Université d'Orsay - Paris-Saclay **OLMA203E**
LEU – 1er semestre 2021-2022

Cours de Topologie 1 : topologie et fonctions de plusieurs variables

Table des matières

Préliminaires : Rappels de théorie des ensembles.

Nous commençons par revoir les définitions de base de la théorie des ensemble. Le but est d'acquérir l'intuition du sens de ces notions (notamment par de petits schémas). Rien de neuf ici, mais le langage que nous allons rappeler est utilisé tout au long de ce module : il est essentiel de savoir le manipuler couramment.

1. Sous-ensembles ou parties.

Une *partie* (ou un *sous-ensemble*) de l'ensemble $\mathbb{R}$ des réels est formé de réels vérifiant une propriété plus ou moins particulière : par exemple $\mathbb{R}^+ = \{x \in \mathbb{R}, x \geq 0\}$, $]2;5] = \{x \in \mathbb{R}, 2 < x \leq 5\}$. En général, pour annoncer que l'on considère l'ensemble des réels qui vérifient une certaine propriété $\mathcal{P}$, on utilise $\boxed{\text{accolades-barre}}$ (ou : accolades-virgule), $\{x \in \mathbb{R} \,|\, x \text{ vérifie la propriété } \mathcal{P}\}$. Ainsi $\{x \in \mathbb{R} \,|\, \sin(x) \geq \frac{1}{2}\}$ se lit "l'ensemble des réels x tels que $\sin(x) \geq \frac{1}{2}$".

On peut se représenter schématiquement un sous-ensemble : cela demande un effort de compréhension de la propriété qui définit la partie... La récompense est que, lorsque l'on a bien compris la propriété, on ne fait plus d'erreur sur la partie !

De même, dans un ensemble X quelconque, une *partie* (ou *sous-ensemble*) de X est formé d'éléments de X vérifiant une certaine propriété. Par exemple pour $X =$ l'ensemble des étudiants de l'amphi, je peux considérer $A =$ la partie formée des étudiant.e.s ayant déjà pris l'avion au moins 2 fois dans leur vie, donc A est une partie ou un sous-ensemble de X, on note $A \subset X$. Deux sous-ensembles A, B de X sont égaux si ils ont exactement les mêmes éléments : on écrit alors $\boxed{A = B}$.

Rappelons que pour X un ensemble et p un élément de X, on note $\{p\}$ la partie de X dont le seul élément est p ce genre de partie s'appelle des *singletons*).

Prenons $X = \mathbb{R}^2$, $A = \{(x,y) \in \mathbb{R}^2 \,|\, x^2 \leq 1 \text{ et } y^2 \leq 1\}$, puis $B = \{(x,y) \in \mathbb{R}^2 \,|\, \max(|x|,|y|) \leq 1\}$. Alors A et B sont des parties de X et en fait $A = B$.

Noter que les propriétés qui définissent A et B ne s'énoncent pas de la même façon, et pourtant ces propriétés produisent le même sous-ensemble !

Lorsque chaque élément d'une partie A est automatiquement un élément d'une partie B, on dit que A *est inclus dans* B, noté $\boxed{A \subset B}$.

2. Opérations sur les sous-ensembles.

2.1. Réunion : "ou" / $\exists$.

Définition 2.1. Soit X un ensemble.

(1) Pour A, B deux parties de X la *réunion* de A et de B est $\{x \in X \,|\, x \in A \text{ ou } x \in B\}$. La réunion est notée $A \cup B$.

(2) Si $A_1, A_2, \dots, A_n$ sont n parties de X ($n \in \mathbb{N}$), leur *réunion* est $\{x \in X \,|\, x \in A_1, \text{ ou } x \in A_2, \text{ ou } \dots, \text{ ou } x \in A_n\}$. La réunion est notée $\cup_{i=1}^{i=n} A_i$ (ou alors : $A_1 \cup A_2 \cup \cdots \cup A_n$). On remarque que l'on dispose aussi d'une description alternative :

$$\cup_{i=1}^{i=n} A_i = \{x \in X \,|\, \exists i \in \{1, \dots, n\}, x \in A_i\}.$$

(3) Soit $(A_i)_{i \in I}$ une famille de parties de X. Leur *réunion* est $\{x \in X \,|\, \exists i \in I, x \in A_i\}$. On la note $\cup_{i \in I} A_i$.

Remarque 2.2. On rappelle qu'une *famille de parties de* X correspond à la donnée d'un ensemble I (*ensemble d'indices*) et, pour chaque $i \in I$, une partie A_i de X. Souvent on considère des familles finies, ou dénombrables, c'est à dire que l'on prend $I = \{1, \dots, n\}$ ou $I = \mathbb{N}$.

Exemple 2.3. Ici $I = \mathbb{N}$. Pour $n \in \mathbb{N}$ soit $A_n = \{x \in \mathbb{R}, |x - \frac{1}{2^n}| \leq \frac{1}{1+n^2}\}$. Alors

$$\cup_{n \in \mathbb{N}} A_n = \{x \in \mathbb{R} \text{ tels qu'il existe un entier } n \geq 0 \text{ vérifiant } |x - \frac{1}{2^n}| \leq \frac{1}{1+n^2}\}.$$

2.2. Intersection : "et" / $\forall$.

Définition 2.4. Soit X un ensemble.

(1) Pour A, B deux parties de X l'*intersection* de A et de B est $\{x \in X \,|\, x \in A \text{ et } x \in B\}$. L'intersection est notée $A \cap B$.

(2) Si $A_1, A_2, \dots, A_n$ sont n parties de X ($n \in \mathbb{N}$), leur *intersection* est $\{x \in X \,|\, x \in A_1, \text{ et } x \in A_2, \text{ et } \dots, \text{ et } x \in A_n\}$. L'intersection est notée $\cap_{i=1}^{i=n} A_i$ (ou $A_1 \cap A_2 \cap \cdots \cap A_n$). On remarque que l'on dispose aussi d'une description alternative :

$$\cap_{i=1}^{i=n} A_i = \{x \in X \,|\, \forall i \in \{1, \dots, n\}, x \in A_i\}.$$

(3) Soit $(A_i)_{i \in I}$ une famille de parties de X. Leur *intersection* est $\{x \in X \,|\, \forall i \in I, x \in A_i\}$. On la note $\cap_{i \in I} A_i$.

Exemple 2.5. Ici $X = \mathbb{R}^2$ et $I = \mathbb{R}$. Pour $t \in \mathbb{R}$ soit $A_t = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 y^2 \leq \frac{1}{10^t}\}$. Alors

$$\cap_{t \in \mathbb{R}} A_t = \{(x, y) \in \mathbb{R}^2 \text{ tels que pour tout réel } t \text{ on a } x^2 y^2 \leq \frac{1}{10^t}\}.$$

2.3. Complémentaire : "non".

Définition 2.6. Soit X un ensemble. Soit $A \subset X$ une partie. Alors le *complémentaire de A dans X* est la partie $\{x \in X \,|\, x \notin A\}$ ($x \notin A$ se lit : x n'appartient pas à A). Le complémentaire sera noté $X \setminus A$.

Exemple 2.7. Ici $X = \mathbb{R}$ et $A = \mathbb{Q}$. Alors $X \setminus A$ est l'ensemble des nombres irrationnels.

Remarque 2.8.

3. Opérations utilisant des fonctions.

3.1. Image réciproque.

Définition 3.1. Soient X, Y deux ensembles. Soit $B \subset Y$ une partie. Alors l'*image réciproque de B par f* est simplement $\{x \in X \,|\, f(x) \in B\}$. Autrement dit : $f^{-1}(B)$ est l'ensemble des antécédents par f de tous les points possibles de B.

Remarque 3.2. Attention : la notation f^{-1} ne signifie pas ici "la bijection réciproque" ! En effet ici f^{-1} est appliqué à une partie, pas à un élément. D'ailleurs, on ne suppose pas f bijective.

Exemple 3.3. Ici $X = \mathbb{R} = Y$, $f : \mathbb{R} \to \mathbb{R}$ est la fonction $f(x) = x^2 - 2x$ et $A = [3; +\infty[$. Donc $f^{-1}(A) = \{x \in \mathbb{R} \,|\, x^2 - 2x \geq 3\}$. Or

$$x^2 - 2x \geq 3 \iff (x-1)^2 \geq 4 \iff |x-1| \geq 2 \iff (x-1 \geq 2) \text{ ou } (1-x \geq 2) \iff x \geq 3 \text{ ou } x \leq -1.$$

Ainsi $f^{-1}(A) = [3; +\infty[\cup] - \infty; -1]$.

3.2. **Image directe.**

Définition 3.4. Soient X, Y deux ensembles. Soit $A \subset X$ une partie. Alors l'*image directe de A par f* est $\{y \in Y \,|\, \exists x \in A, f(x) = y\}$. Autrement dit : $f^{-1}(B) = \cup_{x \in A}\{f(x)\}$ est l'ensemble des images par f de tous les points de A.

Exemple 3.5. Ici $X = \mathbb{R} = Y$, $f : \mathbb{R} \to \mathbb{R}$ est la fonction $f(x) = \cos(x)$ et $A = [-\frac{\pi}{2}; \frac{\pi}{2}]$. Alors l'étude de f (et l'utilisation du théorème des valeurs intermédiaires) montre que $f(A) = [0; 1]$.

Première partie 1. **Distance et topologie sur $\mathbb{R}$.**

1. Rappels réels.

1.1. **Description des réels.**

Les réels sont des nombres familiers : la calculatrice nous en donne une très bonne intuition, lorsqu'elle affiche *un nombre avec un signe, qui a une partie entière, puis une partie fractionnaire illimitée et quelconque*. Pour ce semestre, ce descriptif peut être retenu comme une définition des réels !

Nous admettons l'existence de l'ensemble $\mathbb{R}$ dont nous rappelons maintenant les propriétés. Nous n'allons pas ici expliquer comment on peut construire ou définir rigoureusement les réels à partir des nombres élémentaires que sont les naturels, les décimaux ou les rationnels.

(Et pourtant un jour il faudra bien apprendre une construction explicite de ces nombres auxquels nous croyons tous, et en plus vérifier que toutes les constructions raisonnables amènent au "même" ensemble $\mathbb{R}$.)

Tout d'abord, l'ensemble $\mathbb{R}$ des réels contient $\mathbb{N}$ (les entiers naturels), $\mathbb{Z}$ (les entiers relatifs), $\mathbb{D}$ (les nombres décimaux relatifs) et aussi $\mathbb{Q}$ (les rationnels).

Comme $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{D}$ et $\mathbb{Q}$, l'ensemble $\mathbb{R}$ est formé de *nombres*, c'est à dire qu'on peut y effectuer les opérations $+, -, \times, \div$ (mais pas $a \div 0$ bien entendu). Les calculs sur $\mathbb{R}$ généralisent ceux sur $\mathbb{Q}$, et vérifient les mêmes règles bien connues :

$$a+b = b+a, a+0 = a, a+(b+c) = (a+b)+c, \text{ tout réel } a \text{ admet un opposé noté } -a \text{ tel que } a+(-a) = 0$$

$$a \times b = b \times a, a \times 1 = a, a \times (b \times c) = (a \times b) \times c, \text{ tout réel } a \neq 0 \text{ admet un inverse noté } \frac{1}{a} \text{ tel que } a \times \frac{1}{a} = 1$$

$$a \times (b+c) = a \times b + a \times c$$

Comme $\mathbb{Q}$, l'ensemble $\mathbb{R}$ est muni d'un *ordre* $\leq$ (une relation de comparaison, lue "inférieure ou égale") : $\leq$ est défini par $x \leq y \iff y - x$ est de signe $+$.

L'ordre se lit facilement sur l'écriture décimale : si x, y sont des réels ≥ 0, alors on a $x \leq y \iff x = y$, ou bien le premier chiffre de y différent de celui de x lui est supérieur.

Voici les propriétés de $\leq$. D'abord, on peut toujours comparer deux réels a, b : ou bien $a \leq b$ ou bien $b \leq a$, dans ce dernier cas on note aussi $a \geq b$. Lorsqu'on a simultanément $a \leq b$ et $b \leq a$ alors $a = b$. Si $a \leq b$ [$a \geq b$] mais $a \neq b$ on note $a < b$ [$a > b$]. La comparaison est transitive : si $a \leq b$ et $b \leq c$ alors $a \leq c$.

Ensuite l'ordre est compatible avec les opérations algébriques au sens où :

$$a \leq b \Rightarrow a + c \leq b + c \ ; \ a \leq b \text{ et } c \geq 0 \Rightarrow a \times c \leq b \times c \ ; \ a \leq b \iff -b \leq -a$$

Définition 1.1 (intervalles)**.** Soit a un réel. On pose

$[a; +\infty[= \{x \in \mathbb{R}, a \leq x\},]a; +\infty[= \{x \in \mathbb{R} a < x\},]-\infty; a] = \{x \in \mathbb{R}, x \leq a\},]-\infty; a[= \{x \in \mathbb{R}, x < a\}$

Soient a, b deux réels tels que $a \leq b$. On pose

$[a; b] = \{x \in \mathbb{R}, a \leq x \leq b\},]a; b[= \{x \in \mathbb{R}, a < x < b\}, [a; b[= \{x \in \mathbb{R}, a \leq x < b\},]a; b] = \{x \in \mathbb{R}, a < x \leq b\}$

Avec $\mathbb{R} =]-\infty; +\infty[$ et $\emptyset$, ces huit sortes de parties sont appelées *intervalles*. Les quatre derniers intervalles sont dits *bornés*.

Tout comme l'ordre sur $\mathbb{Q}$, l'ordre sur $\mathbb{R}$ est *archimédien*, c'est à dire que : pour tout réel $a \geq 0$ il existe un entier N tel que $a \leq N$. On peut remarquer que ceci implique des propriétés bien utiles :

(1) Pour tout réel a il existe un (unique) entier relatif $m \in \mathbb{Z}$ tel que $m \leq a < m + 1$ (cet entier m est la *partie entière de* a, notée $\lfloor a \rfloor$) (pouvez-vous le déduire d'archimédien ?).

(2) Pour tout réel $a > 0$ il existe un entier $n > 0$ tel que $0 < \dfrac{1}{n} < a$ (pourquoi ?).

(3) Pour deux réels a, b quelconques tels que $a < b$ il existe un rationnel $r = \dfrac{p}{q}$ tel que $a < r < b$ (exercice !).

Définition 1.2 (majoré, minoré, borné)**.** Soit $X \subset \mathbb{R}$ une partie. On dit que X est :

(1) *minorée* par $b \in \mathbb{R}$ si $X \subset [b; +\infty[$

(2) *majorée* par $a \in \mathbb{R}$ si $X \subset]-\infty; a]$

(3) *bornée* s'il existe deux réels a, b tels que $X \subset [a; b]$ (i.e. X est majorée par b et minorée par a)

Quand ils existent les majorants et les minorants ne sont jamais uniques ! On remarquera que les intervalles qu'on a appelé bornés sont effectivement bornés au sens ci-dessus.

Aucune des propriétés énoncées jusqu'à présent ne différencie $\mathbb{R}$ de $\mathbb{Q}$. Voici la propriété qui fait que $\mathbb{R}$ est infiniment plus riche que $\mathbb{Q}$.

Axiome 1.3 (de la borne supérieure)**.** *Soit $E \subset \mathbb{R}$ une partie non vide majorée. Alors l'ensemble de tous les majorants de E possède un plus petit élément.*

Rappelons ce que cela signifie. Pour cela, commençons par noter $\mathrm{Maj}(E)$ l'ensemble de tous les nombres $M \in \mathbb{R}$ tels que $\forall x \in E$, on a $x \leq M$ (un tel M est justement ce qu'on appelle un *majorant* de E). Dire que E est *majorée*, c'est dire que E admet au moins un majorant, autrement dit $\mathrm{Maj}(E) \neq \emptyset$. Pour $E = \emptyset$, E est majoré, on a même $\mathrm{Maj}(E) = \mathbb{R}$: ce n'est pas très intéressant et il n'y a pas de plus petit élément dans $\mathrm{Maj}(E)$.

Si $y - x > 1$, $\exists n \in \mathbb{Z}$ entre x et y

Si $y - x < 1$, Dichotomie.

Mais dans $\mathbb{R}$ dès que E est majorée ET non vide, $\text{Maj}(E)$ contient un élément M_0 qui est le plus petit, c'est à dire tel que pour tout $M \in \text{Maj}(E)$, on a $M_0 \leq M$. Ce nombre M_0 est noté $\sup(E)$, appelé le *supremum* de E, ou plus souvent *borne supérieure* de E. On peut voir que $\text{Maj}(E)$ coïncide avec l'intervalle $[\sup(E); +\infty[$.

