im z_k.$$

Définition 10.3. Une *série* est une somme infinie des éléments d'une suite. On dit que la suite (a_k) est le *terme général* de la série :

$$\sum_{k \geqslant 0} a_k \equiv \sum_{k=0}^{\infty} a_k.$$

Définition 10.4. On dit que la série $\sum_{k=0}^{\infty} a_k$ *converge* si la suite des *sommes partielles* dans $\mathbb{K}$ définie par :

$$s_n = \sum_{k=0}^{n} a_k$$

converge dans $\mathbb{K}$. La limite de cette suite est appelée la *somme* de la série et peut se noter des manières suivantes=

$$\sum_{k \geqslant 0} a_k = \sum_{k=0}^{\infty} a_k = \sum_{k} a_k = \sum a_k.$$

Si cette suite des sommes partielles ne converge pas, on dit que la série diverge.

Proposition 10.5.

- *La série $\sum_k a_k$ converge si et seulement si pour tout m, la série $\sum_{k \geqslant m} a_k$ converge ;*[13]
- *soit $(x_k) \subset \mathbb{R}^+$. S'il existe $K \in \mathbb{N}$ et $\epsilon > 0$ tels que pour tout $k \geqslant K : x_k > \epsilon$, alors la série diverge ;*
- *si la série $\sum_k a_k$ converge, alors le* reste *défini par $\sum_{k=n}^{\infty} a_k$ tend vers 0 pour $n \to +\infty$.*

Proposition 10.6. *Soient $(a_k) \subset \mathbb{K}$ et $\sum_k a_k$. Si la série $\sum_k a_k$ converge, alors $a_k \to 0$ pour $k \to +\infty$.*

Démonstration. Par définition, si la série converge, c'est que la suite des sommes partielles converge. Dès lors, la suite s_n est une suite de Cauchy. On peut donc dire que pour $\epsilon > 0$, il existe $N \in \mathbb{N}$ tel que :

$$\forall m > n > N : |s_m - s_n| = \left| \sum_{k \geqslant n+1} \right| < \epsilon.$$

[13] Mais la somme de ces deux séries n'est pas identique.

En particulier, on peut exprimer :

$$|s_{n+1} - s_n| = |a_{n+1}| < \epsilon.$$

On y voit donc bien que $a_k \to 0$ pour $k \to +\infty$. $\square$

Lemme 10.7 (Formule de la somme d'une suite géométrique)**.** *Soit* $(z_k) \subset \mathbb{K}$ *une suite géométrique. On peut donc dire que* $z_k = z^k$ *pour* $z \in \mathbb{K}$*. La* n*ème somme partielle de la série* $\sum_k z_k$ *est donnée par :*

$$s_n = \frac{1 - z^{n+1}}{1 - z}$$

Démonstration. Par définition de la suite des sommes partielles, on trouve :

$$\sum_{k=0}^{n+1} z^k = s_n + z^{n+1} = 1 + \sum_{k=0}^{n} z^{k+1} = 1 + zs_n.$$

On a donc :

$$s_n + z^{n+1} = 1 + zs_n,$$

ce qui revient à écrire :

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

$\square$

Corollaire 10.8. *Une série de suite géométrique* $(z^k) \subset \mathbb{K}$ *converge si et seulement si*$|z| \lneqq 1$*.*

Démonstration. Si$|z| < 1$, alors pour $n \to +\infty$, on a $z^{n+1} \to 0$, et donc $s_n \to \frac{1}{1-z}$.

Si$|z| \geqslant 1$, alors le terme général z_k ne tend pas vers 0 pour $k \to +\infty$, et donc par la proposition 10.6, on sait que la série ne converge pas. $\square$

10.2 Critères de convergence

Théorème 10.9 (Convergence des séries à termes positifs)**.** *Soit* $(a_k) \subset \mathbb{R}$ *une suite positive (*$\forall k \in \mathbb{N} : a_k \geqslant 0$*). La série* $\sum_k a_k$ *converge si et seulement si la suite des sommes partielles est bornée.*

Démonstration. La suite des sommes partielles est croissante. Par convergence des suites, on sait qu'elle converge si et seulement si elle est bornée, et dans ce cas, $s_n \to \sup s_n$. Sinon, $s_n \to +\infty$. $\square$

Théorème 10.10 (Critère de Leibniz)**.** *Soit* $(a_k) \subset \mathbb{R}$ *une suite définie positive. Si* $\lim_{k\to+\infty} a_k = 0$ *et si* a_k *est décroissante, alors la série* $\sum_k (-1)^k a_k$ *converge.*

Démonstration. En *divisant* la suite des sommes partielles en 2 (les indices pairs et impairs), on trouve une suite croissante et une suite décroissante :

$$s_{2n+2} - s_{2n} = \sum_{k=2n+1}^{2n+2} (-1)_k^a = -a_{2n+1} + a_{2n+2} \leqslant 0,$$

et :

$$s_{2n+3} - s_{2n+1} = \sum_{k=2n+2}^{2n+3} (-1)^k a_k = a_{2n+2} - a_{2n+3} \geqslant 0.$$

Ces deux suites réelles convergent dans $\mathbb{R}$ si et seulement si elles sont respectivement minorée et majorée. On a supposé a_k décroissante. Dès lors :

$$s_{2n+1} \leqslant s_{2n} \leqslant s_0,$$

et :

$$s_{2n} \geqslant s_{2n+1} \geqslant s_1.$$

On note alors $s \coloneqq \inf s_{2n}$ et $t \coloneqq \sup s_{2n+1}$. On observe ensuite :

$$t - s = \lim_{n\to\infty} s_{2n+1} - s_{2n} = -\lim_{n\to+\infty} a_{2n+1} = 0.$$

On a donc $t = s$. Ce la veut dire que pour tout $\epsilon > 0$, il existe N_1 et N_2 dans $\mathbb{N}$ tels que :

$$\begin{aligned} \forall n \geqslant N_1 &: |s_{2n+1} - t| < \epsilon, \\ \forall n \geqslant N_2 &: |s_{2n} - s| < \epsilon. \end{aligned}$$

On en déduit que pour $n \geqslant \max\{N_1, N_2\}$, on trouve $|s_n - t| < \epsilon$. □

Théorème 10.11 (Critère de Cauchy pour les séries). *Soit* $(a_k) \subset \mathbb{K}$. *La série* $\sum_k a_k$ *converge si et seulement si :*

$$\forall \epsilon > 0 : \exists N \in \mathbb{N}\, \text{t. q.}\, \forall m > n \geqslant N : \left| \sum_{k=n+1}^{m} a_k \right| < \epsilon.$$

Démonstration. En considérant la suite des sommes partielles, on sait qu'elle converge si et seulement si elle est de Cauchy. □

10.3 Convergence absolue

Définition 10.12. Soit $(a_k) \subset \mathbb{K}$. On dit que la série $\sum_k a_k$ *converge absolument* si la série $\sum_k |a_k|$ converge.

Remarque. Toute série à termes positifs convergente est absolument convergente.[14]

Proposition 10.13. *Toute série absolument convergente est convergente.*

Démonstration. Soit $\sum_k a_k$ une série absolument convergente. On sait que la série $\sum_k |a_k|$ converge, et est donc une suite de Cauchy. Dès lors, on sait que pour tout $\epsilon > 0$, il existe N tel que

$$\forall m > n \geqslant N : \left| \sum_{k=n+1}^{m} |a_k| \right| < \epsilon.$$

Cependant, on peut exprimer :

$$\epsilon > \left| \sum_{k=n+1} |a_k| \right| = \sum_{k=n+1}^{m} |a_k| \geqslant \left| \sum_{k=n+1}^{m} a_k \right|.$$

On en déduit que la série $\sum_k a_k$ (ou la série des sommes partielles) satisfait également le critère de Cauchy, et est donc convergente. □

[14]Cette remarque peut également se formuler « Pour une suite à termes positifs, la convergence et la convergence absolue sont équivalentes ».

Théorème 10.14. *Si la série $\sum_k a_k$ possède une série majorant $\sum_k b_k$ telle que $\sum_k b_k$ est convergente, alors $\sum_k a_k$ est absolument convergente.*

Démonstration. On prend $\sum_k b_k$ un majorant de $\sum_k a_k$. Par le critère de Cauchy, on sait :

$$\forall \epsilon > 0 : \exists N \text{ t. q. } \forall m > n \geqslant N : \left| \sum_{k=n+1}^{m} b_k \right| < \epsilon.$$

Par définition du majorant, on sait qu'il existe $K \in \mathbb{N}$ tel que :

$$\forall k \geqslant K : |a_k| \leqslant b_k.$$

On en déduit

$$\forall m > n \geqslant \max\{N, K\} : \sum_{k=n+1}^{m} |a_k| \leqslant \sum_{k=n+1}^{m} b_k < \epsilon.$$

Donc $\sum_k |a_k|$ satisfait le critère de Cauchy, elle est donc convergente. □

Remarque. La contraposée de ce théorème affirme que si $\sum_k b_k$ est un majorant de $\sum_k a_k$ et que $\sum_k |a_k|$ n'est pas convergente, alors $\sum_k b_k$ diverge.