Cette propriété entraîne que $\mathbb{R}$ contient beaucoup plus de nombres que $\mathbb{Q}$. Par exemple considérons $E = \{r = \frac{p}{q}, \text{ avec } p, q \in \mathbb{N}, q > 0 \text{ et } p^2 \leq 2q^2\}$. Alors E est non vide $r = 1 \in E$, majoré ($M = 2 \in \text{Maj}(E)$), donc il y a un réel ρ tel que $\rho = \sup(E)$, et on peut montrer que $\rho^2 = 2$. Bref ρ est le réel habituellement noté $\sqrt{2}$, et ce n'est pas un rationnel !

En passant à l'opposé on déduit aisément de l'axiome de la borne supérieure le résultat suivant (avec des définitions correspondantes) :

Axiome 1.4 (de la borne inférieure). *Soit $E \subset \mathbb{R}$ une partie non vide minorée. Alors l'ensemble de tous les minorants de E possède un plus grand élément.*

Ce plus grand élément est noté $\inf(E)$, appelé l'*infimum* de E, ou *borne inférieure* de E.

Par exemple $\sup([2; 3[) = 3$ et $\inf([2; 3[) = 2$. Dans ce cas : l'inf est un <u>minimum</u>, puisqu'il appartient à la partie considérée. Mais le sup n'est pas un <u>maximum</u>.

Savoir que le sup existe ne permet pas de le déterminer, pour cela on rappelle la caractérisation suivante :

Lemme 1.5 (caractérisation du sup). *Soit $X \subset \mathbb{R}$ une partie non vide majorée et soit $M \in \mathbb{R}$ fixé. Alors $M = \sup(X) \iff M$ est un majorant de X et $\forall \varepsilon > 0$, $\exists x \in X$ tel que $M - \varepsilon < x \leq M$.*

Démonstration. $M \neq \sup(X) \iff M$ n'est pas un majorant, ou M est un majorant mais pas le plus petit. □

Proposition 1.6 (caractérisation des intervalles par les sous-intervalles fermés). *Soit $I \subset \mathbb{R}$. Alors :*

I est un intervalle $\iff$ pour deux nombres $u, v \in I$ quelconques tels que $u \leq v$, on a $[u; v] \subset I$.

Démonstration. $\Rightarrow$ est évident au cas par cas. Par exemple si $I = [a; b[$ et $c, d \in I$ alors $a \leq c \leq d < b$ (mettons), donc $[c; d] \subset I$. Montrons $\Leftarrow$. On peut supposer $I \neq \emptyset$, donc que I contient un réel a. Si I n'est pas majorée alors il existe pour tout entier $n \geq a$ un élément $a_n \in I$ tel que $a_n \in I$. Alors $[a; a_n] \subset I$ et donc $[a; n] \subset I$. Ceci étant vrai pour tout $n \geq a$, on en déduit que $[a; +\infty[\subset I$. De même si I n'est pas minoré on obtient que $]-\infty; a] \subset I$.

Supposons I majorée. Alors I est non vide (contient a) et est majoré, donc I admet une borne supérieure. En d'autres termes il existe un réel M tel que tout $x \in I$ est $\leq M$ (M est UN majorant de I) et pour tout $\varepsilon > 0$, l'ensemble I contient un x tel que $M - \varepsilon < x \leq M$. On en déduit que $I \subset]-\infty; M]$ et que $[a; M[\subset I$. Ou bien $M \in I$, et alors $I \cap [a; +\infty[= [a; M]$. Ou bien $M \notin I$ et alors $I \cap [a; +\infty[= [a; M[$.

On peut faire le même raisonnement si I est minorée : pour $m = \inf(I)$ on obtient : ou bien $m \in I$, et alors $I \cap]\infty; a] = [m; a]$, ou bien $m \notin I$, et alors $I \cap]\infty; a] =]m; a]$.

A partir de là on conclut en utilisant le fait que $(m; a] \cup [a; M) = (m; M)$ (tous les choix possibles). □

On peut remarquer que la proposition 1.6 est fausse dans l'ensemble $\mathbb{Q}$, comme le montre la partie $X = \{x \in \mathbb{Q}, x^2 < 2\}$. Clairement $X =]-\sqrt{2}; \sqrt{2}[\cap \mathbb{Q}$, donc on ne pourra pas trouver deux rationnels s, t tels que $X = \{x \in \mathbb{Q}, s < x < t\}$. Et pourtant X vérifie bien la propriété de stabilité par intervalles (rationnels) fermés $[u; v]$.

1.2. **Valeur absolue et distance entre deux réels.**

La fonction *valeur absolue* $|\cdot| : \mathbb{R} \to \mathbb{R}$ est définie comme suit : si $u \geq 0$ on pose $|u| = u$, sinon on pose $|u| = -u$ (qui est ≥ 0 !). Des propriétés de l'ordre on déduit immédiatement :

(1) Pour tout $u \in \mathbb{R}$ on a $|u| \geq 0$. On a $|u| = 0 \iff u = 0$.
(2) Pour tout $u \in \mathbb{R}$ on a $|-u| = |u|$.
(3) Pour $u, v \in \mathbb{R}$ on a $|u+v| \leq |u| + |v|$ (avec égalité si et seulement si u et v sont de même signe).

On remarque alors que majorer la valeur absolue d'un nombre revient à encadrer ce nombre :

Lemme 1.7.

Pour $x \in \mathbb{R}$ et $M \geq 0$ on a $|x| \leq M \iff -M \leq x \leq M$ et $|x| < M \iff -M < x < M$.

Démonstration. En effet $|x| \leq M \iff \big(x \geq 0$ et $x \leq M\big)$ ou $\big(x \leq 0$ et $-x \leq M$, soit $-M \leq x\big)$, ce qui permet de conclure (même démarche avec les inégalités strictes). □

Corollaire 1.8. *Soit $X \subset \mathbb{R}$. Alors X bornée $\iff \exists M \geq 0$ tel que $\forall x \in X, |x| \leq M$.*

Démonstration. Car pour $a \leq b$, en posant $M = \max(|a|, |b|)$, on a toujours $[a; b] \subset [-M; M]$. □

Définition 1.9. La *distance* entre deux réels u et v est le nombre $d(u, v)$ défini par $d(u, v) = |v - u|$.

En utilisant les propriétés de $|\cdot|$ on obtient :

Lemme 1.10.

(1) La distance est définie-positive : $d(u, v) \geq 0$ et $d(u, v) = 0 \iff u = v$.
(2) La distance est symétrique : $d(u, v) = d(v, u)$.
(3) La distance vérifie l'inégalité triangulaire : $d(u, v) \leq d(u, w) + d(w, v)$ (avec égalité si et seulement si $w \in [u; v]$ (si $u \leq v$) ou $w \in [v; u]$ (si on est dans le cas où $v \leq u$)).

Lemme 1.11. *Soit $X \subset \mathbb{R}$. Alors X bornée $\iff$ $\{d(0, x), x \in X\}$ est borné (majoré) $\iff$ $\{d(x, x'), x, x' \in X\}$ est borné (majoré).*

Lorsque $X \subset \mathbb{R}$ est bornée non vide, on appelle diamètre *de X le réel* $\sup(\{d(x, x'), x, x' \in X\})$, *noté diam$(X)$.*

Si E n'est pas bornée on pose $\mathrm{diam}(E) = +\infty$. Enfin la partie vide est bornée, mais son diamètre n'est pas défini.

Démonstration. Les ensembles de nombres considérés sont dans $\mathbb{R}^+$, donc pour ces ensembles : borné $\iff$ majoré. La première équivalence est donnée par le Corollaire 1.8. L'implication " $\{d(0, x), x \in X\}$ borné " $\Rightarrow$ " $\{d(x, x'), x, x' \in X\}$ borné " découle de ce que $d(x, x') \leq d(x, 0) + d(0, x') = d(0, x) + d(0, x')$. Réciproquement si $\{d(x, x'), x, x' \in X\}$ est borné soit M un majorant de $\{d(x, x'), x, x' \in X\}$: alors pour tout $x_0 \in X$ fixé on aura $d(0, x) \leq d(0, x_0) + d(x_0, x) = |x_0| + d(x_0, x) \leq |x_0| + M$ et donc $\{d(0, x), x \in X\}$ est borné. □

Définition 1.12. Soit $r > 0$ donné. Nous dirons que deux réels u et v sont *r-proches* si $d(u, v) < r$.

Le résultat suivant est immédiat, il fait le lien entre une condition d'analyse (la proximité) et un lieu géométrique (l'intervalle) :

Lemme 1.13. *Soit $a \in \mathbb{R}$ un réel fixé, et soit $r > 0$ un nombre strictement positif fixé. Alors l'ensemble des nombres r-proches de a est l'intervalle $]a-r; a+r[$, qu'on appelle l'*intervalle ouvert de centre a et de rayon r.

Démonstration. x est r-proche de $a \iff |x-a| < r \iff -r < x-a < r \iff a-r < x < a+r$. □

Remarque 1.14. Soient b, c deux réels tels que $b < c$. Alors $]b; c[$ est l'intervalle $]a-r; a+r[$ en prenant $a = \dfrac{b+c}{2}$ et $r = \dfrac{c-b}{2}$.

L'inégalité triangulaire est l'outil fondamental pour obtenir des majorations de distances, comme l'illustrent les deux propositions suivantes.

Proposition 1.15 (Valeurs approchées et opérations algébriques)**.** *Soit $r > 0$ un réel fixé. Supposons que x est r-proche de a et que y est r-proche de b. Alors :*

(1) La somme $x+y$ est $2r$-proche de $a+b$. La différence $x-y$ est $2r$-proche de $a-b$.

(2) Le produit xy est r'-proche de ab, pour $r' = (|a|+|b|+r)r$.

(3) Si $b \neq 0$ et $r < |b|$ alors $y \neq 0$ et de plus le quotient $\dfrac{x}{y}$ est r''-proche de $\dfrac{a}{b}$, avec $r'' = \dfrac{(|a|+|b|)r}{|b|(|b|-r)}$.

Démonstration. (1) On a $|(x+y)-(a+b)| = |(x-a)+(y-b)| \leq |x-a|+|y-b| < r+r = 2r$.

(2) On a $|xy-ab| = |(xy-xb)+(xb-ab)| \leq |xy-xb|+|xb-ab| = |x||y-b|+|b||x-a|$. Or en écrivant $x = x-a+a$ on obtient $|x| \leq |x-a|+|a| < r+|a|$. On obtient donc

$$|xy-ab| \leq |x||y-b|+|b||x-a| < (r+|a|)r+|b|r = (|a|+|b|+r)r$$

(3) D'abord en écrivant $|b| = |(b-y)+y|$ on obtient $|b| \leq |b-y|+|y|$. Or $|b-y| < r$, donc on obtient $|b| < r+|y|$, ce qui nous donne $|y| > |b|-r$, qui est > 0 par le choix de r. Donc $y \neq 0$ comme annoncé.

Ensuite $|\dfrac{1}{y}-\dfrac{1}{b}| = |\dfrac{b-y}{yb}| < \dfrac{r}{|b|(|b|-r)}$. Donc $|\dfrac{a}{y}-\dfrac{a}{b}| < \dfrac{|a|r}{|b|(|b|-r)}$. Enfin

$$|\frac{x}{y}-\frac{a}{b}| = |(\frac{x}{y}-\frac{a}{y})+(\frac{a}{y}-\frac{a}{b})| \leq |\frac{x}{y}-\frac{a}{y}|+|\frac{a}{y}-\frac{a}{b}| < \frac{1}{|y|}|x-a|+\frac{|a|r}{|b|(|b|-r)} < \frac{r}{|b|-r}+\frac{|a|r}{|b|(|b|-r)}$$

ce qui donne la majoratoin annoncée en réduisant au même dénominateur.

□

Proposition 1.16 (valeurs approchées à la queue leu-leu)**.**

Si a est r-proche de b et b est s-proche de c, alors a et c sont $(r+s)$-proches.

Si a est r-proche de b, b est s-proche de c et c est t-proche de d, alors a et d sont $(r+s+t)$-proches.

Plus généralement si a_0 est r_1-proche de a_1, a_1 est r_2-proche de a_2, $\ldots$, a_{n-1} est r_n-proche de a_n, alors a_0 et a_n sont r-proches pour $r = r_1+\cdots+r_n$.

Démonstration. On démontre par récurrence sur $n \geq 1$ que, si a_i, a_{i+1} sont r_{i+1}-proches pour $i = 0, \ldots, n-1$, alors $|a_0-a_n| \leq r_1+\cdots+r_n$. Pour $n = 1$ il n'y a rien à démontrer.

Supposons la propriété vraie au rang n. Alors en utilisant l'inégalité triangulaire et l'hypothèse de récurrence au rang n :

$$|a_0 - a_{n+1}| = |(a_0 - a_n) + (a_n - a_{n+1})| \leq |a_0 - a_n| + |a_n - a_{n+1}| \leq (r_1 + \cdots + r_n) + r_{n+1}$$

ce qui démontre la propriété au rang $n + 1$. □

2. Suites.

Définition 2.1 (suites réelles)**.** Une *suite de réels* (ou : une *suite numérique réelle*) est une application $u : \mathbb{N} \to \mathbb{R}$. Le *k-ième terme* de la suite est $u(k)$, que l'on note plutôt u_k. Et l'usage est de noter la suite plutôt $(u_0, u_1, \ldots)$, ou $(u_n)_{n\in\mathbb{N}}$, ou $(u_n)_{n\geq 0}$.

On considère aussi des suites définies sur des parties $I \subset \mathbb{N}$: ce sont des fonctions $u : I \to \mathbb{R}$. Par exemple $u : \{1, 2, \cdots\} = \mathbb{N}^* \to \mathbb{R}$ définie par $u_n = \dfrac{1}{n}$.

En mathématiques on généralise cette notion de suites :

Définition 2.2 (suites d'un ensemble)**.** Soit X un ensemble quelconque. Une *suite de* X est une application $u : \mathbb{N} \to X$. Le *k-ième terme* de la suite est $u(k)$, que l'on note plutôt u_k (on parle aussi du *terme de rang* k). On note la suite plutôt $(u_0, u_1, \ldots)$, ou $(u_n)_{n\in\mathbb{N}}$, ou $(u_n)_{n\geq 0}$. (On considère aussi des suites de X définies sur des parties $I \subset \mathbb{N}$: ce sont des fonctions $u : I \to X$.)

Une suite $(x_n)_n$ est *constante* si $x_0 = x_1 = x_2 = \cdots = x_m = \cdots$.

Une suite $(x_n)_n$ est *stationnaire* si elle est constante à partir d'un certain rang, c'est à dire qu'il existe un entier $k \geq 0$ tel que $x_k = x_{k+1} = x_{k+2} = \cdots = x_{k+m} = \cdots$.

L'ensemble des *valeurs* d'une suite $(u_n)_{n\geq 0}$ est l'ensemble $\{u_0, u_1, \ldots, u_n, \ldots\} \subset X$. En d'autre termes l'ensemble des valeurs de la suite $u : \mathbb{N} \to X$ est simplement l'image directe $u(\mathbb{N})$. Rappelons que l'ensemble des valeurs ne détermine pas du tout la suite $u : \mathbb{N} \to X$: pour que deux suites $u : \mathbb{N} \to X$ et $v : \mathbb{N} \to X$ soient égales, il faut non seulement qu'elles prennent les mêmes valeurs, mais en plus qu'elle prenne une valeur donnée x aux mêmes rangs n ($u^{-1}(\{x\}) = v^{-1}(\{x\})$).

Pour étudier les suites réelles on dispose d'un outil qui n'existe pas dans n'importe quel ensemble X : la distance.

Exemple 2.3 (suites réelles majorées, minorées, bornées)**.** Soit $(x_n)_n$ une suite réelle. On dit que $(x_n)_n$ est *majorée* / *minorée* / *bornée* lorsque l'ensemble de ses valeurs est majoré / minoré / borné, ce qui revient à dire :

$$\exists M \in \mathbb{R}, \forall n \geq 0, u_n \leq M \ / \ M \leq u_n \ / \ -M \leq u_n \leq M.$$

Les suites réelles constantes ou stationnaires sont très particulières. Mais on obtient une immense classe de suites très intéressantes, si l'on considère les suites "de plus en plus (égales à une) constante" :

Définition 2.4 (suite convergente)**.** Soit $(x_n)_n$ une suite réelle.

(1) On dit que *la suite* $(x_n)_n$ *tend vers* 0 lorsque

pour tout $\varepsilon > 0$ il existe un rang $N \in \mathbb{N}$ tel que pour tout $n \geq N$ on a $|x_n| < \varepsilon$

(soit : "x_n est de plus en plus proche de 0"). On note $x_n \to 0$.

(2) On dit que la suite $(x_n)_n$ est *convergente* s'il existe un réel ℓ tel que la suite des distances $(|x_n - \ell|)_n$ tend vers 0 (soit : "x_n est de plus en plus proche de ℓ").

Rappelons quelques faits :

Proposition 2.5. *(1) Si la suite $(x_n)_n$ tend vers 0 alors $(x_n)_n$ est convergente [en prenant $\ell = 0$].*

(2) (Unicité de la limite) Si $(x_n)_n$ est convergente alors il existe un unique réel ℓ tel que $|x_n - \ell| \to 0$. Ce réel ℓ est appelé la limite de la suite $(x_n)_n$, on note $\ell = \lim_{n\to\infty} x_n$, ou encore $x_n \to \ell$. Par définition on a donc :

$$\lim_{n\to\infty} x_n = \ell \iff \forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tel que pour tout } n \geq N, \text{ on a } |x_n - \ell| < \varepsilon.$$

(3) Si la suite $(x_n)_n$ est constante avec $x_n = \ell$ pour tout $n \geq 0$ alors $x_n \to \ell$. Plus généralement si la suite $(x_n)_n$ est stationnaire avec $x_n = \ell$ pour tout n suffisamment grand, alors $x_n \to \ell$.

(4) Si la suite $(x_n)_n$ est convergente alors elle est bornée.

(5) Si la suite $(x_n)_n$ est croissante et majorée alors elle est convergente. De plus dans ce cas la limite de $(x_n)_n$ est $\sup\{x_n, n \in \mathbb{N}\}$.

(6) Supposons que $(x_n)_n$ et $(y_n)_n$ sont deux suites convergentes, avec $\lim_{n\to\infty} x_n = a$ et $\lim_{n\to\infty} y_n = b$. Alors :

(a) La suite somme $(x_n + y_n)_n$ est convergente et $\lim_{n\to\infty} x_n + y_n = a + b$.

(b) La suite produit $(x_n y_n)_n$ est convergente et $\lim_{n\to\infty} x_n y_n = ab$.

indications de preuve. (1) Evident.

(2) Si ℓ, ℓ' vérifient la définition de la convergence pour $(x_n)_n$ alors $d(\ell, \ell') \leq d(\ell, x_n) + d(x_n, \ell')$. Alors pour tout $\varepsilon > 0$ on peut prendre n tellement grand que $d(\ell, x_n) < \frac{1}{2}\varepsilon$ et $d(x_n, \ell') < \frac{1}{2}\varepsilon$. Cela donne $d(\ell, \ell') < \varepsilon$, et ce pour tout $\varepsilon > 0$. On en déduit que $d(\ell, \ell') \leq 0$, donc $d(\ell, \ell') = 0$ et $\ell = \ell'$.

(3) Evident.

(4) Si $x_n \to \ell$, alors $\exists N_1$ tel que $n \geq N_1 \Rightarrow d(x_n, \ell) \leq 1$. A fortiori, comme $d(0, x_n) \leq d(0, \ell) + d(\ell, x_n)$, on aura $d(0, x_n) \leq |\ell| + 1$ pour tout $n \geq N_1$. Et finalement $\{|x_n|, n \in \mathbb{N}\}$ est majoré par le plus grand des $N_1 + 1$ nombres $|x_0|, |x_1|, \ldots, |x_{N_1 - 1}|, |\ell| + 1$.

(5) Supposons $(x_n)_n$ croissante majorée, et posons alors $M = \sup\{x_n, n \in \mathbb{N}\}$. Nous savons alors que $\forall n \geq 0$ on a $x_n \leq M$ et de plus $\forall \varepsilon > 0\ \exists N \in \mathbb{N}$ tel que $M_\varepsilon < x_N \leq M$. Mais comme $(x_n)_n$ est croissante en fait pour tout $n \geq N$, on aura aussi $M_\varepsilon < x_n \leq M$, de sorte que $|M - x_n| < \varepsilon$.

(6) (a) On a $(x_n + y_n) - (a + b) = (x_n - a) + (y_n - b)$, donc $|(x_n + y_n) - (a + b)| \leq |x_n - a| + |y_n - b|$. Fixons $\varepsilon > 0$. Alors pour n suffisamment grand on aura à la fois $|x_n - a| < \frac{1}{2}\varepsilon$ et $|y_n - b| < \frac{1}{2}\varepsilon$, de sorte que $|(x_n + y_n) - (a + b)| < \varepsilon$.

(b) Pour le produit $x_n y_n$ on peut écrire $x_n y_n - ab = x_n y_n - a y_n + a y_n - ab$ et donc $|x_n y_n - ab| \leq |x_n y_n - a y_n| + |a y_n - ab|$. On a $|x_n y_n - a y_n| = |y_n||x_n - a|$ et $|a y_n - ab| = |a||y_n - b|$. Comme $(y_n)_n$ converge elle est bornée, mettons par M. Alors $|x_n y_n - ab| \leq M|x_n - a| + a|y_n - b|$. On se retrouve à ajouter deux suites qui tendent vers 0 ($|x_n - a|$ et $|y_n - b|$), multipliées par deux constantes (M et a), ce qui permet de conclure.

$\square$

Intuitivement : dire que $x_n \to \ell$ c'est dire que $(x_n)_n$ est de plus en plus stationnaire : si on remplaçait tous les termes de x_n par la constante ℓ pour n assez grand, on ne commettrait pas une grande erreur !

Converger vers ℓ, c'est être de plus en plus proche de ℓ : mais comment savoir qu'une suite converge quand on ne connaît pas la limite ℓ a priori ? Il y a bien un critère (dit "de Cauchy"), il utilise une formulation de "presque stationnaire" utilisant la distance :

Définition 2.6 (suite de Cauchy)**.** Soit $(x_n)_n$ une suite réelle. On dit que $(x_n)_n$ est une suite *de Cauchy* si elle vérifie la propriété suivante :

$$\text{Pour tout } \varepsilon > 0, \text{ il existe un rang } N \text{ tel que pour } p \geq N \text{ et } q \geq N \text{ quelconques, on a } |x_p - x_q| < \varepsilon.$$

Théorème 2.7. *Soit $(x_n)_n$ une suite réelle. Alors les propriétés suivantes sont équivalentes :*

(1) $(x_n)_n$ est convergente.

(2) $(x_n)_n$ est de Cauchy.

Démonstration. Nous démontrons seulement "Convergente $\Rightarrow$ Cauchy", la réciproque sera étudiée l'an prochain !

Supposons donc $(x_n)_n$ convergente, avec $x_n \to \ell$. Commençons par un argument rédigé vaguement, mais correct. Comme $x_n \to \ell$, quel que soit n choisi suffisamment grand, le réel x_n est très proche de ℓ. Donc si p et q sont choisis également très grands les réels x_p et x_q sont très proches du même nombre ℓ : ils sont donc très proches l'un de l'autre, ce qui correspond à la définition d'être de Cauchy.

Passons à la preuve formalisée. Pour montrer que $(x_n)_n$ est de Cauchy fixons-nous un $\varepsilon > 0$ arbitraire. Comment être sûr que $|x_p - x_q| < \varepsilon$? Comme $x_n \to \ell$ nous savons que pour tout $\varepsilon' > 0$ il existe un rang N tel que pour tout $n \geq N$ le réel x_n est ε'-proche de ℓ. Alors pour des entiers p, q tous les deux $\geq N$, nous aurons que x_p et x_q sont $2\varepsilon'$-proches. Nous voyons alors qu'il nous suffit de prendre $\varepsilon' = \dfrac{1}{2}\varepsilon$ pour assurer la majoration $|x_p - x_q| < \varepsilon$. □

3. Topologie de $\mathbb{R}$.

3.1. Parties ouvertes, parties fermées.

Définition 3.1 (voisinage)**.** Soit V une partie de $\mathbb{R}$ et soit $a \in V$.

On dit que V *est voisinage de* a s'il existe $r > 0$ tel que V contient $]a - r; a + r[$. En d'autres termes : V est voisinage de a lorsque V contient tous les réels r-proches de a (pour un certain $r > 0$).

Exemple 3.2. L'intervalle $[1; 2[$ n'est pas voisinage de 1, mais est voisinage de $1,5$ ou $1,9$.

En effet quel que soit $r > 0$ l'intervalle $]1-r; 1+r[$ n'est pas contenu dans $[1; 2[$, puisque $1 - \dfrac{r}{2}$ n'est pas dans $[1; 2[$. Pour $r_1 = \dfrac{1}{2}$ on a bien $]1,5 - r_1; 1,5 + r_1[\subset [1; 2[$. Enfin $]1,9 - r_1; 1,9 + r_1[\not\subset [1; 2[$, mais si on prend un rayon plus petit, mettons $r_2 = \dfrac{1}{10}$, alors on retrouve une inclusion valide : $]1,9 - r_2; 1,9 + r_2[\subset [1; 2[$.

Définition 3.3 (ouvert, fermé)**.** Soit O une partie de $\mathbb{R}$.

On dit que O *est une partie ouverte, ou un ouvert* si O est voisinage de chacun de ses points. Autrement dit si pour tout a de O, il existe r $(= r_a)$ tel que $]a - r; a + r[\subset O$.

Soit F une partie de $\mathbb{R}$.

On dit que F *est une partie fermée, ou un fermé,* si $\mathbb{R} \setminus F$ est ouvert. Autrement dit si pour tout $a \notin F$, il existe r $(= r_a)$ tel que $]a - r; a + r[\cap F = \emptyset$.

La *topologie* de $\mathbb{R}$ est simplement la famille de toutes les parties ouvertes de $\mathbb{R}$.

Attention à l'erreur classique : fermé = tout ce qui n'est pas ouvert ! C'est une confusion sur l'usage du mot "complémentaire". L'exemple de $[1; 2[$ ci-dessus donne une partie qui n'est ni ouverte, ni fermée.

Lemme 3.4 (le passage au complémentaire est une involution)**.** *Soit $X \subset \mathbb{R}$ une partie quelconque.*

Alors X est ouverte $\iff$ $\mathbb{R} \setminus X$ est fermée.

Démonstration. Posons $Y = \mathbb{R} \setminus X$. Alors par définition : Y est fermée $\iff$ $\mathbb{R} \setminus Y$ est ouverte. Or $\mathbb{R} \setminus Y = \mathbb{R} \setminus \big(\mathbb{R} \setminus X\big) = X$. □

Lemme 3.5. *(1) $\mathbb{R}$ et $\emptyset$ sont ouverts et fermés.*

(2) Les intervalles ouverts $]-\infty; b[$, $]a; b[$, $]a; +\infty[$ sont ouverts.

(3) Les intervalles fermés $]-\infty; b]$, $[a; b]$, $[a; +\infty[$ sont fermés.

Démonstration. Le vide et $\mathbb{R}$ sont ouverts, puisqu'ils sont évidemment voisinages de chacun de leurs points. Par passage au complémentaire $\emptyset$ et $\mathbb{R}$ sont également fermés !

Montrons que $]a; b[$ est ouvert (les cas $]-\infty; b[$ et $]a; +\infty[$ sont semblables et plus simples !) Soit donc $x \in]a; b[$ (ce que l'on visualise par un dessin...)

Ou bien x est plus proche de a que de b, ou bien x est plus proche de b que de a. En formule cela signifie : ou bien $|b - x| \leq |x - a|$ (soit $b - x \leq x - a$), ou bien $|b - x| \geq |x - a|$ (soit $b - x \geq x - a$). Dans le premier cas on pose $r = b - x > 0$, dans le deuxième cas on pose $r = x - a > 0$. Alors dans tous les cas $]x - r; x + r[\subset]a; b[$ puisque $a \leq x - r \leq x + r \leq b$.

Les complémentaires de $]-\infty; b]$ et $[a; +\infty[$ sont des intervalles ouvertes, donc sont ouverts. Enfin $\mathbb{R} \setminus [a; b] =]-\infty; a[\cup]b; +\infty[$, qui est clairement ouvert. □

Proposition 3.6 (propriétés ensemblistes)**.**

(1) Une réunion quelconque d'ouverts est un ouvert. Une intersection finie *d'ouverts est un ouvert.*

(2) Une réunion finie *de fermés est un fermé. Une intersection quelconque de fermés est un fermé.*

Démonstration. Soit $(O_i)_{i \in I}$ une famille quelconque d'ouvert, posons $O = \cup_{i \in I} O_i$ et soit $x \in O$. Il existe donc un $i_0 \in I$ tel que $x \in O_{i_0}$. Comme O_{i_0} est ouvert il existe $r > 0$ tel que $]x - r; x + r[\subset O_{i_0}$. Comme $O_{i_0} \subset O$ on a aussi $]x - r; x + r[\subset O$ et O est voisinage de x.

Soit $\{O_1, \ldots, O_n\}$ un ensemble fini d'ouverts, posons $O = O_1 \cap \cdots \cap O_n$ et soit $x \in O$. Pour chaque $i = 1, \ldots, n$ on a $x \in O_i$, donc comme O_i est ouvert il existe $r_i > 0$ tel que $]x - r_i; x + r_i[\subset O_i$. Soit r le minimum de l'ensemble fini $\{r_1, \cdots, r_n\}$, on a $r > 0$ puisque r est l'un des r_i. Pour tout $i = 1, \ldots, n$ on a $]x - r; x + r[\subset]x - r_i; x + r_i[$, donc $]x - r; x + r[\subset O_i$. Ainsi $]x - r; x + r[\subset \cap_i O_i$.

Les propriétés sur les fermés découlent des propriétés sur les ouverts grâce aux propriétés du passage au complémentaire :

$$\mathbb{R} \setminus \big(\cap_{i \in I} F_i \big) = \cup_{i \in I} \big(\mathbb{R} \setminus F_i\big) \text{ et } \mathbb{R} \setminus \big(\cup_{i=1}^{i=n} F_i \big) = \cap_{i=1}^{i=n} \big(\mathbb{R} \setminus F_i\big)$$

□

3.2. **Suites convergentes et topologie.**

Lemme 3.7. *Soit $O \subset \mathbb{R}$ un ouvert. Soit $(x_n)_n$ une suite convergente, soit ℓ la limite, et supposons que $\ell \in O$. Alors pour n assez grand $x_n \in O$ [$\exists N$ tel que pour $n \geq N$ on a $x_n \in O$].*

Démonstration. En effet O est voisinage de ℓ, donc $\exists r > 0$ tel que $]\ell - r; \ell + r[\subset O$. Mais pour n assez grand x_n est r-proche de ℓ, c'est à dire est dans $]\ell - r; \ell + r[$, donc a fortiori dans O. $\square$

En utilisant les suites on obtient une caractérisation intrinsèque des fermés, c'est à dire qui n'utilise plus les ouverts (et ne nécessite donc pas le passage au complémentaire) :

Proposition 3.8 (caractérisation séquentielle des fermés)**.**

Soit F une partie de $\mathbb{R}$. Alors les propriétés suivantes sont équivalentes :

(1) F est un fermé.

(2) Pour toute suite $(x_n)_n$ qui d'une part est à valeur dans F et qui d'autre part est convergente, on a $\lim_{n\to\infty} x_n \in F$.

Démonstration. (1) $\Rightarrow$ (2) Supposons F fermé, donc $\mathbb{R} \setminus F$ est ouvert. Soit $(x_n)_n$ une suite de F qui est convergente, notons ℓ la limite. Supposons par l'absurde $\ell \notin F$, soit ℓ appartient à l'ouvert $\mathbb{R} \setminus F$. Alors d'après le Lemme 3.7 pour n assez grand on aurait $x_n \in \mathbb{R} \setminus F$, ce qui est absurde.

(2) $\Rightarrow$ (1), ou, ce qui revient au même : $\neg(1) \Rightarrow \neg(2)$. Supposons donc que F n'est pas fermé. Donc $\mathbb{R} \setminus F$ n'est pas ouvert, et $\mathbb{R} \setminus F$ contient au moins un point y dont il n'est pas voisinage. Alors pour tout entier $n \in \mathbb{N}$ l'intervalle $]y - \frac{1}{n+1}; y + \frac{1}{n+1}[$ n'est pas contenu dans $\mathbb{R} \setminus F$, ce qui signifie exactement que F contient un point x_n dans $]y - \frac{1}{n+1}; y + \frac{1}{n+1}[$. La suite $(x_n)_n$ est formée d'éléments de F, et elle tend vers y avec $y \notin F$. $\square$

3.3. **Complément : densité.**

Cette partie n'est pas au programme : elle est là pour ceux qui veulent déjà aller plus loin.