Corollaire 10.15 (Critère d'équivalence)**.** *Soient $\sum_k a_k$ et $\sum_k b_k$ deux séries réelles à termes positifs. Supposons que $\forall k \in \mathbb{N} : b_k \neq 0$. On note :*

$$\alpha := \lim_{k \to +\infty} \frac{a_k}{b_k} \in \mathbb{R}^+ \cup \{+\infty\}.$$

1. *Si $0 < \alpha < +\infty$, alors $\sum_k a_k$ converge (respectivement diverge) si et seulement si $\sum_k b_k$ converge (respectivement diverge) ;*
2. *si $\alpha = 0$ et $\sum_k b_k$ converge, alors $\sum_k a_k$ converge ;*
3. *si $\alpha = +\infty$ et $\sum_k b_k$ diverge, alors $\sum_k a_k$ diverge.*

Remarque. Dans le cas 1, on remarque que α n'est ni nul ni infini, et donc que $\sum a_k$ et $\sum_k b_k$ tendent vers 0 à la même vitesse.

Démonstration. Preuve du point 1. Par hypothèse, il existe $K \in \mathbb{N}$ tel que :

$$\forall k \geqslant K : -\epsilon + \alpha \leqslant \frac{a_k}{b_k} \leqslant \alpha + \epsilon.$$

En prenant ϵ suffisamment petit, on peut poser :

$$\begin{cases} \beta & := \alpha - \epsilon > 0, \\ \gamma & := \alpha + \epsilon. \end{cases}$$

On en déduit qu'il existe N tel que :

$$\forall k \geqslant N : \beta b_k \leqslant a_k \leqslant \gamma b_k.$$

On a alors γb_k, un majorant de a_k, donc par le critère de comparaison, si $\sum_k b_k$ converge, alors $\sum_k a_k$ converge. De même, on a a_k un majorant de βb_k, donc par le critère de comparaison, si $\sum_k a_k$ converge, alors $\sum_k b_k$ converge.

Les points 2 et 3 se démontrent de manière similaire. □

Théorème 10.16 (Critère de la racine)**.** *Soit* $(a_k) \subset \mathbb{K}$*. On note :*

$$\alpha \coloneqq \limsup_{k\to+\infty} \sqrt[k]{|a_k|}.$$

1. *Si* $\alpha < 1$*, alors* $\sum_k a_k$ *converge absolument ;*
2. *si* $\alpha > 1$*, alors* $\sum_k a_k$ *diverge.*

Remarque. Le critère de la racine ne donne aucune information sur le cas où $\alpha = 1$ car il est possible de trouver des exemples convergents, absolument convergents et divergents.

Démonstration. Preuve du point 1. Soit $a \in (\alpha, 1)$. Par hypothèse, il existe K tel que :

$$\forall k \geqslant K : \sqrt[k]{|a_k|} < a,$$

ou encore$|a_k| < a^k$. Or, on sait la série géométrique $\sum_k a^k$ convergente (car $a \in (0,1)$), et par le critère de comparaison, comme $\sum_k a^k$ est un majorant de $\sum_k a_k$, le critère de comparaison dit que $\sum_k a_k$ converge.

Preuve du point 2. Pour tout $n \in \mathbb{N}$, il existe $k \geqslant n$ tel que$|a_k| \geqslant 1$. Dès lors, on observe :

$$\lim_{k\to+\infty} |a_k| \geqslant 1.$$

Le terme général de la série ne tend donc pas vers 0, et par la proposition 10.6, on sait que $\sum_k a_k$ diverge. □

Théorème 10.17 (Critère du quotient)**.** *Soient* $(a_k) \subset \mathbb{K}$ *et* $K_0 \in \mathbb{N}$ *tels que :*

$$\forall k \geqslant K_0 : a_k \neq 0.$$

1. *S'il existe* $K \geqslant K_0$ *et* $a < 1$ *tels que :*
$$\forall k \geqslant K = \frac{|a_{k+1}|}{|a_k|} < a,$$
alors la série $\sum_k a_k$ *est convergente ;*
2. *s'il existe* $K \geqslant K_0$ *tel que :*
$$\forall k \geqslant K : \frac{|a_{k+1}|}{|a_k|} \geqslant 1,$$
alors la série $\sum_k a_k$ *est divergente.*

Démonstration. Preuve du point 1. Commençons par montrer que pour tout $k \geqslant K+1$, on a :

$$|a_k| \leqslant \frac{|a_K|}{a^K} a^k.$$

Pour $k = K+1$, on a par hypothèse :

$$\frac{|a_k|}{|a_K|} \leqslant a,$$

ou encore :

$$|a_k| \leqslant |a_K|\, a = |a_K|\, a \frac{a^K}{a^K} = |a_K| \frac{a^k}{a^K}.$$

Ensuite, supposons cette inégalité vraie pour $k > K$, et montrons-la pour $k+1$:

$$|a_{k+1}| \leqslant a|a_k| = a\frac{|a_K|}{a^K} a^k = \frac{|a_K|}{a^K} a^{k+1}.$$

On peut donc exprimer, pour tout $k \geqslant K$:

$$|a_k| \leqslant \beta a^k,$$

où $\beta = \frac{|a_K|}{a^K} \in \mathbb{R}$ est une constante. On peut donc dire que la série $\sum_k a_k$ est majorée par la série $\beta \sum_k a^k$, qui converge car $a < 1$, et donc par le critère de comparaison, on peut dire que $\sum_k a_k$ converge.

Preuve du point 2. Pour tout $m \in \mathbb{N}$, on sait que $|a_{K+m}| \geqslant |a_K|$. Donc $|a_k|$ ne tend pas vers 0 pour $k \to +\infty$. Par la proposition 10.6, on sait que $\sum_k a_k$ ne converge pas. □

10.4 Critères de Dirichlet et d'Abel

Théorème 10.18 (Critère de Dirichlet). *Soient $(a_k) \subset \mathbb{R}$ et $(b_k) \subset \mathbb{C}$ telles que :*

(i) $\lim_{k\to+\infty} a_k = 0$;

(ii) $\sum_k |a_{k+1} - a_k|$ *est convergente ;*

(iii) *la suite des sommes partielles de (b_k) est bornée dans $\mathbb{C}$.*

Alors la série $\sum_k (a_k b_k)$ est convergente dans $\mathbb{C}$.

Démonstration. On note (B_n) la suite des sommes partielles de $\sum_k b_k$ et (s_n) la suite des sommes partielles de $\sum_k a_k b_k$. On écrit s_n sous la forme :

$$s_n = \sum_{k=0}^{n} a_k b_k = \sum_{k=0}^{n} a_k B_k - \sum_{k=0}^{n-1} a_{k+1} B_k = \sum_{k=0}^{n-1} B_k (a_k - a_{k+1}) - a_n B_n.$$

On sait par hypothèse que B_n est bornée, et donc il existe $M \gneqq$ tel que :

$$\left|(a_j - a_{j+1}) B_j\right| \leqslant M \left|a_j - a_{j+1}\right|.$$

Or, par hypothèse également, la série $\sum_j \left|a_j - a_{j+1}\right|$ converge. Comme la série convergente $M \sum_k |a_k - a_{k+1}|$ est un majorant de la série $\sum_k (a_j - a_{j+1}) B_j$. Par le critère de comparaison, on sait que :

$$\sigma \coloneqq \lim_{n\to+\infty} \left(\sum_{k=0}^{n-1} (a_k - a_{k+1}) B_k \right) \in \mathbb{R}.$$

On sait également par hypothèse que la suite (a_k) tend vers 0 pour $k \to +\infty$. Par les règles de calcul des suites, on sait que la suite $(a_n B_n)$ tend également vers 0 pour $n \to +\infty$. Dès lors, par les règles de calcul sur les suites, on trouve que la suite des sommes partielles de $\sum_k a_k b_k$ converge. □

Remarque. Lorsque (a_k) est décroissante, le critère de Dirichlet peut s'exprimer comme étant de critère d'Abel.

Théorème 10.19 (Critère d'Abel). *Soient $(a_k) \subset \mathbb{R}$ et $(b_k) \subset \mathbb{C}$ telles que :*

(i) $\lim_{k\to+\infty} a_k = 0$;

(ii) $a_{k+1} < a_k$ *pour tout k ;*

(iii) *la suite des sommes partielles de (b_k) est bornée dans $\mathbb{C}$.*

Alors la série $\sum_k (a_k b_k)$ est convergente dans $\mathbb{C}$.