Définition 3.9 (partie dense)**.** Soit $E \subset \mathbb{R}$ une partie. On dit que E est dense dans $\mathbb{R}$ si pour tout $x \in \mathbb{R}$, et pour tout $r > 0$, il existe $e \in E$ tel que e est r-proche de x, soit $|x - e| < r$.

Une partie dense E est donc par définition une partie dont les éléments permettent d'approcher aussi précisément que possible n'importe quel réel. Bien sûr $\mathbb{R}$ est dense dans $\mathbb{R}$! Mais l'intérêt de la notion, c'est que la partie dense E peut être beaucoup plus petite que $\mathbb{R}$ lui-même, comme le montre l'exemple suivant :

Exemple 3.10. $\mathbb{D}$ ou $\mathbb{Q}$ sont denses dans $\mathbb{R}$.

Comme $\mathbb{D} \subset \mathbb{Q}$ il suffit de démontrer que $\mathbb{D}$ est dense. Or quand on fixe $x \in \mathbb{R}$ et que l'on considère l'écriture décimale de x, on réalise que pour tout $n \in \mathbb{N}$ on a un encadrement $d_n \leq x < d_n + \frac{1}{10^n}$ où d_n est un nombre décimal ayant n chiffres après la virgule (d_n est l'écriture décimale de x, tronquée au n-ième chiffre après la virgule ; en formules $d_n = \frac{\lfloor 10^n x \rfloor}{10^n}$). On a $|x - d_n| < \frac{1}{10^n}$, ce qui conclut.

Par exemple si on applique la construction précédente au nombre réel $\sqrt{2} = 1,414213...$ (nombre qui n'est pas rationnel !), la suite de décimaux obtenue est

$$d_0 = 1 \ ; \ d_1 = 1,4 \ ; \ d_2 = 1,41 \ ; \ d_3 = 1,414 \ ; \ d_4 = 1,4142 \ ; \ \text{... et on a :}$$

$$\sqrt{2} - d_0 = 0,414213... < 1 \ ; \ \sqrt{2} - d_1 = 0,014213... < \frac{1}{10} \ ; \ \sqrt{2} - d_2 = 0,004213... < \frac{1}{100} \ ;$$
$$\sqrt{2} - d_3 = 0,000213... < \frac{1}{1000} \ ; \ \sqrt{2} - d_4 = 0,000013... < \frac{1}{10000} \ ; \ ...$$

Proposition 3.11. *Soit $E \subset \mathbb{R}$ une partie. Sont équivalentes :*

(1) *Pour tout x dans $\mathbb{R}$ et pour tout $r > 0$ l'intersection $]x-r; x+r[\cap E$ est non vide (c'est à dire E est dense dans $\mathbb{R}$).*

(2) *Pour tout x dans $\mathbb{R}$ il existe une suite $(e_n)_n$ d'éléments de E tels que $e_n \to x$.*

(3) *Le seul fermé de $\mathbb{R}$ qui contient E est $\mathbb{R}$.*

Démonstration. (1) $\Rightarrow$ (2) : car pour $x \in \mathbb{R}$ fixé, en prenant $r = \dfrac{1}{n+1}$ au lieu de $r > 0$ quelconque on trouve un $e_n \in E$ tel que $e_n \in]x - \dfrac{1}{n+1}; x + \dfrac{1}{n+1}[$, c'est à dire $d(e_n, x) < \dfrac{1}{n+1}$. La suite $(e_n)_n$ est formée d'éléments de E et tend vers x.

(2) $\Rightarrow$ (3) : si $E \subset F$ avec F fermé, alors pour $x \in \mathbb{R}$ prenons une suite $(e_n)_n$ formée d'éléments de E et qui tend vers x. A fortiori $(e_n)_n$ est formée d'éléments de F. Par la caractérisation séquentielle des fermés la limite de $(e_n)_n$ est dans F, soit $x \in F$.

(3) $\Rightarrow$ (1) , ou plutôt $\neg(1) \Rightarrow \neg(3)$. Si E n'est pas dense, soit $x_0 \in \mathbb{R}$ tel qu'il existe $r_0 > 0$ avec $]x_0 - r_0; x_0 + r_0[\cap E = \emptyset$. Alors soit $F = \mathbb{R} \backslash]x_0 - r_0; x_0 + r_0[$. On a $E \subset F$, $F \neq \mathbb{R}$ et F fermé comme complémentaire d'un intervalle ouvert. $\square$

4. Fonctions continues.

Dans cette dernière section nous rappelons sans entrer dans les détails ce qui concerne la continuité : tout cela sera repris en profondeur pour les fonctions $f : \mathbb{R}^m \to \mathbb{R}$.

4.1. Définitions équivalentes.

Proposition 4.1 (continuité en un point)**.** *Soit $E \subset \mathbb{R}$ est une partie quelconque et soit $f : E \to \mathbb{R}$ une fonction. Soit $a \in E$ fixé. Les propriétés suivantes sont équivalentes :*

(1) *Pour toute suite réelle $(x_n)_n$ telle que $x_n \in E$ $(\forall n \in \mathbb{N})$ et $x_n \to a$, la suite réelle $(f(x_n))_n$ est convergente, de limite $f(a)$.*

(2) *Pour tout $\varepsilon > 0$, il existe un $\alpha > 0$ tel que : pour tout $x \in E$ vérifiant $|x - a| < \alpha$ [x est α-proche de a], on a $|f(x) - f(a)| < \varepsilon$ [$f(x)$ est ε-proche de $f(a)$].*

Démonstration. (1) $\Rightarrow$ (2) : on montre plutôt la contraposée, c'est à dire $\neg(2) \Rightarrow \neg(1)$. La propriété $\neg(2)$ donne : $\exists \varepsilon_0 > 0$ tel que pour tout $\alpha > 0$, il existe un $x \in E$ avec $|x - a| < \alpha$, et simultanément $|f(x) - f(a)| \geq \varepsilon_0$. On applique cette propriété pour tous les $\alpha > 0$ de la forme $\alpha = \dfrac{1}{1+n}$ $(n \in \mathbb{N})$. Alors il existe une suite $(x_n)_n$ d'éléments de E tels que $|x_n - a| < \dfrac{1}{1+n}$ et pourtant $|f(x_n) - f(a)| \geq \varepsilon_0$. Nous avons bien obtenu une suite $(x_n)_n$ de E qui tend vers a, et telle que $(f(x_n))_n$ ne tend pas vers $f(a)$: donc $\neg(1)$ est vraie.

Réciproquement démontrons (2) $\Rightarrow$ (1). Supposons (2) vraie, et montrons (1) - pour cela considérons une suite $(x_n)_n$ de E telle que $x_n \to a$. Fixons $\varepsilon > 0$: il existe donc $\alpha > 0$ tel que si $x \in E$ et $|x - a| < \alpha$ alors $|f(x) - f(a)| < \varepsilon$. Or, comme $x_n \to a$ il existe un rang N tel que pour $n \geq N$ on a

$|x_n - a| < \alpha$. Alors, par la propriété (2) explicitée précédemment, nous savons que $|f(x_n) - f(a)| < \varepsilon$. Nous avons bien vérifié que $f(x_n) \to f(a)$.

$\square$

Définition 4.2 (continuité)**.** Soit $E \subset \mathbb{R}$ est une partie quelconque et soit $f : E \to \mathbb{R}$ une fonction.

(1) Soit $a \in E$ fixé. Lorsque f vérifie en a les conditions équivalentes de la proposition 4.1, on dit que f est *continue* au point a.

(2) Lorsque f est continue en tout point $a \in E$, on dit que f est *continue (sur E)*.

Par exemple n'importe quelle fonction $f : \mathbb{Z} \to \mathbb{R}$ est continue.

Une fonction f est donc continue si elle vérifie d'une manière précise la propriété floue suivante : "pour x, a proches, les images $f(x), f(a)$ sont proches également". Il y a une condition simple pour satisfaire cette propriété :

Définition 4.3 (fonctions Lipschitziennes)**.** Soit $E \subset \mathbb{R}$ une partie quelconque et soit $f : E \to \mathbb{R}$ une fonction. Pour un réel $K \geq 0$, on dit que f *est K-Lipschitzienne sur E* si :

$$\text{pour } x, x' \text{ quelconques dans } E, \text{ on a } |f(x') - f(x)| \leq K|x' - x|.$$

Et f est dite *Lipschitzienne sur E* s'il existe un réel $K \geq 0$ tel que f est K-Lipschitzienne sur E.

Lemme 4.4. *[Lipschitz $\Rightarrow$ Continue] Si f est Lipschitzienne sur E alors f est continue sur E.*

Démonstration. Si f est K-Lipschitzienne, f est même *uniformément* continue, car elle vérifie :

$$\forall \varepsilon, \exists \alpha, \forall a : |x - a| < \alpha \Rightarrow |f(x) - f(a)| < \varepsilon \text{ - il suffit en effet de considérer } \alpha = \frac{\varepsilon}{K}.$$

Or $\forall \varepsilon, \exists \alpha, \forall a : \ldots$ est une propriété plus forte que $\forall a \forall \varepsilon, \exists \alpha, : \ldots$! $\square$

En utilisant les limites à gauche ou à droite, on peut définir le nombre dérivée à gauche et à droite, et si ces nombre sont égaux on obtient le nombre dérivée.

Théorème 4.5 (théorème des accroissements finis - TAF)**.** *Soit $f : [a; b] \to \mathbb{R}$ une fonction continue, dérivable sur $]a; b[$. Alors pour c, d deux nombres distincts compris entre a et b il existe un réel ξ (strictement) compris entre c et d tel que $\dfrac{f(c) - f(d)}{c - d} = f'(\xi)$.*

Si f est dérivable en a, alors elle est comme "Lipschitzienne en a", et donc elle est continue (l'argument suit vraiment celui du Lemme 4.4). Mais il y a mieux :

Proposition 4.6 (dérivée bornée $\Rightarrow$ Lipschitzienne)**.**

Supposons que f est continue sur $[a; b]$ et dérivable sur $]a; b[$. Supposons de plus que $K = \sup\limits_{x \in]a;b[} |f'(x)|$ est fini. Alors f est K-Lipschitzienne sur $[a; b]$.

Démonstration. Pour $x, y \in [a; b]$, $x \neq y$, nous savons par le TAF qu'il existe un ξ entre x et y tel que $\dfrac{f(x) - f(y)}{x - y} = f'(\xi)$, d'où $\dfrac{|f(x) - f(y)|}{|x - y|} \leq K$ et pour finir $|f(x) - f(y)| \leq K|x - y|$. $\square$

Corollaire 4.7. *Soit $f : [a; b] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Alors f est Lipschitzienne sur $[a; b]$.*

4.2. **Topologie et continuité.** Quand l'ensemble de définition de la fonction est ouvert on peut également caractériser la continuité de façon très élégante (adieux : suites, $\forall \varepsilon, \exists \alpha$...) :

Proposition 4.8. *Soit $U \subset \mathbb{R}$ un ouvert (par exemple $U = \mathbb{R}$) et soit $f : U \to \mathbb{R}$ une fonction. Alors f est continue sur $U \iff$ pour tout ouvert O, l'image réciproque $f^{-1}(O)$ est un ouvert.*

Démonstration. Supposons f continue sur U. Soit O un ouvert et soit $x \in f^{-1}(O)$, i.e. $f(x) \in O$. Comme O est ouvert, $\exists \varepsilon > 0$ tel que $]f(x) - \varepsilon; f(x) + \varepsilon[\subset O$. Comme f est continue en x, $\exists \alpha > 0$ tel que si $d(x', x) < \alpha$ (i.e. si $x' \in]x - \alpha; x + \alpha[$), alors $d(f(x'), f(x)) < \varepsilon$, soit $f(x') \in]f(x) - \varepsilon; f(x) + \varepsilon[$ - et a fortiori $f(x') \in O$, soit $x' \in f^{-1}(O)$. Ceci montre que $f^{-1}(O)$ contient $]x - \alpha; x + \alpha[$, donc est voisinage de x.

Réciproquement supposons que pour tout ouvert O, l'image réciproque $f^{-1}(O)$ est un ouvert. Pour $x \in U$ et $\varepsilon > 0$ fixés, posons $O =]f(x) - \varepsilon; f(x) + \varepsilon[$: c'est un ouvert, donc $f^{-1}(O)$ est un ouvert. Un ouvert qui contient x, donc $\exists \alpha > 0$ tel que $]x - \alpha; x + \alpha[\subset f^{-1}(O)$. Ceci signifie que $\forall x' \in]x - \alpha; x + \alpha[$ on a $f(x') \in O =]f(x) - \varepsilon; f(x) + \varepsilon[$, c'est à dire $d(f(x'), f(x)) < \varepsilon$.

□

Corollaire 4.9. *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction. Alors f est continue $\iff$ pour tout fermé F, l'image réciproque $f^{-1}(F)$ est un fermé.*

Démonstration. Car pour toute partie $Y \subset \mathbb{R}$ on a $\mathbb{R} \setminus f^{-1}(Y) = f^{-1}(\mathbb{R} \setminus Y)$. □

Exemple 4.10. Pour $a, b \in \mathbb{R}$ fixés, et pour n'importe quelle fonction continue $f : \mathbb{R} \to \mathbb{R}$, les parties $\{x \in \mathbb{R}, a \leq f(x)\}$, $\{x \in \mathbb{R}, f(x) \leq b\}$ et $\{x \in \mathbb{R}, a \leq f(x) \leq b\}$ sont fermées. Il s'agit en effet des images réciproques $f^{-1}([a; +\infty[)$, $f^{-1}(] - \infty; b])$ et $f^{-1}([a; b])$ respectivement. Or f est continue et $[a; +\infty[$, $] - \infty; b]$ et $[a; b]$ sont fermés.

De la même façon les parties $\{x \in \mathbb{R}, a < f(x)\}$, $\{x \in \mathbb{R}, f(x) < b\}$ et $\{x \in \mathbb{R}, a < f(x) < b\}$ sont ouvertes, comme images réciproques par f continue des intervalles ouverts $]a; +\infty[$, $]-\infty; b[$ et $]a; b[$.

Par exemple l'ensemble $\{x \in \mathbb{R}, \frac{1}{4} \leq x^3 - x \leq \frac{3}{4}\}$ est fermé - comme image réciproque par la fonction continue $x \mapsto x^3 - x$ de l'intervalle fermé $[\frac{1}{4}; \frac{3}{4}]$.

Et $\{x \in \mathbb{R}, -\frac{2}{3} < \sin(x) < \frac{2}{3}\}$ est ouvert- comme image réciproque par la fonction continue $x \mapsto \sin(x)$ de l'intervalle ouvert $] - \frac{2}{3}; \frac{2}{3}[$.

Proposition 4.11. *Soit $f : F \to \mathbb{R}$ une fonction continue, avec $F = \mathbb{R}$ ou plus généralement F est un fermé de $\mathbb{R}$. Soit $t \in \mathbb{R}$. Alors l'ensemble $\mathcal{L}_t(f) = \{x \in \mathbb{R}, f(x) = t\}$ est une partie fermée de $\mathbb{R}$.*

Démonstration. $\mathcal{L}_t(f)$ est séquentiellement fermé. Car si $x_n \to \ell$ ($\ell \in \mathbb{R}$) et $\forall n, x_n \in \mathcal{L}_t(f)$ (soit $f(x_n) = t$), alors 1) $\ell \in F$ (car F fermé) puis 2) $f(\ell) = \lim_{n\to\infty} f(x_n)$ car f continue sur F. Alors 3) $f(\ell) = t$ soit $t \in \mathcal{L}_t(f)$. □

Corollaire 4.12 (zéros des fonctions continues)**.** *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue. L'ensemble des solutions de l'équation $f(x) = 0$ forme une partie fermée.*

Corollaire 4.13 (points fixes des fonctions continues)**.** *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue. L'ensemble des solutions de l'équation $f(x) = x$ est une partie fermée.*

Corollaire 4.14 (lieu d'égalité de deux fonctions continues)**.** *Soit $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ deux fonctions continues. L'ensemble des solutions de l'équation $f(x) = g(x)$ forme une partie fermée.*

Démonstration. Car $\{x \in \mathbb{R}, f(x) = g(x)\}$ est l'ensemble $\mathcal{L}_0(f-g)$, fermé d'après la Proposition 4.11 puisque $x \mapsto f(x) - g(x)$ est continue. □

Corollaire 4.15 (prolongement des identités par densité)**.** *Soit $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ deux fonctions continues. On suppose qu'il existe une partie dense $E \subset \mathbb{R}$ telle que $\forall t \in E, f(t) = g(t)$. Alors $f = g$, c'est à dire que $\forall x \in \mathbb{R}, f(x) = g(x)$.*

Démonstration. En effet $\{x \in \mathbb{R}, f(x) = g(x)\}$ est fermé et contient une partie dense. □

Exemple 4.16. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue telle que pour x, y quelconques on a $f(x+y) = f(x) + f(y)$. Alors il existe un réel K tel que pour tout $x \in \mathbb{R}$ on a $f(x) = Kx$.

En effet posons $K = f(1)$ (on n'a pas tellement le choix) et montrons que pour tout $x \in \mathbb{R}$ on a $f(x) = g(x)$, où l'on a posé $g(x) = Kx$ (pour tout $x \in \mathbb{R}$). On montre d'abord que $f(x) = g(x)$ pour tout $x \in \mathbb{Q}$, puis on prolongera cette identité entre fonctions continues par densité des rationnels.

Soit F l'ensemble des x tels que $f(x) = g(x)$, et montrons que $\mathbb{Q} \subset F$. D'abord $1 \in F$. Puis dès que $x, x' \in F$, on a $f(x) = g(x)$ et $f(x') = g(x')$ alors $f(x + x') = f(x) + f(x') = g(x) + g(x') = Kx + Kx' = K(x + x') = g(x + x')$, donc $x + x' \in F$. Par exemple $f(x + 1) = f(x) + f(1) = Kx + K = K(x + 1)$. Donc en fait par récurrence à partir de $x = 1$ on voit que $f(x) = g(x)$ pour tout $x \in \mathbb{N}, x > 0$.

Par récurrence, on voit que la propriété fonctionnelle de f a une autre conséquence : pour tout entier $n \in \mathbb{N}, n > 0$ on a $f(nx) = nf(x)$. De $f(nx) = nf(x)$ (et $g(nx) = ng(x)$) nous tirons aussi que si $x \in F$ alors $nx \in F$. Et si $nx \in F$ alors $x \in F$. Autrement si $y \in F$ alors $\dfrac{y}{n} \in F$. En particulier si $q \in \mathbb{N}^*$ alors $\dfrac{1}{q} \in F$. En combinant les deux stabilités nous voyons que pour p, q deux entiers > 0 nous avons $\dfrac{p}{q} \in F$. Donc F contient tous les rationnels > 0. Pour conclure nous remarquons que la propriété fonctionnelle de f donne d'abord $f(x + 0) = f(x) + f(0)$, donc $f(x) = f(x) + f(0)$, et pour finir $f(0) = 0$. En particulier $0 \in F$. Enfin $f(0) = f(x + (-x)) = f(x) + f(-x)$, donc $0 = f(x) + f(-x)$ et $f(-x) = -f(x)$: comme bien entendu $g(-x) = -g(x)$ aussi, nous obtenons que si $x \in F$ alors $-x \in F$. En conclusion F contient $\mathbb{Q}^+$ et $\mathbb{Q}^-$, donc en fait $\mathbb{Q} \subset F$.

Deuxième partie 2. Normes, distances et topologie dans $\mathbb{R}^m$.

1. Exemples de normes, distances associées.

1.1. **Normes euclidiennes dans le plan ou l'espace.** On rappelle que dans $\mathbb{R}^2$ étant donné deux vecteurs $u = (x, y)$ et $u' = (x', y')$, on définit le *produit scalaire* de u et u' par la formule :

$$u \cdot u' = xx' + yy'.$$

De même dans $\mathbb{R}^3$ étant donné deux vecteurs $u = (x, y, z)$ et $u' = (x', y', z')$, on définit le produit scalaire de u et u' par la formule :

$$u \cdot u' = xx' + yy' + zz'.$$

Le produit scalaire est *bilinéaire* :

$$(au + bv) \cdot (a'u' + b'v') = aa' \ u \cdot u' + ab' \ u \cdot v' + ba' \ v \cdot u' + bb' \ v \cdot v'$$

et *symétrique* :

$$u \cdot u' = u' \cdot u.$$

Dans $\mathbb{R}^2$ ou $\mathbb{R}^3$, on constate que $u \cdot u$ est ≥ 0, car c'est une somme de carrés de nombres réels ($x^2 + y^2$ ou $x^2 + y^2 + z^2$).

Et d'ailleurs $u \cdot u = 0$ si et seulement si toutes les composantes de u sont nulles, autrement dit $u = 0$.

On peut donc définir pour tout vecteur $u \in \mathbb{R}^{2 \text{ ou } 3}$

$\|u\|_2 = \sqrt{u \cdot u}$: si $u = (x, y) \in \mathbb{R}^2$ alors $\|u\|_2 = \sqrt{x^2 + y^2}$, si $u = (x, y, z) \in \mathbb{R}^3$ alors $\|u\|_2 = \sqrt{x^2 + y^2 + z^3}$

Cette quantité $\|u\|_2$ s'appelle la *norme euclidienne* du vecteur u.

En élevant au carré, il est facile de vérifier $\boxed{\textit{l'inégalité de Cauchy-Schwartz} : |u \cdot v| \leq \|u\|_2 \|v\|_2}$ (pour deux vecteurs u, v quelconques). De Cauchy-Schwartz on déduit $\|u + v\|_2^2 \leq (\|u\|_2 + \|v\|_2)^2$, et donc

$$(N_+) \ \|u + v\|_2 \leq \|u\|_2 + \|v\|_2$$

D'autre part bien sûr on a (pour tout vecteur u et tout réel λ) :

$$(N_\times) \ \|\lambda \times u\|_2 = |\lambda| \times \|u\|_2.$$

1.2. **Normes générales, cas fondamental des normes $\|\cdot\|_\infty$ et $\|\cdot\|_1$.**

On utilise comme axiomes les propriétés constatées sur $\|\cdot\|_2$:

Définition 1.1. Une *norme* sur $\mathbb{R}^m$ est une application $\|\cdot\| : \mathbb{R}^m \to \mathbb{R}^+$ qui vérifie :

1. $(N_{>0})$: Si $u \neq 0$ alors $\|u\| > 0$; et $\|0\| = 0$ ($\|\cdot\|$ est *définie-positive*).
2. (N_+) : Pour deux vecteurs $u, v \in \mathbb{R}^m$ quelconques, on a $\|u + v\| \leq \|u\| + \|v\|$ ($\|\cdot\|$ est *sous-additive*, ou vérifie *l'inégalité triangulaire*).
3. $(N_\times)$: Pour tout vecteur u et tout scalaire réel λ, on a $\|\lambda \times u\| = |\lambda| \times \|u\|$ ($\|\cdot\|$ est *scalairement multiplicative*, ou *homogène* - ne pas oublier la valeur absolue !)

On peut noter que $(N_\times)$ donne pour $\lambda = 0$ (et u quelconque) que $\|0\| = 0$. Donc la partie importante de $(N_{>0})$ est que $\|u\| > 0$ si $u \neq 0$. En appliquant $n - 1$ fois la sous-additivité on obtient :

$$\|u_1 + \cdots + u_n\| \leq \|u_1\| + \cdots + \|u_n\|$$

Plus généralement en combinant (N_+) et $(N_\times)$ on voit qu'on peut majorer la norme des combinaisons linéaires :

$$(N_{\text{CL}}) : \|\lambda_1 u_1 + \cdots + \lambda_k u_k\| \leq |\lambda_1|\|u_1\| + \cdots + |\lambda_k|\|u_k\|$$

Définition 1.2 ($\|\cdot\|_1$ et $\|\cdot\|_\infty$)**.** Pour $u = (x_1, \cdots, x_m)$ un vecteur de $\mathbb{R}^m$, on pose :
$\|u\|_1 = |x_1| + \cdots + |x_m|$
$\|u\|_\infty = \max(|x_1|, \cdots, |x_m|)$

Lemme 1.3. *Sur $\mathbb{R}^m$, les fonctions $u \mapsto \|u\|_1$ et $u \mapsto \|u\|_\infty$ sont deux normes.*

Démonstration. Par construction pour tout vecteur u on a $\|u\|_1 \geq 0$ et $\|u\|_\infty \geq 0$.

Pour $u = (x_1, \ldots, x_m)$ si $\|u\|_1 = 0$ cela signifie la nullité de la somme des $|x_i|$, qui sont des nombres ≥ 0. Si l'un des $|x_i|$ est > 0, alors la somme serait > 0. Donc $|x_1| = \cdots = |x_m| = 0$ et on a bien $u = 0$.

Pour $u = (x_1, \ldots, x_m)$ si $\|u\|_\infty = 0$ alors pour tout $i = 1, \ldots, d$ on a $0 \leq |x_i| \leq \max(|x_i|) = 0$, donc $|x_i| = 0$ et ici encore $u = 0$.

On vient de vérifier $(N_{>0})$ pour $\|\cdot\|_1$ et $\|\cdot\|_\infty$.

Soient $u, v \in \mathbb{R}^m$ deux vecteurs quelconques, avec $u = (x_1, \ldots, x_m)$ et $v = (y_1, \ldots, y_m)$, de sorte que $u + v = (x_1 + y_1, \ldots, x_m + y_m)$. Par l'inégalité triangulaire de la valeur absolue, pour tout $i = 1, \ldots, d$, on obtient $|x_i + y_i| \leq |x_i| + |y_i|$. En ajoutant ces inégalités on trouve :

$$\|u + v\|_1 = \sum_{i=1}^{i=m} |x_i + y_i| \leq \sum_{i=1}^{i=m} (|x_i| + |y_i|) = \Big(\sum_{i=1}^{i=m} |x_i|\Big) + \Big(\sum_{i=1}^{i=m} |y_i|\Big) = \|u\|_1 + \|v\|_1.$$

De $|x_i + y_i| \leq |x_i| + |y_i|$ on tire aussi :

$\forall i = 1, \cdots, d, |x_i + y_i| \leq \max(|x_i|) + \max(|y_i|) = \|u\|_\infty + \|v\|_\infty,$ donc $\|u + v\|_\infty = \max(|x_i + y_i|) \leq \|u\|_\infty + \|v\|_\infty$.

On vient de vérifier (N_+) pour $\|\cdot\|_1$ et $\|\cdot\|_\infty$.

Si $u = (x_1, \ldots, x_m)$ et $\lambda \in \mathbb{R}$ on a $\lambda u = (\lambda x_1, \ldots, \lambda x_m)$. Donc

$$\|\lambda u\|_1 = |\lambda x_1| + \cdots + |\lambda x_m| = |\lambda||x_1| + \cdots + |\lambda||x_m| = |\lambda|(|x_1| + \cdots + |x_m|) = |\lambda|\|u\|_1, \text{ et}$$

$$\|\lambda u\|_\infty = \max\big(|\lambda x_1|, \cdots, |\lambda x_m|\big) = \max\big(|\lambda||x_1|, \cdots, |\lambda||x_m|\big) = |\lambda| \max\big(|x_1|, \cdots, |x_m|\big) = |\lambda|\|u\|_\infty.$$

On vient de vérifier $(N_\times)$ pour $\|\cdot\|_1$ et $\|\cdot\|_\infty$. $\square$

Remarque 1.4 (norme euclidienne sur $\mathbb{R}^m$)**.** Pour $u = (x_1, \cdots, x_m)$ un vecteur de $\mathbb{R}^m$, on peut poser $\|u\|_2 = \sqrt{x_1^2 + \cdots + x_m^2}$. Pour $\|\cdot\|_2$ les axiomes $(N_{>0})$ et $(N_\times)$ sont faciles à vérifier. Et avec plus de travail on peut démontrer ici aussi que l'axiome (N_+) est satisfait (comme dans le cas particulier de $\mathbb{R}^2$ ou $\mathbb{R}^3$) : on parle de *la norme euclidienne standard de* $\mathbb{R}^m$. Ici aussi on déduit (N_+) de l'inégalité de Cauchy-Schwarz, mais pour établir CS c'est une autre histoire : une histoire d'algèbre bilinéaire, ce sera vu en MEU252 (S4).

Remarque 1.5 (normes sur $\mathbb{R}$)**.** Dans $\mathbb{R}^2$ nous connaissons trois normes distinctes $\|\cdot\|_1, \|\cdot\|_\infty, \|\cdot\|_2$ (elles valent 1 en $(1, 0)$ mais diffèrent en $(1, 1)$).

En revanche sur $\mathbb{R}$ il existe une unique norme qui vaut 1 en 1 : c'est la valeur absolue de $\mathbb{R}$! Les autres normes sur $\mathbb{R}$ sont simplement les fonctions $u \mapsto k|u|$ (pour $k \in \mathbb{R}^{+*}$ une constante fixée) : cela résulte de l'axiome $(N_\times)$ (on aura $N(u) = N(u \times 1) = |u|N(1) = k|u|$ avec $k = N(1)$).

1.3. **Distance associée à une norme, boules.**

Définition 1.6.

Soit $u \mapsto \|u\|$ une norme sur $\mathbb{R}^m$. La *distance sur* $\mathbb{R}^m$ associée à cette norme est la fonction $d : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^+$ définie par $d(v, w) = \|w - v\|$.

En particulier on notera d_1, d_∞, d_2 les distances associées à $\|\cdot\|_1, \|\cdot\|_\infty, \|\cdot\|_2$. Donc :

$d_1\big((x_1, \cdots, x_m), (y_1, \cdots, y_m)\big) = |y_1 - x_1| + \cdots + |y_m - x_m|$ $(= |x_1 - y_1| + \cdots + |x_m - y_m|)$

$d_\infty\big((x_1, \cdots, x_m), (y_1, \cdots, y_m)\big) = \max(|y_1 - x_1|, \cdots, |y_m - x_m|)$ $(= \max(|x_1 - y_1|, \cdots, |x_m - y_m|))$

$d_2\big((x_1, \cdots, x_m), (y_1, \cdots, y_m)\big) = \sqrt{(y_1 - x_1)^2 + \cdots + (y_m - x_m)^2}$ $(= \sqrt{(x_1 - y_1)^2 + \cdots + (x_m - y_m)^2})$

Par exemple, dans l'espace $\mathbb{R}^3$, pour $A = (1, 2, 1)$ et $B = (0, 1, 2)$ on a $B - A = (-1, -1, 1)$, et donc

$$d_1(A, B) = 3, d_\infty(A, B) = 1, d_2(A, B) = \sqrt{3}$$

Lemme 1.7 (propriétés caractéristiques des distances)**.** *Soit d la distance sur $\mathbb{R}^m$ associée à une norme $u \mapsto \|u\|$. Alors*

d est symétrique *: Pour $v, w \in \mathbb{R}^m$ on a $d(v, w) = d(w, v)$.*

d est définie positive *: Pour $v, w \in \mathbb{R}^m$ tel que $v \neq w$ on a $d(v, w) > 0$. Pour $v \in \mathbb{R}^m$ on a $d(v, v) = 0$.*

*d vérifie l'*inégalité triangulaire *: Pour $u, v, w \in \mathbb{R}^m$ on a $d(u, w) \leq d(u, v) + d(v, w)$.*

Démonstration.

(1) $d(v, w) = \|w - v\| = \|-(v - w)\| = |-1|\|v - w\| = \|v - w\| = d(w, v)$.

(2) Pour $v \neq w$ on a $d(v, w) = \|w - v\|$, qui est la norme d'un vecteur non nul, donc est > 0. Et si $w = v$ on a $d(v, w) = \|0\| = 0$.

(3) $d(u, w) = \|w - u\| = \|(v - u) + (w - v)\| \leq \|v - u\| + \|w - v\| = d(v, u) + d(v, w)$

□

Remarque 1.8.

(1) Pour tout vecteur $u \in \mathbb{R}^m$ on a $d(0, u) = \|u\|$.

(2) Le caractère défini-positif revient à dire que

Pour $v, w \in \mathbb{R}^m$ on a $d(v, w) \geq 0$, et $d(v, w) = 0 \iff v = w$.

(3) Un cas particulier de l'inégalité triangulaire appliquée à $0, u, u + v$ redonne $\|u + v\| \leq \|u\| + \|v\|$, ce qui explique pourquoi cette inégalité sur les normes est également appelée inégalité triangulaire.