Démonstration. On observe que :

$$\sum_{k=0}^{n}|a_k - a_{k+1}| = \sum_{k=0}^{n}(a_k - a_{k+1}) = a_0 + \sum_{k=0}^{n} a_k - \sum_{k=0}^{n} a_k - a_{n+1} = a_0 - a_{n+1}.$$

Dès lors :

$$\lim_{n\to+\infty}\sum_{k=0}^{n}|a_k - a_{k+1}| = a_0 - \lim_{n\to+\infty} a_{n+1} = a_0.$$

Donc la série $\sum_k |a_k - a_{k+1}|$ converge. En appliquant le critère de Dirichlet, on trouve effectivement que $\sum_k a_k b_k$ converge. □

10.5 Opérations sur les séries

Définition 10.20. Soient la suite $(a_k) \subset \mathbb{K}$ et $\sigma : \mathbb{N} \to \mathbb{N}$ bijective. On dit que la série $\sum_k a_{\sigma(k)}$ est un *réarrangement* de la série $\sum_k a_k$.

Remarque. Si σ est une permutation de $\mathbb{N}$ telle que seul un nombre fini de valeurs k ne sont pas envoyées sur elle-même, alors la série $\sum_k a_{\sigma(k)}$ converge si et seulement si $\sum_k a_k$ converge et leur somme sera la même.

Théorème 10.21 (Théorème de réarrangment de Riemann)**.** *Soit $(a_k) \subset \mathbb{R}$ telle que $\sum_k a_k$ converge sans converger absolument. Alors pour tout $s \in \mathbb{R}$, il existe $\sigma : \mathbb{N} \to \mathbb{N}$ bijective telle que $\sum_k a_{\sigma(k)} = s$. De plus, il existe σ une permutation de $\mathbb{N}$ telle que $\sum_k a_{\sigma(k)}$ diverge.*

Remarque. Ce théorème n'est pas démontré ici car il dépasse le niveau du cours.

Définition 10.22. Soit $(a_k) \subset \mathbb{R}$ telle que $\sum_k a_k$ converge sans converger absolument. Alors on dit que la convergence de $\sum_k a_k$ est *conditionnelle*. Soit $(b_k) \subset \mathbb{R}$ telle que $\sum_k b_k$ converge absolument. Alors on dit que la convergence de $\sum_k b_k$ est *inconditionnelle*

Remarque. Cette dénomination vient du fait qu'une suite qui converge sans converger absolument peut converger vers n'importe quel réel (et peut ne pas converger) selon le réarrangement. Une série qui converge absolument ne dépend pas du réarrangement (voir théorème 10.23).

Théorème 10.23. *Si $\sum_k a_k$ est une série absolument convergente et σ est une permutation de $\mathbb{N}$, alors $\sum_k a_{\sigma(k)}$ est absolument convergente et la somme ne dépend pas de σ.*

Démonstration. Commençons par montrer que $\sum_k a_{\sigma(k)}$ converge absolument, et nous montrerons que la somme ne dépend pas de σ ensuite.

Soit $n \in \mathbb{N}$. On pose $\tilde{n} := \max_{0\leqslant i\leqslant n} \sigma(i)$. Cela implique que $\sigma([\![0,n]\!]) \subset [\![0,\tilde{n}]\!]$. On en déduit :

$$\sum_{k=0}^{n}\left|a_{\sigma(k)}\right| \leqslant \sum_{k=0}^{\tilde{n}}|a_k| \leqslant \sum_{k\geqslant 0}|a_k|.$$

La série est donc bornée et est donc absolument convergente.

On note respectivement $(\hat{s}_n)$ et (s_n) les suites des sommes partielles de $\sum_k\left|a_{\sigma(k)}\right|$ et $\sigma_k|a_k|$. Soit $\epsilon > 0$. Par le critère de Cauchy, on sait qu'il existe N tel que :

$$\forall m > n \geqslant N : \sum_{k=n+1}^{m}|a_k| < \frac{\epsilon}{2}.$$

On prend alors $\hat{N} \geqslant N$ tel que :

$$[\![0, N]\!] \subseteq \sigma\left([\![0, \hat{N}]\!]\right).$$

Soit $n \geqslant \hat{N}$. On pose $A_n \coloneqq \sigma\left([\![0, n]\!]\right) \setminus [\![0, N]\!]$. On peut exprimer :

$$|s_n - \hat{s}_n| = \left|\sum_{k=0}^{n} a_k - \sum_{k=0}^{n} a_{\sigma(k)}\right| = \left|\sum_{k=\hat{N}+1}^{n} a_k + \sum_{k \in A_n} a_k\right| \leqslant \sum_{k=\hat{N}+1}^{n} |a_k| + \sum_{k \in A_n} |a_k| \leqslant \sum_{k=\hat{N}+1}^{n} |x_k| + \sum_{k=\hat{N}+1}^{\tilde{n}} |a_k| < \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

On a donc montré que pour tout $n \geqslant \hat{N}$ et $\epsilon > 0$, on a $|s_n - \hat{s}_n|$, ou encore $s_n = \hat{s}_n$. En prenant la limite, l'égalité reste et on montre l'égalité des sommes des suites. □

Définition 10.24. Soit $(a_{ij})_{i,j \in \mathbb{N}}$ une *suite double*. La *série double* $\sum_{i,j} a_{ij}$ dite *sommable* si :

$$\sup_{n \in \mathbb{N}} \sum_{i,j=0}^{n} |a_{ij}| < +\infty$$

Proposition 10.25. *Si $\sum_k a_k$ et $\sum_k b_k$ sont deux séries absolument convergentes. Alors la série double $\sum_{i,j} a_{ij}$ est sommable.*

Théorème 10.26 (Théorème de Fubini pour les séries)**.** *Soit $(a_{ij}) \subset \mathbb{K}$. Si la série double $\sum_{i,j} a_{ij}$ est sommable, alors pour tout $\sigma : \mathbb{N} \to \mathbb{N}^2$[15], la série $\sum_k a_{\sigma(k)}$ est absolument convergente et la somme s est indépendante de σ.*

De plus, pour tout $i, j \in \mathbb{N}$, les séries $\sum_k a_{ik}$ et $\sum_k a_{kj}$ sont absolument convergentes et :

$$\sum_i \left(\sum_j a_{ij}\right) = \sum_j \left(\sum_i a_{ij}\right) = s.$$

Remarque. Ce résultat n'est pas démontré.

Définition 10.27. Le *produit de Cauchy* de deux séries $\sum_k a_k$ et $\sum_k b_k$ est défini par la série :

$$\sum_n \sum_{\sum_{k=0}^{n}} a_k b_{n-k}$$

Remarque. Une conséquence du théorème de Cauchy pour les suites est que si $\sum_k a_k$ et $\sum_k b_k$ sont deux séries absolument convergentes, alors leur produit de Cauchy est une série absolument convergente qui a pour somme $\left(\sum_k a_k\right)\left(\sum_k b_k\right)$.

De plus, le produit de Cauchy de deux suites convergentes de manière conditionnelle peut de pas converger.

10.6 Séries de puissances

Définition 10.28. Soit $(a_k) \subset \mathbb{K}$ une suite. On appelle la série $\sum_k a_k (z - z_0)^k$ une série de puissances où $z_0 \in \mathbb{K}$.

[15] $\mathbb{N}$ et $\mathbb{N}^2$ ont bien le même cardinal, pour s'en convaincre, on peut imaginer une bijection qui remplit $\mathbb{N}^2$ par ses diagonales. Il y a également la fonction de couplage de Cantor bien connue qui le montre.

Théorème 10.29. *Soit* $(a_k) \subset \mathbb{K}$*. Alors il existe* $\rho_a \in [0, +\infty]$ *, appelé* rayon de convergence *tel que pour tout* z *tel que*$|z - z_0| < \rho_a$*, la série de puissances* $\sum_k a_k(z-z_0)^k$ *converge absolument et pour tout* z *tel que*$|z - z_0| > \rho_a$*, la série de puissances diverge. De plus, la* formule de Hadamard *donne :*

$$\rho_a = \frac{1}{\limsup_{k\to+\infty} \sqrt[k]{|a_k|}}.$$

Démonstration. On observe :

$$\limsup_{k\to+\infty} \sqrt[k]{\left|a_k(z-z_0)^k\right|} = |z - z_0| \limsup_{k\to+\infty} |a_k| = \frac{|z-z_0|}{\rho_a}.$$

Il faut donc, par le critère de la racine, que$|z - z_0| \lneqq \rho_a$ pour une convergence absolue et$|z - z_0| \gneqq \rho_a$ pour une divergence. □

Définition 10.30. Le disque de contre z_0 et de rayon ρ_a est appelé *rayon de convergence*.