Lemme 1.9 (minoration triangulaire)**.**

Soit d la distance sur $\mathbb{R}^m$ associée à une norme $u \mapsto \|u\|$. Pour $u, v, w \in \mathbb{R}^m$ on a

$$d(u, w) \geq |d(u, v) - d(v, w)|.$$

Démonstration. L'inégalité demandée revient à $-d(u, w) \leq d(u, v) - d(v, w) \leq d(u, w)$.

Or $-d(u, w) \leq d(u, v) - d(v, w) \iff d(v, w) \leq d(u, v) + d(u, w)$, ce qui est vrai par inégalité triangulaire. Et $d(u, v) - d(v, w) \leq d(u, w) \iff d(u, v) \leq d(u, w) + d(v, w)$, à nouveau vrai par inégalité triangulaire. □

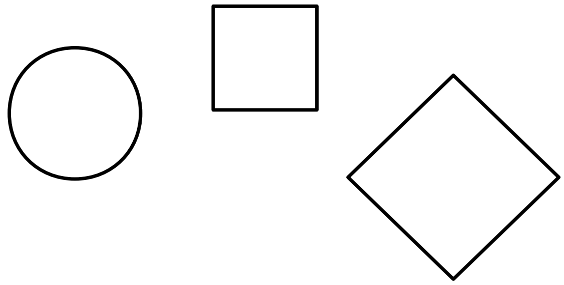


FIGURE 1. Trois boules dans $\mathbb{R}^2$, successivement pour les normes $\|\cdot\|_2$, $\|\cdot\|_\infty$ et $\|\cdot\|_1$.

(On peut en fait vérifier que la propriété de minoration triangulaire n'est pas seulement une conséquence de l'inégalité triangulaire : elle lui est équivalente.)

Définition 1.10 (boules et sphères)**.** On suppose que $\mathbb{R}^m$ est muni d'une norme $\|\cdot\|$ et de la distance $d(\cdot,\cdot)$ associée à cette norme. Soit $r \in \mathbb{R}^+$ un réel fixé.

(1) On dira que $v, w \in \mathbb{R}^m$ sont *r-proches (pour d)* si $d(v, w) < r$.

(2) Soit $u \in \mathbb{R}^m$ un point fixé.

  (a) On appelle *boule ouverte de centre u et de rayon r* la partie $\{v \in \mathbb{R}^m, d(u, v) < r\}$, notée $B(u, r)$. Donc $B(u, r)$ est l'ensemble des $v \in \mathbb{R}^m$ qui sont r-proches de u.

  (b) On appelle *boule fermée de centre u et de rayon r* la partie $\{v \in \mathbb{R}^m, d(u, v) \leq r\}$ notée $\bar{B}(u, r)$.

  (c) On appelle *sphère de centre u et de rayon r* la partie $\{v \in \mathbb{R}^m, d(u, v) = r\}$ notée $S(u, r)$.

Bien entendu les boules ouvertes de rayon $r = 0$ sont vides, et $\bar{B}(u, 0) = \{u\}$.

La terminologie de boule et de sphère se justifie dans le cas de l'espace $\mathbb{R}^3$, en prenant la distance euclidienne d_2. Dans $\mathbb{R}^2$ muni de sa distance euclidienne d_2 habituelle on parle de *disque* plutôt que de boule, et de *cercle* plutôt que de sphère.

Remarque 1.11. Les notions de norme et de distance associée ont un sens sur un espace vectoriel quelconque E, mais on n'en parlera pas dans cette UE. Nous pouvons tout de même remarquer que si nous avons une norme sur $\mathbb{R}^m$, alors pour tout sous-espace vectoriel $E \subset \mathbb{R}^m$, la restriction $\|\cdot\|_{|E}$ vérifie les axiomes $(N_{>0})$, (N_+) et $(N_\times)$ - et c'est ce qui compte pour obtenir une norme.

Exemple 1.12 (la r-proximité pour $\|\cdot\|_\infty$)**.** Soient v, w deux points de $\mathbb{R}^m$. Donc $v = (v_1, \ldots, v_n)$ et $w = (w_1, \ldots, w_n)$. Alors dire que v et w sont r-proches pour la norme $\|\cdot\|_\infty$ revient à dire que v_1, w_1 sont r-proches, ... , et v_n, w_n sont r-proches. C'est la notion de proximité la plus intuitive !

A comparer avec l'expression de la r-proximité si à la place de $\|\cdot\|_\infty$ on prend une autre norme, comme $\|\cdot\|_1$ ou $\|\cdot\|_2$...

De même que la notion de proximité dépend de la norme choisie, de même bien sûr les boules dépendent de la norme ! Voir la figure 1 pour les normes $\|\cdot\|_2, \|\cdot\|_\infty, \|\cdot\|_1$.

2. Comparaison de normes.

Sur $\mathbb{R}^m$ il y a beaucoup de normes, et donc de distances, assez différentes : quand deux points A et B sont r-proches pour $\|\cdot\|_\infty$, ils ne sont pas nécessairement r-proches pour une autre norme - mettons $\|\cdot\|_1$. On aimerait pouvoir comparer les distances (et donc les notions de proximité) associées à des normes différentes.

Proposition 2.1. *Soit $\|\cdot\|$ une norme quelconque sur $\mathbb{R}^m$. Notons $(e_1, \dots, e_m)$ la base canonique de $\mathbb{R}^m$.*

(1) Si on pose $M_1 = \max\{\|e_k\|, k = 1, \dots, m\}$, alors pour tout vecteur $u \in \mathbb{R}^m$ on a $\|u\| \leq M_1\|u\|_1$.

(2) Si on pose $M_\infty = \|e_1\| + \cdots + \|e_m\|$, alors pour tout vecteur $u \in \mathbb{R}^m$ on a $\|u\| \leq M_\infty\|u\|_\infty$.

(3) Si on pose $M_2 = \sqrt{\|e_1\|^2 + \cdots + \|e_m\|^2}$, alors pour tout vecteur $u \in \mathbb{R}^m$ on a $\|u\| \leq M_2\|u\|_2$.

Démonstration. Soit $u = (x_1, \dots, x_m)$ c'est à dire $u = x_1e_1 + \cdots + x_me_m$. Alors

$$\|u\| \leq \|x_1e_1\| + \cdots + \|x_me_m\| = |x_1|\|e_1\| + \cdots + |x_m|\|e_m\|$$

On va alors obtenir les majorations de $\|u\|$ par les autres normes en contrôlant le terme $|x_1|\|e_1\| + \cdots + |x_m|\|e_m\|$.

Pour $M_1 = \max\{\|e_k\|, k = 1, \dots, m\}$ on a

$$|x_1|\|e_1\| + \cdots + |x_m|\|e_m\| \leq |x_1|M_1 + \cdots + |x_m|M_1 = M_1(|x_1| + \cdots + |x_m|) = M_1\|u\|_1.$$

Pour $M_\infty = \|e_1\| + \cdots + \|e_m\|$ on a

$$|x_1|\|e_1\| + \cdots + |x_m|\|e_m\| \leq \max(|x_k|, k = 1, \dots, m)(\|e_1\| + \cdots + \|e_m\|) = \|u\|_\infty M_\infty.$$

Reste la majoration par $\|\cdot\|_2$. Pour cela on utilise l'inégalité de Cauchy-Schwartz (déjà connue pour $m = 2$ ou $m = 3$, et valable en fait pour tout m, comme on l'a admis). Ainsi on a toujours :

$$|aa' + bb'| \leq \sqrt{a^2 + b^2}\,\sqrt{a'^2 + b'^2},$$

$$|aa' + bb' + cc'| \leq \sqrt{a^2 + b^2 + c^2}\,\sqrt{a'^2 + b'^2 + c'^2},$$

et plus généralement

$$|a_1a_1' + \cdots + a_ma_m'| \leq \sqrt{a_1^2 + \cdots + a_m^2}\,\sqrt{a_1'^2 + \cdots + a_m'^2}.$$

On interprète $|x_1|\|e_1\| + \cdots + |x_m|\|e_m\|$ comme le produit scalaire $u.u'$ avec $u = (|x_1|, \dots, |x_m|)$ et $u' = (\|e_1\|, \cdots, \|e_m\|)$. L'inégalité de CS donne alors

$$u.u' \leq |u.u'| \leq \sqrt{|x_1|^2 + \cdots + |x_m|^2}\,\sqrt{\|e_1\|^2 + \cdots + \|e_m\|^2}$$

soit exactement

$$|x_1|\|e_1\| + \cdots + |x_m|\|e_m\| \leq \|u\|_2M_2 = M_2\|u\|_2.$$

□

Corollaire 2.2. *Soit $u \in \mathbb{R}^m$.*

(1) On a $\|u\|_1 \leq m\|u\|_\infty$ et $\|u\|_\infty \leq \|u\|_1$.

(2) Si $m = 2$ ou $m = 3$ on a $\|u\|_2 \leq \|u\|_1$.

(3) Si $m = 2$ on a $\|u\|_2 \leq 2\|u\|_\infty$ et si $m = 3$ on a $\|u\|_2 \leq 3\|u\|_\infty$.

(4) $(4.a)$ *Si* $m = 2$ *on a* $\|u\|_1 \leq \sqrt{2}\|u\|_2$ *et si* $m = 3$ *on a* $\|u\|_1 \leq \sqrt{3}\|u\|_2$.
$(4.b)$ *Si* $m = 2$ *on a* $\|u\|_\infty \leq \sqrt{2}\|u\|_2$ *et si* $m = 3$ *on a* $\|u\|_\infty \leq \sqrt{3}\|u\|_2$.

Remarque 2.3. Les majorations $(1), (2), (4.a)$ sont optimales : il y a toujours un vecteur non nul u qui réalise l'égalité dans l'inégalité. En revanche on peut un peu améliorer les majorations (3) et $(4.b)$. On voit immédiatement que si $m = 2$ on a $\|u\|_2 \leq \sqrt{2}\|u\|_\infty$ et si $m = 3$ on a $\|u\|_2 \leq \sqrt{3}\|u\|_\infty$. Et d'autre part clairement $\|u\|_\infty \leq \|u\|_2$. On obtient les encadrements :

$$\|u\|_\infty \leq \|u\|_1 \leq m\|u\|_\infty, \text{ ou } \frac{1}{m}\|u\|_1 \leq \|u\|_\infty \leq \|u\|_1 \;;\; \|u\|_2 \leq \|u\|_1 \leq \sqrt{m}\|u\|_2, \text{ ou } \frac{1}{\sqrt{m}}\|u\|_1 \leq \|u\|_2 \leq \|u\|_1$$

$$\text{et enfin } \|u\|_\infty \leq \|u\|_2 \leq \sqrt{m}\|u\|_\infty, \text{ ou } \frac{1}{\sqrt{m}}\|u\|_2 \leq \|u\|_\infty \leq \|u\|_2.$$

En résumé nous avons obtenu que si N, N' sont deux des trois normes $\|\cdot\|_1, \|\cdot\|_\infty, \|\cdot\|_2$ sur $\mathbb{R}^m$, alors il existe toujours une constante $K > 0$ telle que pour tout $u \in \mathbb{R}^m$ on a $N(u) \leq KN'(u)$: on dit que les normes $\|\cdot\|_1, \|\cdot\|_\infty, \|\cdot\|_2$ sont *équivalentes* sur $\mathbb{R}^m$.

Comme le montre le résultat suivant, c'est un fait valable pour deux normes **quelconques**, mais nous admettrons ce théorème (démonstration en L3...) :

Théorème 2.4 (équivalence de toutes les normes)**.** *Soient N, N' deux normes sur $\mathbb{R}^m$, alors il existe toujours une constante $K > 0$ telle que pour tout $u \in \mathbb{R}^m$ on a $N(u) \leq KN'(u)$ et $N'(u) \leq KN(u)$.*

Les conséquences de l'équivalence des normes sont très importantes (même si on se restreint à l'équivalence de $\|\cdot\|_1, \|\cdot\|_\infty, \|\cdot\|_2$).

Corollaire 2.5 (proximité pour une norme quelconque)**.** *Soit $\|\cdot\|$ une norme sur $\mathbb{R}^m$. Alors il existe une constante réelle $K > 0$ avec la propriété suivante. Pour $r > 0$ quelconque et deux éléments $v, w \in \mathbb{R}^m$:*

(1) Si v, w sont r-proches pour $\|\cdot\|$ alors v, w sont Kr-proches composantes par composantes : $|v_1 - w_1| < Kr, \ldots, |v_n - w_n| < Kr$.

(2) Si v, w sont r-proches composantes par composantes alors v, w sont Kr-proches pour $\|\cdot\|$.

Pour $\|\cdot\| = \|\cdot\|_1$ on peut prendre $K = m$ et pour $\|\cdot\| = \|\cdot\|_2$ on peut prendre $K = \sqrt{m}$.

Démonstration. On utilise l'équivalence de $\|\cdot\|$ avec $\|\cdot\|_\infty$. On sait donc qu'il existe une constante réelle $K > 0$ telle que pour tout $u \in \mathbb{R}^m$ on a $\|u\|_\infty \leq K\|u\|$ et $\|u\| \leq K\|u\|_\infty$. □

Corollaire 2.6 (parties bornées)**.** *Soit $X \subset \mathbb{R}^m$ une partie quelconque. Les propriétés suivantes sont équivalentes*

(0) X est contenu dans $[-M; M]^m = \{u = (x_1, \ldots, x_m) \in \mathbb{R}^m, \forall i = 1, \ldots, m$ on a $|x_i| \leq M\}$.

(1) X est contenu dans une boule pour la norme $\|\cdot\|_\infty$.

(1') L'ensemble de réels $\{\|v - w\|_\infty, v, w \in X\}$ est borné.

(2) Il existe une norme $\|\cdot\|_0$ telle que la partie X est contenu dans une boule pour la norme $\|\cdot\|_0$.

(2') Il existe une norme $\|\cdot\|_0$ telle que l'ensemble de réels $\{\|v - w\|_0, v, w \in X\}$ est borné.

(3) Pour toute norme $\|\cdot\|$ la partie X est contenu dans une boule pour la norme $\|\cdot\|$.

(3') Pour toute norme $\|\cdot\|$ l'ensemble de réels $\{\|v - w\|, v, w \in X\}$ est borné.

Une partie $X \subset \mathbb{R}^m$ qui vérifie ces propriétés équivalentes est dite bornée. *Si de plus $X \neq \emptyset$, le* diamètre *de X (pour la distance d associée à la norme $\|\cdot\|$) est alors défini par :*

$$diam(X, \|\cdot\|) := \sup\{d(v,w) \ (= \|v-w\|), v, w \in X\}$$

(On note aussi $diam(X)$ s'il n'y a pas de confusion possible sur la norme.)

Le centre de la boule considérée, son caractère ouvert ou fermé n'a pas d'importance, comme on va le voir dans l'argument. Donc pour finir X est borné si et seulement si X est contenue dans un hypercube $[-M; M]^m$.

Démonstration. Pour N une norme quelconque sur $\mathbb{R}^m$, dire que X est contenu dans la boule $B(A,r)$ (ou $\bar{B}(A,r)$) revient à dire que l'ensemble $\Delta_A(N) = \{N(v-A), v \in X\}$ est borné. Comme pour un autre point B on a $N(v-B) \leq N(v-A) + N(A-B)$ on voit que $\Delta_A(N)$ est borné ssi $\Delta_B(N)$ l'est. En particulier $(0) \iff (1)$. D'autre part si $N' \leq KN$ on voit que ($\Delta_A(N)$ borné) $\Rightarrow$ ($\Delta_A(N')$ borné). Donc en utilisant l'équivalence des normes les énoncés $(1), (2), (3)$ sont équivalents.

Montrons l'équivalence avec l'énoncé ayant un '. Si $X = \emptyset$ il n'y a rien à faire, supposons donc que X contient au moins un point A. Pour N une norme fixée sur $\mathbb{R}^m$ posons $\Delta(N) = \{N(v-w), v, w \in X\}$. Comme $A \in X$ on a $\Delta_A(N) \subset \Delta(N)$, donc si $\Delta(N)$ est borné alors X est contenu dans une boule ($\Delta_A(N)$ est borné). Réciproquement supposons que $\Delta_A(N)$ est borné, soit M un de ses majorants. Alors par inégalité triangulaire on a $N(v-w) \leq N(v-A) + N(A-w) \leq 2M$, donc $\Delta(N)$ est borné. □

Ainsi donc les boules (ouvertes ou fermées) de n'importe quelle norme sont toujours bornées. Les parties finies sont bornées. Une réunion finie de parties bornées est bornée. Une droite n'est pas bornée.