Proposition 10.31. *Soit* $(a_k) \subset \mathbb{K}$ *telle que :*

$$\lim_{k\to+\infty} \left|\frac{a_k}{a_{k+1}}\right|$$

existe dans $\mathbb{R} \cup \{\pm\infty\}$*. Alors le rayon de convergence de la série de puissances* $\sum_k a_k(z-z_0)^k$ *est donné par :*

$$\rho_a = \lim_{k\to+\infty} \left|\frac{a_k}{a_{k+1}}\right|.$$

Démonstration. On calcule :

$$\lim_{k\to+\infty} \left|\frac{a_{k+1}(z-z_0)^{k+1}}{a_k(z-z_0)^k}\right| = \frac{|z-z_0|}{\lim_{k\to+\infty}\left|\frac{a_k}{a_{k+1}}\right|}.$$

On en conclut que la série $\sum_k a - K(z-z_0)^k$ converge si $|z - z_0| < \lim_{k\to+\infty}\left|\frac{a_k}{a_{k+1}}\right|$ et diverge si $|z - z_0| > \lim_{k\to+\infty}\left|\frac{a_k}{a_{k+1}}\right|$. Par le théorème 10.29, on en déduit que :

$$\rho_a = \lim_{k\to+\infty} \left|\frac{a_k}{a_{k+1}}\right|.$$

□

10.7 L'ensemble des réels n'est pas dénombrable

Définition 10.32. Un ensemble X est dit :

- *fini* s'il existe $m \in \mathbb{N}$ et une bijection $f : X \to [\![1, m]\!]$;
- *infini* s'il n'est pas fini ;
- *dénombrable* s'il existe une bijection $f : X \to \mathbb{N}$;
- *au plus dénombrable* s'il est fini ou dénombrable.

Définition 10.33. Deux ensembles X et Y sont dits *équipotents* s'il existe $f : X \to Y$ bijective.

Proposition 10.34. *Soient X et Y deux ensembles au plus dénombrable. Alors les ensembles suivantes :*

- $A \subset X$;
- $X \cup Y$;
- $X \times Y$,

sont au plus dénombrables.

Démonstration. Si A est fini, il n'y a rien à démontrer. Si A est infini (et donc X est infini), alors on sait qu'il existe $g : X \to \mathbb{N}$ une bijection. On peut restreindre cette bijection à $g|_A : A \to g(A) \subseteq \mathbb{N}$, qui reste une bijection. A est donc dénombrable.

Si X et Y sont finis, on sait que $X \cup Y$ est fini. Si X est fini et Y infini (le cas symétrique s'en déduit), alors il existe $m \in \mathbb{N}$ et $f : X \to [\![1, m]\!]$ et il existe également $g : Y \to \mathbb{N}$. On définit alors :

$$h : X \cup Y \to \mathbb{N} : x \mapsto f(x-1)1_X(x) + g(x)1_Y(x).$$

On remarque que h est bijective. De même, si X et Y sont tous deux infinis, il existe $f : X \to \mathbb{N}$ et $g : Y \to \mathbb{N}$ bijectifs. On définit à nouveau :

$$h : X \cup Y \to \mathbb{N} : x \mapsto 2f(x)1_X(x) + (2g(x)+1)1_Y(x).$$

On remarque également que h est bijective.

Si X et Y sont tous deux finis, alors $X \times Y$ est fini également. Si X et Y sont tous deux infinis, alors l'idée de la preuve est de construire une fonction $\mathbb{N} \to X \times Y$ qui remplit le *plan discret* $X \times Y$ en diagonale, et qui, par construction est bijective.[16] Si X est fini et Y infini (ou symétriquement), l'argument est le même. $\square$

Proposition 10.35. *Une union dénombrable d'ensembles au plus dénombrables est toujours dénombrable.*

Proposition 10.36. *L'ensemble $\mathbb{Q}$ des rationnels est dénombrable.*

Définition 10.37. On prend $f : \mathbb{Q} \to \mathbb{Z} \times \mathbb{Z}_0 : \frac{p}{q} \mapsto (p, q)$. On voit que f est injective, et donc $f|_{\mathbb{Q}}$ est une bijection avec un sous-ensemble de $\mathbb{Z} \times \mathbb{Z}_0$ qui, lui, est dénombrable. On sait également $\mathbb{N} \subset \mathbb{Q}$ et donc $\mathbb{Q}$ ne peut être fini. $\mathbb{Q}$ doit donc forcément être dénombrable.

Remarque. La théorie des séries permet de justifier la notion de développement décimal illimité. En effet :

$$1.999\ldots = 1 + \sum_{k \geqslant 1} \frac{9}{10} = 1 + 9 \sum_{k \geqslant 1} 10^{-k} = 1 + 9\frac{10^{-1}}{1 - 10^{-1}} = 1 + 1 = 2.$$

Tout nombre $x \in [0, 1]$ peut se développer dans une base $b \geqslant 2$ de la manière suivante :

$$x = \sum_{k \geqslant 1} x_k b^{-k}.$$

Théorème 10.38. *L'ensemble $\mathbb{R}$ des nombres réels n'est pas dénombrable.*

Démonstration. Montrons que $[0, 1)$ est indénombrable. Supposons par l'absurde qu'il existe une bijection $f : \mathbb{N} \to [0, 1)$. Cette fonction *ordonne* l'intervalle $[0, 1)$ par un ensemble $f(0), f(1), \ldots$ Chacun de ces nombres

[16] À nouveau, la fonction de couplage de Cantor peut être utilisée, et c'est une fonction $X \times Y \to \mathbb{N}$ donc qui prend le problème dans l'autre sens.

a une représentation par développement décimal unique :

$$f(n) = \sum_{k \geqslant 1} (x_n)_k 10^{-k}.$$

On considère la suite (a_k) définie par :

$$a_k = \begin{cases} 0 & \text{si } (a_{k-1})_k \neq 0 \\ 1 & \text{si } (a_{k-1})_k = 0 \end{cases} = 1 - \delta_{0(a_{k-1})_k}.$$

On prend $a \in [0,1)$ défini par $\sum_{k \geqslant 1} a_k 10^{-k}$. Par construction, on sait pour tout $k \in \mathbb{N}$ que $a \neq f(k)$. On peut donc dire $a \notin f(\mathbb{N}) = [0,1)$, or f est supposée injective. C'est une contradiction, et donc $[0,1)$ n'est pas dénombrable. □

Corollaire 10.39. *L'ensemble $\mathbb{R} \setminus \mathbb{Q}$ des nombres irrationnels est indénombrable.*

Démonstration. Supposons par l'absurde que $\mathbb{R} \setminus \mathbb{Q}$ est dénombrable. Par la proposition 10.34, on sait que $\mathbb{R} = (\mathbb{Q} \setminus \mathbb{Q}) \cup \mathbb{Q}$ est dénombrable. Or on vient de voir dans le théorème 10.38 que $\mathbb{R}$ est indénombrable. Il faut donc $\mathbb{R} \setminus \mathbb{Q}$ indénombrable. □

Remarque. Il est assez intuitif de se dire qu'en retirant un ensemble *uniquement dénombrable* d'un ensemble indénombrable, la différence ensembliste reste indénombrable.

11 Approximation de Taylor

11.1 Approximation polynomiale des fonctions d'une variable

Théorème 11.1 (Théorème de Taylor)**.** *Soient* $f : I \to \mathbb{R}$, $a \in \operatorname{int} I$, *et* $k \geqslant 1$. *Si* f *est* k *fois dérivable en* a, *alors il existe une unique polynôme :*

$$P(x) = \sum_{i=0}^{k} c_i (x-a)^i$$

tel que :

$$\lim_{x \to a} \frac{f(x) - P(x)}{(x-a)^k} = 0.$$

De plus, $P = T_k(f, a)$ *où :*

$$T_k(f, a)(x) \coloneqq \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!} (x-a)^n.$$

Définition 11.2. On appelle le polynôme $T_k(f, a)$ le *polynôme de Taylor de* f *autour de* a.

Définition 11.3. Si $f \in C^\infty$, on définit la *série de Taylor de* f *autour de* a par :

$$T(f, a)(x) \coloneqq \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n.$$

Remarque. La série de Taylor est une série de puissance et converge donc pour tout x tel que$|x - a| < \rho$, où :

$$\rho = \frac{1}{\limsup_{n \to +\infty} \frac{f^{(n)}(a)}{n!}}.$$

Si $\rho \gneqq 0$, la série $T(f, a)$ définit une fonction (car il y a convergence) sur un voisinage de a. Il faut alors montrer que $T(f, a) = f$.

Définition 11.4. On dit que la fonction $f : I \to \mathbb{R}$ est *analytique* si :

$$\forall a \in I : \exists \delta \text{ t. q. } \forall x \in (a - \delta, a + \delta) : f(x) = T(f, a)(x).$$

Preuve du théorème de Taylor. Montrons d'abord l'unicité, et montrons ensuite l'existence sur base de l'unicité.

Soient P_1, P_2 deux polynômes de degré $\leqslant k$ tels que :

$$\lim_{x \to a} \frac{f(x) - P_1(x)}{(x-a)^k} = \lim_{x \to a} \frac{f(x) - P_2(x)}{(x-a)^k} = 0.$$

Par les règles de calcul sur les limites, on a :

$$\lim_{x \to a} \frac{P_1(x) - P_2(x)}{(x-a)^k} = 0.$$

On note $P(h) \coloneqq P_1(a+h) - P_2(a+h)$. On peut également écrire :

$$P(h) = \sum_{i=0}^{k} a_i^i.$$

On a donc $P(h) \to 0$ pour $h \to 0$. Il faut donc avoir $a_i = 0$ pour tout i. On a donc $P_1(x) - P_2(x) = 0$, ce qui montre que le polynôme est unique.

Montrons maintenant qu'il existe. Plus précisément, montrons que $T_k(f, a)$ est ce polynôme.

On pose $R_k(f, a)(x) = f(x) - T_k(f, a)(x)$. Montrons que $\frac{R_k(f,a)(x)}{x-a} \to 0$ pour $x \to a$. Prouvons-cela par induction sur k. Pour $k = 1$, par définition de la dérivée, on trouve :

$$\lim_{x\to a} \frac{f(x) - T_1(f, a)}{x - a} = \lim_{x\to a} \frac{f(x) - f(a) - f'(a)(x - a)}{x - a} = \lim_{x\to a} \frac{f(x) - f(a)}{x - a} - f'(a) = 0.$$

Pour le pas de récurrence, on calcule :

$$R_{k+1}(f, a)(x) = R_{k+1}(f, a)(x) - R_{k+1}(f, a)(a).$$

Par le théorème de la moyenne, on sait qu'il existe ξ compris strictement entre x et a tel que :

$$R_{k+1}(f, a)(x) = R_{k+1}(f, a)'(\xi)(x - a).$$

En calculant le terme de droite, on trouve :

$$\begin{aligned} R_{k+1}(f, a)'(\xi) &= f'(\xi) - \sum_{n=1}^{k+1} \frac{f^{(n)}(\xi)}{n!} n(x - a)^{n-1} = f'(\xi) - \sum_{n=0}^{k} \frac{f^{(n+1)}(\xi)}{n!}(x - a)^n \\ &= f'(\xi) - \sum_{n=0}^{k} \frac{(f')^{(n)}(\xi)}{n!}(x - a)^n = f'(\xi) - T_k(f', a)(\xi) = R_k(f', a)(\xi). \end{aligned}$$

On procède ensuite à la récurrence :

$$\frac{R_{k+1}(f, a)(x)}{(x - a)^{k+1}} = \frac{R_k(f', a)(\xi)(x - a)}{(x - a)^{k+1}} = \frac{R_k(f', a)(\xi)}{(\xi - a)^k} \frac{(\xi - a)^k}{(x - a)^k}.$$

On en déduit alors :

$$\lim_{x\to a} \left| \frac{R_{k+1}(f, a)(x)}{(x - a)^{k+1}} \right| = \epsilon(k) \lim_{\xi\to a} \left| \frac{R_k(f', a)(\xi)}{(\xi - a)^k} \right|,$$

où :

$$\epsilon(k) = \frac{\left|(\xi - a)^k\right|}{\left|(x - a)^k\right|} \leqslant 1.$$

On a donc bien :

$$\lim_{x\to a} \left| \frac{R_{k+1}(f, a)(x)}{(x - a)^{k+1}} \right| = 0.$$

□

Théorème 11.5 (Formule du reste de Lagrange)**.** *Sous les hypothèses du théorème de Taylor, si de plus $f^{(k)}$ est continue et $f^{(k+1)}$ existe, alors pour tout x, il existe ζ strictement entre x et a tel que :*

$$R_k(f, a)(x) = \frac{f^{(k+1)}(\zeta)}{(k + 1)!}(x - a)^{k+1}.$$

Corollaire 11.6. *Soit $f \in C^\infty(I = (x_1, x_2))$. Si :*

$$\forall x_0 \in I : \exists \delta > 0, M > 0 \text{ t. q.}([x_0 \pm \delta] \subset I) \wedge (\forall k \in \mathbb{N}, x \in (x_0 \pm \delta) : \left|f^{(k)}(x)\right| \leqslant k!M^k),$$

alors f est analytique.

Démonstration. On remarque que le rayon de convergence est $\geqslant \frac{1}{M}$ par :

$$\frac{1}{\rho} = \sqrt[k]{\frac{f^{(k)}(a)}{k!}} \leqslant \sqrt[k]{M^k} = M.$$

Dès lors :

$$\sum_{k\geqslant 0} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

converge pour tout x tel que$|x-a| < \frac{1}{M}$. On calcule ensuite :

$$\left|R_k(f,a)(x)\right| = \frac{1}{k!}\left|f^{(k)}(a)(x-a)^k\right| \leqslant \frac{1}{k!}\left|k!M^k(x-a)^k\right| = M^k(x-a)^k.$$

Or on sait$|x-a| \gneqq \frac{1}{M}$. Dès lors :

$$\left|M(x-a)\right|^k \to 0.$$

□

11.2 Fonctions de plusieurs variables

Définition 11.7. Soit $f : E \subset \mathbb{R}^m \to \mathbb{R}$. On dit que f est *continûment dérivable* si les dérivées partielles de f sont continues. Cela se note $f \in C^1$.

Remarque. Comme pour les fonction $f : \mathbb{R} \to \mathbb{R}$, on définit $f \in C^n$ par récurrence si $f \in C^{n-1}$ est les dérivées partielles de $f^{(n)}$ existent et sont continues ($f^{(n)}$ est évidemment un vecteur de dérivées partielles).

Définition 11.8. On dit que $f : E \subset \mathbb{R}^m \to \mathbb{R}$ est k fois continûment dérivable (ou $f \in C^k$) si toutes les dérivées partielles

$$\frac{\partial^k}{\partial x_{i_k}\partial x_{i_{k-1}}\cdots\partial x_{i_2}\partial x_{i_1}}$$

existent et sont continues. On appelle ces dérivées partielles les *dérivées partielles d'ordre* k *de* f.

Définition 11.9. La fonction $f : E \subset \mathbb{R}^m \to \mathbb{R}^n$ est dite k fois continûment dérivable si toutes ses composantes f_i avec $1 \leqslant i \leqslant n$ sont de classe C^k.

Théorème 11.10 (Théorème de Clairault)**.** *Soit* $f : E \subset \mathbb{R}^m \to \mathbb{R}$ *une fonction définie sur un ouvert* E *de classe* C^2*. Alors pour tout* $a \in E$ *et* $1 \leqslant i, j \leqslant n$*, on a :*

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(a) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(a).$$

Remarque. Ce résultat n'est pas démontré.

Corollaire 11.11. *Soit* $f : E \subset \mathbb{R}^m \to \mathbb{R}^n$ *une fonction de classe* C^k *définie sur un ouvert* E*. Pour tout* $\gamma \leqslant k$*, les dérivées partielles d'ordre* γ *de* f *:*

$$\frac{\partial^\gamma f}{\partial x_{i_\gamma}\partial x_{i_{\gamma-1}}\cdots\partial x_{i_2}\partial x_{i_1}}$$

ne dépendent pas de l'ordre de dérivation (l'ordre des i_j*).*

Définition 11.12. Soit $\alpha \in \mathbb{N}^m$. On dit que α est un *multi-indice* et sa longueur est définie par :

$$\overline{\alpha} \coloneqq \sum_{k=1}^{m} \alpha_m.$$

Définition 11.13. Soient $f : E \subset \mathbb{R}^m \to \mathbb{R}$ une fonction de classe C^k définie sur un ouvert E et $\alpha \in \mathbb{N}^m$ un multi-indice tel que $\overline{\alpha} = k$. On introduit la notation suivante :

$$\partial^\alpha f \coloneqq \frac{\partial^{\overline{\alpha}} f}{\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}}.$$

Définition 11.14. Si $\alpha \in \mathbb{N}^m$ est un multi-indice et $y \in \mathbb{R}^m$ un vecteur, on définit :

1. $\mathbb{R} \ni \alpha! \coloneqq \prod_{k=1}^{m} \alpha_k!$;
2. $\mathbb{R} \ni y^\alpha \coloneqq \prod_{k=1}^{n} y_k^{\alpha_k}$.