Le diamètre d'une partie bornée est toujours un réel ≥ 0, mais il dépend de la norme considérée.

Corollaire 2.7 (parties denses)**.** *Soit $X \subset \mathbb{R}^m$ une partie. Les propriétés suivantes sont équivalentes :*

(1) Il existe une norme $\|\cdot\|_0$ sur $\mathbb{R}^m$ telle que pour tout $u \in \mathbb{R}^m$, pour tout $\varepsilon > 0$, l'intersection $B_{\|\cdot\|_0}(u,\varepsilon) \cap X$ est $\neq \emptyset$ (en d'autres termes : il existe $v \in X$ tel que $\|u-v\|_0 < \varepsilon$).

(2) Pour toute norme $\|\cdot\|$ sur $\mathbb{R}^m$, pour tout $u \in \mathbb{R}^m$, pour tout $\varepsilon > 0$, l'intersection $B_{\|\cdot\|}(u,\varepsilon) \cap X$ est $\neq \emptyset$ (en d'autres termes : il existe $v \in X$ tel que $\|u-v\| < \varepsilon$).

Une partie X qui possède les propriétés ci-dessus est dite dense dans $\mathbb{R}^m$.

Ainsi une partie est dense si elle visite toutes les boules ouvertes de $\mathbb{R}^m$.

Exemple 2.8. Le sous-ensemble $\mathbb{Q}^2$ est dense dans $\mathbb{R}^2$.

3. Suites convergentes dans $\mathbb{R}^m$.

Commençons par une remarque sur une interprétation utile des suites de $\mathbb{R}^m$. Par exemple soit $u : \mathbb{N} \to \mathbb{R}^2$ une suite de points du plan. Alors pour tout $k \in \mathbb{N}$ on peut écrire u_k sous la forme $u_k = (x_k, y_k)$. Nous obtenons ainsi deux suites réelles $(x_k)_k$ et $(y_k)_k$. Ce sont les suites formées des composantes du vecteurs u_k. Réciproquement si nous nous donnons deux suites réelles $(a_k)_k$ et $(b_k)_k$ nous pouvons considérer la suite $(A_k)_k$ de points du plan $\mathbb{R}^2$ définie par $A_k = (a_k, b_k)$. Pour finir : on voit qu'une suite de $\mathbb{R}^2$ revient à la donnée d'un couple de suites réelles. Et plus généralement une suite de $\mathbb{R}^m$ revient à la donnée de m suites réelles (formées de chacune des composantes des points de la suite). Les propriétés des suites $(u_n)_n$ de $\mathbb{R}^m$ peuvent toujours s'exprimer par des propriétés sur les m suites formées par leurs composantes.

Définition 3.1 (suites bornées)**.** Considérons une suite $u : \mathbb{N} \to \mathbb{R}^m$. On dit que u est *bornée* si l'ensemble de ses valeurs $V(u) = \{u_n, n \in \mathbb{N}\}$ est borné. Autrement dit $\exists M \in \mathbb{R}^+$ tel que $\forall n \in \mathbb{N}$ on a $u_n \in [-M; M]^m$.

On remarque qu'une suite $(u_n)_n$ est bornée si et seulement si chacune des m suites formées par les composantes de u_n est bornée. Par exemple une suite $\big(u_n = (x_n, y_n)\big)_n$ de $\mathbb{R}^2$ est bornée exactement lorsque $(x_n)_n$ et $(y_n)_n$ sont bornées.

Définition 3.2. L'espace $\mathbb{R}^m$ est muni d'une norme $\|\cdot\|$, on note d la distance associée.

Soit $u : \mathbb{N} \to \mathbb{R}^m$ une suite et soit $v \in \mathbb{R}^m$ un vecteur. On dit que $(u_k)_k$ *converge vers* v *(pour la distance* d*)* si la suite $(d(u_k, v))_k$ tend vers 0, autrement dit $\|u_k - v\| \to 0$.

Lemme 3.3 (unicité de la limite)**.** *Si $(u_k)_k$ converge vers v et vers v' (pour la même distance d), alors $v = v'$.*

On dira que v est **la** limite *de la suite $(u_k)_k$, et on notera $u_k \to v$, ou $\lim_{k\to\infty} u_k = v$.*

Démonstration. Par inégalité triangulaire $d(v, v') \leq d(v, u_k) + d(u_k, v')$. La suite $\big(d(v, u_k) + d(u_k, v')\big)_k$ tend vers 0, donc $d(v, v') \leq 0$, donc $d(v, v') = 0$ et $v = v'$. □

Exemple 3.4. Dans $\mathbb{R}^2$ muni de la norme $\|\cdot\|_\infty$, on considère la suite $u_k = (\frac{2k}{1+k}, \frac{k^2 - \sqrt{k}}{(1+k)^2})$. Alors $(u_k)_k$ converge et $u_k \to (2, 1)$.

Lemme 3.5. *Soient N, N' deux normes sur $\mathbb{R}^m$ et soit $K > 0$ un réel tel que $N' \leq K.N$. Soit d'autre part $(u_k)_k$ une suite de $\mathbb{R}^m$ et soit $v \in \mathbb{R}^m$ un vecteur fixé.*

Si $\lim_{k\to\infty} u_k = v$ pour la distance d, alors $\lim_{k\to\infty} u_k = v$ pour la distance d'.

Démonstration. Evident puisque $d'(u_k, v) \leq K\, d(v, u_k)$ et $Kd(v, u_k) \to 0$. □

Corollaire 3.6 (unicité de la convergence)**.** *Soit $(u_k)_k$ une suite de $\mathbb{R}^m$ et soit $v \in \mathbb{R}^m$ un vecteur fixé. Les propriétés suivantes sont équivalentes :*

(1) $u_k \to v$ pour une norme sur $\mathbb{R}^m$.

(2) $u_k \to v$ pour toute norme sur $\mathbb{R}^m$.

Ainsi la convergence d'une suite ne dépend pas de la norme choisie sur $\mathbb{R}^m$. En appliquant le corollaire pour la norme $\|\cdot\|_\infty$ nous obtenons une caractérisation très concrète de la convergence des suites, qui n'utilise plus aucune norme :

Proposition 3.7 (convergence composante par composante)**.** *Soit $(u_k)_k$ une suite de $\mathbb{R}^m$ et soit $v \in \mathbb{R}^m$ un vecteur fixé.*

(1) Si $m = 2$, $u_k = (x_k, y_k)$ et $v = (a, b)$, alors $\boxed{u_k \to v \iff x_k \to a \text{ et } y_k \to b}$.

(2) Si $m = 3$, $u_k = (x_k, y_k, z_k)$ et $v = (a, b, c)$, alors $\boxed{u_k \to v \iff x_k \to a,\ y_k \to b \text{ et } z_k \to c}$.

(3) Si m est quelconque alors le même résultat est valable (mais plus délicat à énoncer en notations) : $u_k \to v \iff$

[1] la première composante de u_k tend vers la première composante de v,

[2] la deuxième composante de u_k tend vers la deuxième composante de v,

[ ] ...

[i] la i-ième composante de u_k tend vers la i-ième composante de v,

[ ] ...

[m] et pour finir la m-ième composante de u_k tend vers la m-ième composante de v.

Exemple 3.8. La suite $\left((\frac{2n^2+5n-3}{3n^2+1}, \sqrt{n^2+n+1}-\sqrt{n^2+1})\right)_n$ converge vers $(\frac{2}{3}, \frac{1}{2})$ car la première composante $\frac{2n^2+5n-3}{3n^2+1}$ tend vers $\frac{2}{3}$ et la deuxième composante $\sqrt{n^2+n+1}-\sqrt{n^2+1}$ tend vers $\frac{1}{2}$.

Lemme 3.9. *Soit $(u_n)_n$ une suite de $\mathbb{R}^m$. Si la suite $(u_n)_n$ converge alors la suite $(u_n)_n$ est bornée.*

Démonstration. En effet si $u_k \to v$ alors $d(u_k, v) \to 0$ donc la suite $(d(u_k, v))_k$ est bornée, donc $(u_k)_k$ est contenue dans une boule de centre v. $\square$

Bien sûr la plupart des suites bornées ne sont pas convergentes. De plus, dès que $m > 1$, il n'y a pas de notion de suite monotone dans $\mathbb{R}^m$, car il n'y a pas d'ordre privilégié sur $\mathbb{R}^m$. Il y a cependant une notion de "monotonie" restreinte mais utile (voir exercice 2.12).

Proposition 3.10 (opérations sur les suites). *Soit $(u_k)_k$ une suite de $\mathbb{R}^m$ et soit $v \in \mathbb{R}^m$ un vecteur fixé. On suppose que $u_k \to v$.*

Alors pour tout réel λ la suite "produit" $(\lambda u_k)_k$ converge, et sa limite est le vecteur λv.

En d'autres termes $\boxed{\lim_{k\to+\infty} \lambda u_k = \lambda \lim_{k\to+\infty} u_k}$.

Plus généralement si $(\lambda_k)_k$ est une suite numérique réelle telle que $\lambda_k \to \lambda$ alors

$$\lim_{k\to+\infty} \lambda_k u_k = \lambda \lim_{k\to+\infty} u_k \text{ soit } \lim_{k\to+\infty} \lambda_k u_k = (\lim_{k\to+\infty} \lambda_k)(\lim_{k\to+\infty} u_k)$$

De plus supposons que $(u'_k)_k$ est une autre suite convergente de $\mathbb{R}^m$ et soit $v' \in \mathbb{R}^m$ la limite de $(u'_k)_k$. Alors :

(1) La suite "somme" $(u_k + u'_k)_k$ converge, et sa limite est le vecteur $v + v'$.

En d'autres termes $\boxed{\lim_{k\to+\infty} u_k + u'_k = \lim_{k\to+\infty} u_k + \lim_{k\to+\infty} u'_k}$.

(2) La suite des produits scalaires $(u_k \cdot u'_k)_k$ converge, et sa limite est le scalaire $v \cdot v'$. En d'autres termes $\boxed{\lim_{k\to+\infty} u_k \cdot u'_k = (\lim_{k\to+\infty} u_k) \cdot (\lim_{k\to+\infty} u'_k)}$.

Démonstration. On peut soit fixer une norme, soit raisonner composantes par composantes (c'est à dire utiliser $\|\cdot\|_\infty$). Si N est une norme fixée, on a

$$N(\lambda_k u_k - \lambda v) = N\big(\lambda_k(u_k - v) + (\lambda_k - \lambda)v\big) \leq |\lambda_k| N(u_k - v) + |\lambda_k - \lambda| N(v).$$

Comme $N(u_k - v) \to 0$ et $|\lambda_k - \lambda| \to 0$ on voit que $N(\lambda_k u_k - \lambda v) \to 0$. De même :

$$N\big((u_k + u'_k) - (v + v')\big) = N\big((u_k - v) + (u'_k - v')\big)\big) \leq N(u_k - v) + N(u'_k - v')$$

Et donc si $u_k \to v$ et $u'_k \to 0$ on aura $N\big((u_k + u'_k) - (v + v')\big) \to 0$.

Enfin, en utilisant la décomposition $u_k \cdot u'_k - v \cdot v' = u_k \cdot u'_k - v \cdot u'_k + v \cdot u'_k - v \cdot v'$ puis Cauchy-Schwarz :
$|u_k \cdot u'_k - v \cdot v'| = |(u_k - v) \cdot u'_k + v \cdot (u'_k - v')| \leq |(u_k - v) \cdot u'_k| + |v \cdot (u'_k - v')| \leq \|u_k - v\|_2 \|u'_k\|_2 + \|v\|_2 \|u'_k - v'\|_2$
ce qui permet de conclure puisque comme $u'_k \to v'$, nous savons que cette suite est bornée. □

4. Ouverts et fermés.

Définition 4.1 (voisinage)**.** Soit $\| \cdot \|$ une norme sur $\mathbb{R}^m$. Soit $V \subset \mathbb{R}^m$ une partie. Soit $A \in \mathbb{R}^m$.

On dit que V *est voisinage de* A s'il existe un réel $r > 0$ tel que la boule ouverte $B(A, r)$ est contenue dans V [donc en particulier $A \in V$!].

Définition 4.2 (ouverts, fermés)**.** Soit $\| \cdot \|$ une norme sur $\mathbb{R}^m$.

Soit $O \subset \mathbb{R}^m$ une partie.

On dit que O est *une partie ouverte* (ou un *ouvert*) si, pour tout $A \in O$, O est voisinage de A. En d'autres termes, O est ouvert si O est voisinage de chacun de ses points.

Explicitement : $\boxed{O \text{ est ouvert si pour tout } A \in O \text{ il existe un réel } r\,(= r_A!) > 0 \text{ tel que } B(A, r) \subset O}$.

Soit $F \subset \mathbb{R}^m$ une partie.

On dit que F est *une partie fermée* (ou un *fermé*) si $\mathbb{R}^m \setminus F$ est ouvert. Bref, les fermés sont les complémentaires d'ouverts.

Ainsi, quand un ouvert contient un point A, il contient avec A toute une petite boule ouverte de centre A.

Remarquons que pour tout ouvert O le complémentaire $F := \mathbb{R}^m \setminus O$ est fermé ! Car le complémentaire de F est ... O, qui est bien ouvert.

Enfin de manière évidente $\emptyset$ et $\mathbb{R}^m$ sont ouverts et fermés. (On peut démontrer qu'il n'y a pas d'autres parties à la fois ouvertes et fermées - ce sera fait au S5)

Exemple 4.3 (demi-plan ouvert, demi-plan fermé)**.** On munit $\mathbb{R}^2$ de $\| \cdot \|_\infty$. Soit $O = \{(x, y) \in \mathbb{R}^2, y > 0\}$ et soit $F = \{(x, y) \in \mathbb{R}^2, y \geq 0\}$. Alors O est ouvert et F est fermé.

En effet si $A \in O$ alors $A = (a, b)$ avec $b > 0$. La boule ouvert $B(A, b)$ est contenue dans O, car pour $M = (x, y) \in B(A, b)$ on doit avoir $b - y \leq |b - y| \leq \max(|a - x|, |b - y|) < b$. Et donc $b - y < b$, soit $y > 0$, ce qui montre bien que $M \in O$. Puisqu'avec chacun de ses points A, la partie O contient toute une boule ouverte de centre A, c'est que O est ouvert.

Pour montrer que F est fermé il suffit de constater que $\mathbb{R}^2 \setminus F$ est le demi-plan ouvert $\{(x, y) \in \mathbb{R}^2, y < 0\}$ - tout aussi ouvert que O (par le même argument).

Lemme 4.4. *Les boules ouvertes sont des parties ouvertes, les boules fermées sont des parties fermées. Les singletons $\{u\}$ sont fermés.*

Démonstration. Soit $v \in B(u, r)$. Alors $d(u, v) < r$: posons $\varepsilon = r - d(u, v)$, donc $\varepsilon > 0$. Par inégalité triangulaire si $d(v, w) < \varepsilon$ alors $d(u, w) \leq d(u, v) + d(v, w) < (r - \varepsilon) + \varepsilon = r$. Ceci prouve que $B(v, \varepsilon) \subset B(u, r)$. Donc $B(u, r)$ est voisinage de chacun de ses points.

Vérifions maintenant que $\bar{B}(u, r)$ est fermée, et pour cela que $\mathbb{R}^m \setminus \bar{B}(u, r)$ est ouvert : soit donc $v \in \mathbb{R}^m \setminus \bar{B}(u, r)$. Donc $d(v, u) > r$: posons $\varepsilon = d(u, v) - r$, donc $\varepsilon > 0$. Par minoration triangulaire si $d(v, w) < \varepsilon$ alors $d(u, w) \geq d(u, v) - d(v, w) > (r + \varepsilon) - \varepsilon = r$. Ceci prouve que $B(v, \varepsilon) \subset \mathbb{R}^m \setminus \bar{B}(u, r)$.

Donc $\mathbb{R}^m \setminus \bar{B}(u,r)$ est ouvert. Les singletons $\{u\}$ sont des boules fermées de rayon nul : $\{u\} = \bar{B}(u,0)$, donc les singletons sont fermés. $\square$

On démontre facilement que les sous-espaces vectoriels de $\mathbb{R}^m$ sont fermés. Les sous-espaces affines aussi. Notamment les droites affines (voir exercice 2.14.2.b)).