Définition 11.15. Soient $f : E \subset \mathbb{R}^m \to \mathbb{R}$ une fonction de classe C^k définie sur un ouvert, $a \in E$, et $1 \leqslant r \leqslant k$. On définit le polynôme $P_r(f, a)$ par :

$$P_r(f,a)(x) = \sum_{\alpha \in \mathbb{N}\ \text{t.q.}\ \overline{\alpha} = r} \frac{1}{\alpha!}(x-a)^\alpha \partial^\alpha f(a).$$

Définition 11.16. Soit $f : E \subset \mathbb{R}^m \to \mathbb{R}$ une fonction de classe C^k définie sur un ouvert E. On appelle le *polynôme de Taylor de* f *autour de* a *d'ordre* k le polynôme $T_k(f, a) : E \to \mathbb{R}$ défini par :

$$T_k(f,a)(x) = f(a) + \sum_{i=1}^{k} P_i(f,a)(x).$$

Lemme 11.17. *Soit* $f : B(a, R) \subset \mathbb{R}^m \to \mathbb{R}$ *de classe* C^{k+1}. *Pour tout* $t \in (0, 1), 1 \leqslant r \leqslant k + 1$, *on a :*

$$\frac{1}{r!}\frac{d^r F}{dt^r}(t) = \sum_{\alpha \in \mathbb{N}^m\ \text{t.q.}\ \overline{\alpha} = r} \frac{1}{\alpha!}(x-a)^\alpha \partial^\alpha f(a + t(x-a)).$$

Théorème 11.18 (Théorème de Taylor à n variables)**.** *Soit* $f : E \subset \mathbb{R}^m \to \mathbb{R}$ *de classe* C^{k+1}. *Si pour tout* $a \in \operatorname{int} E, R > 0$, *la boule* $B(a, R) \subset E$, *alors il existe* $C > 0$ *tel que pour tout* $x \in \bar{B}(a, R)$ *:*

$$\left\| f(x) - T_k(f,a)(x) \right\| \leqslant C \|x - a\|^{k+1}.$$

Démonstration. On pose :

$$F(t) = f(a + t(x-a)).$$

On observe que :

$$f(x) - f(a) = F(1) - F(0) = \sum_{r=1}^{k} \frac{d^r F}{dt^r}(0) + R.$$

On sait également qu'il existe $\kappa \in (0, 1)$ tel que :

$$R = \frac{1}{(k+1)!}\frac{d^{k+1}F}{dt^{k+1}}(\kappa).$$

Par le lemme 11.17, on sait que

$$\frac{1}{r!}\frac{d^r F}{dt^r}(0) = P_r(f,a)(x).$$

On peut donc réécrire le reste comme :

$$R(x) = \sum_{\alpha \in \mathbb{N}^m \text{ t.q. } \overline{\alpha}=k+1} \frac{1}{\alpha!}(x-a)\alpha\partial^\alpha f(a+\kappa(x-a)).$$

Par hypothèse, $f \in C^{k+1}$, et donc les fonctions $\partial^\alpha f$ sont continues sur E pour $\overline{\alpha} = k+1$. On peut donc dire que pour tout $x \in \bar{B}(a, R)$, ces fonctions sont bornées. Donc il existe $M > 0$ tel que :

$$\left|\partial^\alpha f(a+t(x-a))\right| < M.$$

On sait également que $\alpha! < 1$ et donc par l'inégalité triangulaire, on peut exprimer :

$$\left|R(x)\right| \leqslant M \sum_{\alpha \in \mathbb{N}^m \text{ t.q. } \overline{\alpha}=k+1} \left|(x-a)\alpha\right|.$$

En observant :

$$|y^\alpha| \leqslant \|y\|^{\overline{\alpha}},$$

on peut alors poser $C \coloneqq MN$ où N est le nombre de multi-indices de longueur $k+1$. On a alors :

$$\left|R(x)\right| \leqslant C\|x-a\|^{k+1}.$$

□

Définition 11.19. On définit la matrice Hessienne de la fonction $f : \mathbb{R}^m \to \mathbb{R}$ de classe C^2 par :

$$H_f(a) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right]_{1 \leqslant i,j \leqslant m}.$$

Remarque. La matrice Hessienne d'une fonction f est symétrique.

De plus, on observe que $T_2(f,a)(x) = f(a) + \left\langle(\nabla f)(a), x-a\right\rangle + \frac{1}{2}(x-a)^T H_f(x-a)$.

11.3 Condition suffisante d'extrémalité

Définition 11.20. Soit $f : E \subset \mathbb{R}^m \to \mathbb{R}$. Un point $a \in E$ est un *maximum local* de f si :

$$\exists \delta > 0 \text{ t. q. } \forall x \in B(a,\delta) \cap E : f(x) \leqslant f(a).$$

Similairement, a est un *minimum local* si :

$$\exists \delta > 0 \text{ t. q. } \forall x \in B(a,\delta) \cap E : f(x) \geqslant f(a).$$

Si a est soit un minimum local, soit un maximum local, on dit que a est un *extremum local*.

Lemme 11.21. *Soient $f : E \subset \mathbb{R}^m \to \mathbb{R}$ et $a \in \operatorname{int} E$. Si f est dérivable en a dans la direction $v \in \mathbb{R}^m$ et si a est un extremum local de f, alors $\partial_v f(a) = 0$.*

Démonstration. Soit $f_v : \mathbb{R} \to \mathbb{R} : t \mapsto f(a+tv)$. On remarque alors que :

$$\frac{df_v}{dt}(0) = \partial_v f(a).$$

De plus, si a est un extremum local de f, alors t est un extremum local de f_v. Donc $\frac{df_v}{dt}(0) = 0$. □

Corollaire 11.22. *Soit* $f : E \subset \mathbb{R}^m \to \mathbb{R}$ *différentiable en* $a \in E$*. Si* a *est un extremum local de* f*, alors* $(\nabla f)(a) = 0$*.*

Démonstration. Par le lemme précédent, on sait que pour tout $v \in \mathbb{R}^m$ où $v \neq 0$, $\partial_v f(a) = 0$. Or on sait, puisque f est différentiable en a, $\frac{\partial f}{\partial v} = \langle (\nabla f)(a), v \rangle$. Il faut donc $(\nabla f)(a) = 0$. □

Définition 11.23. On appelle les points annulant le gradient d'une fonction f les *points critiques* de f.

Remarque. Les extrema locaux sont donc des points critiques d'une fonction f multivariée.
Remarque. L'annulation du gradient est une condition nécessaire mais pas suffisante pour la présence d'un extremum local. Un exemple est la fonction $f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x^2 - y^2$. Son gradient s'annule en $(0, 0)$, or pour tout x, on trouve $f(x, 0) \gneqq f(0, 0) = 0 \gneqq f(0, x)$.

Proposition 11.24. *Soient* $f : I \to \mathbb{R}$ *de classe* C^k *et* $a \in I$*. Supposons* $f^{(n)}(a) = 0$ *pour* $1 \leqslant n < k$ *et* $f^{(k)}(a) \neq 0$*. Alors :*

1. *si* k *est pair et* $f^{(k)}(a) > 0$*, alors* a *est un minimum local de* f *;*
2. *si* k *est pair et* $f^{(k)}(a) < 0$*, alors* a *est un maximum local de* f *;*
3. *si* k *est impair, alors* a *n'est pas un extremum de* f*.*

Définition 11.25. Soit M une matrice symétrique carrée $n \times n$. La forme quadratique :

$$Q_M : \mathbb{R}^n \to \mathbb{R} : x \mapsto x^T M x$$

est dite :

- *définie positive* si $\forall x \neq 0 : Q_M(x) \gneqq 0$;
- *définie négative* si $\forall x \neq 0 : Q_M(x) \lneqq 0$;
- *indéfinie* si elle n'est ni positive ni négative.

Théorème 11.26. *Soit* $f : E \subset \mathbb{R}^n \to \mathbb{R}$ *une fonction de classe* C^3 *définie sur un ouvert. Soit* $a \in E$*, un point critique de* f*. Alors*

1. *Si* $H_f(a)$ *est définie positive, alors* a *est un minimum local de* f *;*
2. *si* $H_f(a)$ *est définie négative, alors* a *est un maximum local de* f *;*
3. *si* $H_f(a)$ *est indéfinie, alors* a *n'est pas un extremum de* f*.*

Démonstration. Par hypothèse, $(\nabla f)(a) = 0$. Donc on peut exprimer l'approximation de Taylor d'ordre 2 de f autour de a par :

$$T_2(f, a)(x) = f(a) + \frac{1}{2}(x - a)^T H_f(a)(x - a),$$

que l'on peut réécrire en :

$$T_2(f, a)(x) - f(a) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x - a)_i (x - a)_j.$$

Par le théorème de Taylor dans $\mathbb{R}^n$, on sait qu'il existe $C > 0, r > 0$ tels que $\bar{B}(a, r) \subset E$ et :

$$\left| f(x) - T_2(f, a)(x) \right| \leqslant C \|x - a\|^3.$$

Quand $\|x-a\| \leqslant 1$, on sait $\|x-a\|^3 \leqslant \|x-a\|^2$. Donc pour $x-a$ suffisamment petit, c'est le comportement de $(x-a)^T H_f(a)(x-a)$ qui domine le comportement de $f(x) - T_2(f,a)(x)$.

Supposons $H_f(a)$ définie positive (le cas définie négative se démontre de la même manière). On pose :

$$S^{n-1} \coloneqq \left\{x \in \mathbb{R}^n \text{ t. q.} \|x\| = 1\right\}.$$

L'ensemble S^{n-1} et fermé borné, et donc par le théorème des bornes atteintes, on peut noter α[17] le minimum de Q restreint à S^{n-1}. On trouve alors pour tout $x \neq 0$:

$$Q(x) = Q\left(\|x\| \frac{x}{\|x\|}\right) = \|x\|^2 Q\left(\frac{x}{\|x\|}\right) \geqslant \|x\|^2 \alpha.$$

Prenons alors $\epsilon > 0$ tel que $\epsilon < \frac{\alpha}{4C}$. Si $\|x-a\| < \epsilon$, on trouve :

$$\left|f(x) - T_2(f,a)(x)\right| \leqslant C\|x-a\|^3 \leqslant C\|x-a\|^2 \frac{\alpha}{4C} = \frac{\alpha}{4}\|x-a\|^2 \leqslant \frac{\alpha}{4} Q(x-a) = \frac{1}{4}(x-a)^T H_f(a)(x-a).$$

Dès lors, prenons $x \in B(a,r)$. On observe :

$$\begin{aligned} f(x) - f(a) &= f(x) - T_2(f,a)(x) + T_2(f,a)(x) - f(a) = f(x) - T_2(f,a)(x) + \frac{1}{2}(x-a)^T H_f(a)(x-a) \\ &\geqslant -\frac{1}{4}(x-a)^T H_f(a)(x-a) + \frac{1}{2}(x-a)^T H_f(a)(x-a) = \frac{1}{4}(x-a)^T H_f(a)(x-a) \geqslant 0. \end{aligned}$$

On en déduit donc que a est un minimum local de f.

Si $H_f(a)$ est indéfinie, on sait qu'il existe $u, v \in \mathbb{R}^n$ tels que :

$$u^T H_f(a) u \gneqq 0 \gneqq v^T H_f(a) v.$$

En raisonnant de la même manière, on trouve, avec $\delta \to 0$:

$$f(a + \delta u) \gneqq f(a) \gneqq f(a + \delta v).$$

On en déduit que a n'est pas un extremum de f. □

[17] Que l'on sait $\gneqq 0$ car $H_f(a)$ est supposée définie positive.

12 Fonctions implicites et réciproques

12.1 Théorème du point fixe de Banach

Définition 12.1. Soit X un ensemble et soit $f : X \to X$. On dit que $x \in X$ est un *point fixe de* f si $f(x) = x$.

Définition 12.2. La fonction $f : \mathbb{R}^m \to \mathbb{R}^n$ est dite *lipschitzienne* s'il existe $L > 0$ tel que :

$$\forall x, y \in \mathbb{R}^m : \left\|f(x) - f(y)\right\| \leqslant L\|x - y\| .$$

Le plus petit $L > 0$ satisfaisant cette condition est appelé la *constante de Lipschitz* de f.

Lemme 12.3. *Toute fonction lipschitzienne est uniformément continue.*

Démonstration. Soit $f : \mathbb{R}^m \to \mathbb{R}^n$ lipschitzienne. On sait donc qu'il existe $L > 0$ tel que :

$$\forall x, y \in \mathbb{R}^m : \left\|f(x) - f(y)\right\| \leqslant L\|x - y\| .$$

Soit $\epsilon > 0$, et posons $\delta = \frac{\epsilon}{L}$. Supposons $\|x - y\| < \delta$. On trouve donc :

$$\left\|f(x) - f(y)\right\| \leqslant L\|y - y\| < L\frac{\epsilon}{L} = \epsilon.$$

$\square$

Définition 12.4. Une fonction de Lipschitz est dite *contractante* si sa constante de Lipschitz est strictement inférieure à 1.

Théorème 12.5 (Théorème du point fixe de Banach)**.** *Une application contractante* $f : E \subset \mathbb{R}^m \to E$ *admet un unique point fixe si* E *est fermé.*

Démonstration. Soit f une telle contraction. On nomme $\alpha \in (0, 1)$ sa constante de Lipschitz. Montrons tout d'abord l'unicité du point fixe. Supposons qu'il existe deux points $x, y \in E$ tels que $F(x) = x$ et $F(y = y)$. On sait :

$$\|x - y\| = \left\|f(x) - f(y)\right\| \leqslant \alpha\|x - y\| ,$$

ce qui est une contradiction car $\alpha \lneqq 1$.

Montrons maintenant que ce point fixe existe. On prend $x_0 \in E$ quelconque, et on définit la suite $(x_n)_{n\in\mathbb{N}^*} \subset E$ par $x_n = F(x_{n-1})$. On remarque que :

$$\|x_{n+1} - x_n\| \leqslant \alpha\|x_n - x_{n-1}\| \leqslant \alpha^2\|x_{n-1} - x_{n-2}\| \leqslant \cdots \leqslant \alpha^n\left|F(x_0) - x_0\right| .$$

On prend ensuite $n > m \in \mathbb{N}^*$ et on observe :

$$\|x_n - x_m\| \leqslant \sum_{k=m+1}^{n} \|x_k - x_{k-1}\| \leqslant \sum_{k=m+1}^{n} \alpha^{k-1}\left\|F(x_0) - x_0\right\| .$$

De plus, la suite des sommes partielles de α^k est convergente, et donc de Cauchy. On prend alors $\epsilon > 0$. On sait qu'il existe $N > 0$ tel que :

$$\forall n \geqslant N : \sum_{k=m+1}^{n} \alpha^{k-1} < \frac{\epsilon}{\left\|F(0) - x_0\right\|}.$$

On trouve donc$|x_n - x_m| < \epsilon$ pour tout $n > m > N$.

Si (x_n) est de Cauchy, alors elle est convergente. Appelons ce point x. De plus :

$$f(x) = f(\lim_{n\to+\infty} x_n) = \lim_{n\to+\infty} f(x_n) = \lim_{n\to+\infty} x_{n+1} = x.$$

x est donc bien un point fixe de f. $\square$

Lemme 12.6. *Si* $g : \bar{B}(0,r) \to \mathbb{R}^n$ *est une application contractante de constante* α *telle que* $g(0) = 0$, *alors* $I + g : \bar{B}(0,r) \to \mathbb{R}^n$ *est injective et la boule* $\bar{B}(0,(1-\alpha)r)$ *est incluse dans son image.*

Démonstration. Montrons que $I + g$ est injective : soient $x, y \in \bar{B}(0,r)$ tels que $x + g(x) = y + g(y)$. On peut réécrire cela par :

$$x - y = g(x) - g(y) \qquad \text{et donc} \qquad \|x - y\| = \left\|g(x) - g(y)\right\|.$$

Or g est contractante. Dès lors, il faut$\|x - y\| = 0$, et donc $x = y$. $I + g$ est donc bien injective.

Maintenant, montrons que la boule $\bar{B}(0,(1-\alpha)r)$ est incluse dans l'image de $I + g$. Pour cela, prenons $y \in \bar{B}(0,(1-\alpha)r)$ et montrons qu'il existe $x \in \bar{B}(0,r)$ tel que $x + g(x) = (I + g)(x) = y$. On cherche donc un point fixe de l'application :

$$x \mapsto y - g(x).$$

On définit $F : \bar{B}(0,r) \to \mathbb{R}^n : x \mapsto y - g(x)$. On peut voir que F est à valeurs dans $\bar{B}(0,r)$ car :

$$\left\|F(x)\right\| \leqslant \|y\| + \left\|g(x)\right\| \leqslant \left|(1-\alpha)r\right| + \alpha|r| = r.$$

De plus, on observe que F est contractante car :

$$\left\|F(x) - F(\hat{x})\right\| = \left\|g(x) - g(\hat{x})\right\| \leqslant \alpha\|x - \hat{x}\|,$$

où $\alpha \lneqq 1$. On peut donc appliquer le théorème de Banach qui garantit l'existence d'un point fixe. On a donc montré que la boule était contenue dans l'image de $I + g$. $\square$

12.2 Théorème de la fonction réciproque

Lemme 12.7. *La préimage d'un ensemble de* $U \subset \mathbb{R}^n$ *ouvert par une fonction* $f : \mathbb{R}^n \to \mathbb{R}^n$ *est un ensemble ouvert de* $V \subset \mathbb{R}^n$ *si et seulement si* f *est continue.*

Théorème 12.8 (Théorème de la fonction réciproque)**.** *Soit* $f : E \subset \mathbb{R}^m \to \mathbb{R}^n$ *de classe* C^1 *et soit* $a \in E$. *Si la matrice* $J_f(a)$ *est inversible, alors :*

- *il existe* $U, V \subset \mathbb{R}^n$ *ouverts tels que* $a \in U$ *et* $f(a) \in V$,
- *et* $f : U \to V$ *est bijective.*

De plus, f^{-1} *est bien définie, est dérivable en* $f(a)$ *et sa matrice jacobienne est donnée par :*

$$J_{f^{-1}}(f(a)) = (J_f(a))^{-1}.$$

Démonstration. Montrons cela par étape, commençons par montrer la formule du jacobien de l'inverse, puis montrons que l'on peut se ramener au cas particulier $a = 0$ et $J_f(0) = I$, et puis démontrons ce cas particulier.

Démonstration de la formule du jacobien de l'inverse On suppose f différentiable en $f(a)$. Par la différentiation d'une composition, on trouve :

$$I = \left[\frac{\partial x_i}{\partial x_j}\right] = [\delta_{ij}] = J_I(a) = J_{f^{-1}\circ f}(a) = J_{f^{-1}}(f(a))J_f(a).$$

On en déduit que $J_{f^{-1}}(f(a)) = (J_f(a))^{-1}$.

Montrons que l'on peut se rapporter au cas particulier On veut montrer qu'il est raisonnable de penser que l'on peut se ramener au cas où $a = 0$, $f(0) = 0$, et $J_f(0) = I$.

Soit f respectant les conditions ci-dessus. On définit $\tilde{f} : E \to \mathbb{R}^n$ telle que :

$$\tilde{f}(x) = f(x) - f(a).$$

On a alors $\tilde{f}(a) = 0$. En supposant le résultat vrai pour $\tilde{f}(a) = 0$, il existe U et $\tilde{V}$ ouverts tels que $a \in U$ et $\tilde{f}(a) \in \tilde{V}$ et $\tilde{f} : U \to \tilde{V}$ est une bijection. On pose $V = \tilde{V} - f(a) = \{v - f(a) \text{ t. q. } v \in \tilde{V}\}$. Dès lors V est ouvert et f met U et V en bijection. De plus, f^{-1} est différentiable en $f(a)$ car $\tilde{f}$ est différentiable en $f(a) = 0$.

De manière similaire, on montre, avec $\bar{f}(x) = f(x + a)$ que l'on peut se ramener au cas $a = 0$.

Maintenant, on pose $\hat{f}$ définie par $\hat{f}(x) = (J_f(0))^{-1}f(x)$. Remarquons que l'application $x \mapsto (J_f(0))^{-1}$ est différentiable, que $\hat{f}(0) = 0$, et que $J_{\hat{f}}(0) = (J_f(0))^{-1}J_f(0) = I$.

On peut donc dire qu'il existe $U, \hat{V}$ ouverts dans $\mathbb{R}^n$ tels que $0 \in U$ et $0 \in \hat{V}$ et $\hat{f} : U \to \hat{V}$ est une bijection. Également, $\hat{f}^{-1}$ est différentiable et son jacobien en $a = 0$ est la matrice identité. On constate alors que f est une bijection entre U et V défini par $J_f(0)\hat{V}$. Et également, f est différentiable en 0.

Montrons le résultat pour les conditions énoncées ci-dessus Maintenant que l'on a montré que le résultat généralisable s'il est vrai pour $a = 0$, $f(a) = 0$ et $J_f(0) = I$, démontrons qu'il l'est.

On pose $g : E \to \mathbb{R}^n$ l'application $x \mapsto f(x) - x$ qui représente la *distance* entre f et I. On y voit bien que :

$$J_g(0) = J_f(0) - J_I(0) = 0.$$

De plus, les dérivées partielles de f sont continues, ce qui implique que les dérivées partielles de g le sont également telles que :

$$\forall 1 \leqslant i, j \leqslant n : \frac{\partial g_i}{\partial x_j}(0) = 0.$$

On en déduit qu'il existe une boule fermée $\bar{B}(0, r) \subset E$ telle que :

$$\forall \xi \in \bar{B}(0, r) : \left|\frac{\partial g_i}{\partial x_j}(\xi)\right|^2 \leqslant \frac{1}{4n^2}.$$

Cela implique :

$$\forall \xi \in \bar{B}(0, r) : \left|(\nabla g_i)(\xi)\right|^2 = \sum_{j=1}^{n}\left|\frac{\partial g_i}{\partial x_j}(\xi)\right|^2 \leqslant \frac{n}{4n^2} = \frac{1}{4n}.$$

Par le théorème fondamental et l'inégalité de Cauchy-Schwartz, on peut exprimer pour $x, y \in \bar{B}(0, r)$:

$$\begin{aligned}\left|g_i(x) - g_j(x)\right| &= \left|\int_0^1 \frac{d}{dt}\left(g_i\left(x + t(y - x)\right)\right) dt\right| = \left|\int_0^1 \left\langle(\nabla g_i)(x + t(y - x)), y - x\right\rangle\right| \\ &\leqslant \int_0^1 \left|\left\langle(\nabla g_i)(x + t(y - x)), y - x\right\rangle\right| dt \leqslant \int_0^1 \left|(\nabla g_i)(x + t(y - x)\right||y - x|\, dt \\ &\leqslant \frac{1}{2\sqrt{n}}|y - x|\,.\end{aligned}$$

Dès lors, on en conclut :

$$\left\|g(x) - g(y)\right\|^2 = \sum_{i=1}^{n} \left|g_i(x) - g_i(y)\right|^2 \leqslant \sum_{i=1}^{n} \left(\frac{1}{2\sqrt{n}}\|y - x\|\right)^2 = \frac{n}{4n}\|y - x\|^2 .$$

On a donc défini une boule de rayon r telle que pour tout x, y dans cette boule, on ait :

$$\left\|g(x) - g(y)\right\| \leqslant \frac{1}{2}|x - y| .$$

On en déduit que g est une contraction de facteur $\alpha = \frac{1}{2}$. Par le lemme 12.6, on sait que $I + g$ est injective sur $\bar{B}(0, r)$ et que son image contient la boule $\bar{B}\left(0, \frac{r}{2}\right)$.

On pose $V = B(0, r)$ et $U = f^{-1}(V)$, ce que l'on peut faire car f^{-1} est bien définie (nous apprend le lemme). $f : U \to V$ est une bijection entre deux ouverts (par le lemme 12.7).

Il reste à vérifier que la fonction inverse f^{-1} est différentiable. Pour cela, calculons :

$$\lim_{x\to 0} \frac{\left\|f^{-1}(x) - f^{-1}(0) - (J_{f^{-1}}(0))^{-1}(x - 0)\right\|}{\|x\|} = \lim_{x\to 0} \frac{\left\|f^{-1}(x) - x\right\|}{\|x\|}.$$

Il faut vérifier que cette limite vaut 0. Il faut donc vérifier que pour toute suite $(x_n) \subset \mathbb{R}$ tendant vers 0 que :

$$\lim_{n\to+\infty} \frac{f^{-1}(x_n) - x_n}{\|x\|_n} = 0.$$

Puisque f est bijective sur U (dans V), on peut travailler sur $y_n \coloneqq f^{-1}(x_n)$. Or on sait f différentiable en 0 de jacobien $J_f(0) = I$. On peut donc écrire :

$$\lim_{n\to+\infty} \frac{\left\|y_n - f(y_n)\right\|}{\|y\|_n} = 0.$$

Puisque $\left\|g(x)\right\| \leqslant \frac{1}{2}\|x\|$ (car $g(0) = 0$), on trouve :

$$\frac{1}{2}\|x\| \leqslant \left\|g(x)\right\| \leqslant \frac{3}{2}\|x\| .$$

Cela veut dire que la suite $\left(\frac{\|y_n\|}{\|x_n\|}\right)_n$ est bornée. Dès lors :

$$\lim_{x\to 0} \frac{f^{-1}(x) - x}{\|x\|} = \lim_{n\to+\infty} \frac{\left\|y_n - f(x_n)\right\|}{\|y_n\|} \frac{\|y_n\|}{\|x_n\|} = 0.$$

En effet, le produit d'une suite bornée avec une suite tendant vers 0 tend vers 0. □

Remarque. Si f est de classe C^k, alors f^{-1} est de classe C^k également.

12.3 Théorème de la fonction implicite

12.4 Règle des multiplicateurs de Lagrange