Proposition 4.5 (propriétés ensemblistes)**.**

1. *Une réunion quelconque d'ouverts est un ouvert. Une intersection* finie *d'ouverts est un ouvert.*
2. *Une réunion* finie *de fermés est un fermé. Une intersection quelconque de fermés est un fermé.*

Démonstration. Soit $(O_i)_{i\in I}$ une famille d'ouverts de $\mathbb{R}^m$, et soit $O = \cup_{i\in I} O_i$ leur réunion. Si $u \in O$ alors il existe un $i_0 \in I$ tel que $u \in O_{i_0}$ et donc comme O_{i_0} est ouvert il existe en plus un $\varepsilon > 0$ tel que $B(u,\varepsilon) \subset O_{i_0}$. Comme $O_{i_0} \subset O$ on a aussi $B(u,\varepsilon) \subset O$: donc O est voisinage de u.

Soit $(O_i)_{i=1,\dots,n}$ une famille finie d'ouverts de $\mathbb{R}^m$, et soit $O = \cap_{i=1}^{i=n} O_i$ leur intersection. Si $u \in O$ alors pour tout $i = 1,\dots,n$ on a $u \in O_i$ et donc comme O_i est ouvert il existe en plus un $\varepsilon_i > 0$ tel que $B(u,\varepsilon_i) \subset O_i$. Posons $\varepsilon = \min(\varepsilon_i, i = 1,\dots,n)$, alors $\varepsilon > 0$ et $B(u,\varepsilon) \subset B(u,\varepsilon_i)$ pour tout $i = 1,\dots,n$. Donc $B(u,\varepsilon) \subset \cap_{i=1}^{i=n} B(u,\varepsilon_i) \subset \cap_{i=1}^{i=n} O_i$. On a donc $B(u,\varepsilon) \subset O$: donc O est voisinage de u.

L'énoncé sur les fermés s'obtient par passage au complémentaire comme dans le cas de $\mathbb{R}$. $\square$

Corollaire 4.6. *Soit $O \subset \mathbb{R}^m$ une partie. Alors O est ouvert $\iff$ O est réunion d'une famille de boules ouvertes.*

Par exemple dans $\mathbb{R}^2$ les ouverts pour $\|\cdot\|_\infty$ sont des unions quelconques de carrés ouverts. Et les ouverts pour $\|\cdot\|_2$ sont des unions quelconques de disques ouverts. La question naturelle qui se pose est alors de savoir si les ouverts pour $\|\cdot\|_\infty$ et les ouverts pour $\|\cdot\|_2$ sont les mêmes parties.

Lemme 4.7 (boules concentriques)**.** *Soit N, N' deux normes sur $\mathbb{R}^m$. Soit K un réel > 0 tel que $N' \leq K.N$.*

1. *Pour tout $u \in \mathbb{R}^m$ et tout réel $r' > 0$ on a $B_N(u, \dfrac{r'}{K}) \subset B_{N'}(u,r')$.*
2. *Si O est ouvert pour N' alors O est ouvert pour N.*

Par exemple, comme $\|\cdot\|_2 \leq M_\infty \|\cdot\|_\infty$ (avec $M_\infty(\|\cdot\|_2) = n$), le Lemme 4.7[2] nous dit que tout disque ouvert est une réunion de carrés ouverts ! Et comme $\|\cdot\|_\infty \leq \|\cdot\|_2$, nous obtenons également que tout carré ouvert est une réunion de disques ouverts.

Démonstration. Si $v \in B_N(u, \dfrac{r'}{K})$ alors $N(v-u) < \dfrac{r'}{K}$. Comme $N' \leq K.N$ nous obtenons $N'(v-u) \leq K.N(v-u) < K\dfrac{r'}{K} = r'$. Donc $v \in B_{N'}(u,r')$. Nous venons de montrer l'inclusion $B_N(u, \dfrac{r'}{K}) \subset B_{N'}(u,r')$ (on obtiendrait de la même manière l'inclusion des boules fermés correspondantes).

Si O est ouvert pour N' alors pour tout $u \in O$, il existe $r' > 0$ tel que O contient la boule ouverte $B_{N'}(u,r')$. Mais comme $B_N(u, \dfrac{r'}{K}) \subset B_{N'}(u,r')$ nous voyons que O contient également la boule ouverte $B_N(u,r)$ avec $r = \dfrac{r'}{K}$. Ainsi O est voisinage de u pour la distance associée à N, et finalement O est ouvert pour N. $\square$

En fait, puisque nous avons admis que deux normes quelconques sont toujours comparables (voir Théorème 2.4) nous déduisons du Lemme 4.7 le résultat fondamental suivant

Théorème 4.8. *Soient N, N' deux normes sur $\mathbb{R}^m$. Soit $O \subset \mathbb{R}^m$ une partie.*

Alors O est ouverte pour $N \iff O$ est ouverte pour N'.

La topologie *de $\mathbb{R}^m$ est alors l'ensemble des parties qui sont ouvertes pour* **une** *- et donc* **toute** *- norme sur $\mathbb{R}^m$.*

Bref : il y a peut-être plusieurs normes, mais il n'y a qu'une notion d'ouverts (associés) sur $\mathbb{R}^m$.

Comme il n'y a qu'une notion d'ouverts, il n'y a par conséquent qu'une seule notion de fermés. Mais il est parfois regrettable de devoir passer au complémentaire pour vérifier la définition de "fermé". La proposition suivante donne une caractérisation *intrinsèque* des fermés (ne faisant pas appel à la notion d'ouvert...) :

Proposition 4.9 (caractérisation par les suites)**.** *Soit $F \subset \mathbb{R}^m$ une partie. Les propriétés suivantes sont équivalentes :*

(1) F est fermée.

(2) Pour toute suite $(u_n)_n$ de F qui est convergente dans $\mathbb{R}^m$, la limite $\lim_n u_n$ appartient à F.

Quand on parle de "suite de F" on veut dire que $(u_k)_k$ est une suite de $\mathbb{R}^m$ telle que pour tout $k \in \mathbb{N}$ on a $u_k \in F$. Cette propriété sur les suites est bien une idée de fermeture : les suites de F convergente dans $\mathbb{R}^m$ ne peuvent pas s'échapper de F, même à la limite.

Démonstration. Posons $O = \mathbb{R}^m \setminus F$. Donc F non-fermée $\iff O$ non ouvert $\iff \exists u \notin F$ tel que toute boule de centre u contient un élément de F.

Si F non fermée, prenons donc un u comme ci-dessus, et pour chaque entier $k \in \mathbb{N}$ soit u_n un point de $F \cap B(u, \dfrac{1}{1+n})$. Alors nous avons obtenu une suite $(u_n)_n$ d'éléments de F, qui vérifie $u_n \to u$, et enfin la limite u n'est pas dans F.

Réciproquement s'il existe une suite $(u_n)_n$ de F qui est convergente dans $\mathbb{R}^m$ vers une limite u n'appartenant pas à F, alors toute boule ouverte $B(u, \varepsilon)$ contient les termes u_n pour tout rang assez grand, en particulier contient des éléments de F, et donc F n'est pas fermé.

$\square$

Exemple 4.10. Pour a, b, c, d des réels tels que $a \leq b$ et $c \leq d$, considérons le "rectangle" $R = [a; b] \times [c; d]$. Alors $(x, y) \in R \iff a \leq x \leq b$ et $c \leq y \leq d$. Soit maintenant $(x_n, y_n)_n$ une suite de R qui converge vers $(\alpha, \beta) \in \mathbb{R}^2$. Alors $x_n \to \alpha$ et $a \leq x_n \leq b$ pour tout entier n : par passage à la limite dans les inégalités larges (ou pour le dire en utilisant la terminologie de cette UE : "comme $[a; b]$ est un fermé de $\mathbb{R}$") on obtient $\alpha \in [a; b]$. Par le même argument $\beta \in [c; d]$. Pour conclure : la limite (α, β) appartient à R.

Nous avons démontré que R est (séquentiellement) fermé.

Proposition 4.11. *Soit $X \subset \mathbb{R}^m$ une partie quelconque. Alors les propriétés suivantes sont équivalentes :*

(1) X est dense

(2) Pour tout $u \in \mathbb{R}^m$ il existe une suite $(u_n)_n$ d'éléments de X telle que $u_n \to u$.

(3) Le seul fermé qui contient X est $\mathbb{R}^m$.

Démonstration. Supposons X dense dans $\mathbb{R}^m$. Pour $u \in \mathbb{R}^m$ fixé il existe par hypothèse un élément u_n dans $X \cap B(u, \dfrac{1}{1+n})$. Alors $d(u, u_n)$ et donc $(u_n)_n$ est une suite de X qui tend vers u. Supposons maintenant que pour tout $u \in \mathbb{R}^m$ il existe une suite $(u_n)_n$ d'éléments de X telle que $u_n \to u$, et soit F un fermé de $\mathbb{R}^m$ qui contient X. Alors pour $u \in \mathbb{R}^m$ considérons une suite $(u_n)_n$ d'éléments de X telle que $u_n \to u$. Comme $X \subset F$ chaque u_n est dans F. Comme $u_n \to u$ et F est fermé nous obtenons $u \in F$. Donc $\mathbb{R}^m \subset F$ et $F = \mathbb{R}^m$.

Enfin montrons (1) $\Rightarrow$ (3) sous la forme contraposée $\neg(3) \Rightarrow \neg(1)$. Soit donc F un fermé $\neq \mathbb{R}^m$ qui contient X. Alors $O = \mathbb{R}^m \setminus F$ est un ouvert non vide, soit donc $u \in O$ et $r > 0$ tel que $B(u,r) \subset O$. Nous avons alors $B(u,r) \cap F = \emptyset$ et a fortiori $B(u,r) \cap X = \emptyset$. □

Troisième partie 3. Fonctions de plusieurs variables à valeurs réelles : continuité.

1. Exemples de fonctions de plusieurs variables, opérations.

Définition 1.1. Soit $X \subset \mathbb{R}^m$. Une *fonction de m variables (à valeur réelles) définie sur X* est une application $f : X \subset \mathbb{R}^m \to \mathbb{R}$, c'est à dire que f associe à tout $u \in X$ un réel $f(u)$. On dit que $f(u)$ est *l'image de u par f*. Si $c \in \mathbb{R}$ et si $f(u) = c$, on dit que u est *un antécédent de c par f*.

L'ensemble des *valeurs* prises par f sur X est l'ensemble $f(X)$ de tous les réels c qui admettent au moins un antécédent par f : $f(X) = \{c \in \mathbb{R}, \exists u \in X \text{ tel que } f(u) = c\}$. On peut considérer des images directes plus générales : pour $Y \subset X$, *l'image directe de Y par f* est $f(Y) = \{c \in \mathbb{R}, \exists u \in Y \text{ tel que } f(u) = c\}$.

Voici des exemples de fonctions de plusieurs variables, chacune est définie par une formule, comme ensemble de départ X on prend le domaine de définition de la formule :

$$(x,y) \mapsto 1 + 2x + 3y \ (X = \mathbb{R}^2)\ ,\ (x,y) \mapsto \sqrt{xy}\ (X = \mathbb{R}^+ \times \mathbb{R}^+ \cup \mathbb{R}^- \times \mathbb{R}^-)\ ,$$

$$(x,y,z) \mapsto \frac{1}{x^2+y^2+z^2}\ (X = \mathbb{R}^3 \setminus \{(0,0,0)\})$$

Définition 1.2 (formes coordonnées, formes linéaires, formes affines)**.** Une *forme affine* est une fonction $f : \mathbb{R}^m \to \mathbb{R}$ de la forme $u = (x_1, \dots, x_m) \mapsto a_1x_1 + \dots + a_mx_m + b$, où $(a_1, \dots, a_m, b) \in \mathbb{R}^{d+1}$ est fixé. Par exemple $(x,y) \mapsto 2x - 3y + 5$ ou $(x,y,z) \mapsto x + 2y - 3z - 6$. Les formes affines de la forme $u = (x_1, \dots, x_m) \mapsto a_1x_1 + \dots + a_mx_m$ sont appelées *formes linéaires* - et ce sont bien des applications linéaires. La i-ième forme coordonnée est la forme linéaire $u = (x_1, \dots, x_m) \mapsto x_i$ (pour $i = 1, \dots, m$).

Définition 1.3 (monômes et polynomes)**.** Pour $k \in \mathbb{N}$, un *monôme de degré k* est une fonction de la forme $u \mapsto aM(u)$, où $a \in \mathbb{R}^*$ est une constante, et $u \mapsto M(u)$ est le produit de k formes coordonnées, non nécessairement distinctes. Par exemple $(x,y) \mapsto 2x$ est un monôme de degré 1 (donc une forme linéaire...), $(x,y) \mapsto -x \times x \times x \times y \times y = -x^3y^2$ est un monôme de degré 5, $(x,y,z) \mapsto 4xyz^2$ est un monôme de degré 4, etc...

Par définition un *polynôme* ou *fonction polynomiale* est une somme finie de monômes. Si c'est une somme vide, c'est la fonction nulle, et par convention on définit le *degré* de la fonction nulle comme $-\infty$. Sinon, $f(u) = \sum_{k=1}^{k=p} M_k(u)$, et on définit le *degré* du polynôme comme le plus haut degré des monômes $M_k(u)$ intervenant dans la somme.

Par exemple $(x,y) \mapsto 1+2x-3y+xy-5x^2y$ est un polynôme de degré 3 en deux variables. Et $(x,y,z) \mapsto 1-x+2y+z+xz+yz+x^2-2y^2-z^2+xyz-x^2y^2z^2+x^3yz^2$ est un polynôme de degré 6 en deux variables.

On remarque que les fonctions constantes non nulles sont les polynômes de degré 0. Les formes affines sont les polynômes de degré ≤ 1.

Définition 1.4 (graphe d'une fonction de deux variables)**.** Soit $X \subset \mathbb{R}^2$ une partie quelconque et soit $f : X \to \mathbb{R}$ une fonction quelconque. Alors le *graphe* de f est l'ensemble $\Gamma(f) \subset \mathbb{R}^3$ défini par

$$\Gamma(f) = \{(x,y,z) \in \mathbb{R}^3, (x,y) \in X \text{ et } z = f(x,y)\}.$$

Le graphe est donc contenu dans le "cylindre" $X \times \mathbb{R}$ (de base horizontale X). Au lieu de "graphe" on dit parfois *surface représentative*.

Exemple 1.5. Considérons la forme linéaire $f(x,y) = 2x+3y$. Alors f est définie sur $\mathbb{R}^2$ et le graphe de f est $\Gamma(f) = \{(x,y,z) \in \mathbb{R}^3, z = 2x+3y\}$. Donc $\Gamma(f)$ est le plan de l'espace d'équation $2x+3y-z=0$ (méditer sur l'analogie avec : "le graphe d'une fonction affine $x \mapsto ax+b$ est un plan de $\mathbb{R}^2$").

Par certaines opérations on peut produire de nouvelles fonctions de plusieurs variables :

Définition 1.6 (somme, produit, composition)**.**

(1) Soit $f, g : X \subset \mathbb{R}^m \to \mathbb{R}$ deux fonctions de d variables.
La *somme de f et g* est la fonction $f+g : u \mapsto f(u)+g(u)$.
Le *produit de f et g* est la fonction $fg : u \mapsto f(u)g(u)$.

(2) Soit $g : X \subset \mathbb{R}^m \to \mathbb{R}$ une fonction de d variables. Soit $\phi : D \subset \mathbb{R} \to \mathbb{R}$ une fonction d'une variable définie sur D tel que $g(X) \subset D$. Alors la *composée* $\phi \circ g$ est définie par la formule $X \ni u \mapsto \phi\big(g(u)\big)$.

Par exemple les polynômes s'obtiennent à partir des fonctions affines (ou même des formes coordonnées) en itérant des produits et des sommes. Et $(x,y) \mapsto \dfrac{1}{x^2+y^2}$ est la composée $\phi \circ f$ du polynôme $f : (x,y) \mapsto x^2+y^2$ avec la fonction $\phi : t \mapsto \dfrac{1}{t}$.

2. Continuité : formulations séquentielles ou métriques.

2.1. Définitions équivalentes.

Lemme 2.1. *Munissons $\mathbb{R}^m$ d'une norme et de la distance associée.*

Soit $f : X \subset \mathbb{R}^m \to \mathbb{R}$ une fonction définie sur une partie X de $\mathbb{R}^m$. Soit A un point de X. Alors les propriétés suivantes sont équivalentes :

(1) (continuité séquentielle) Pour toute suite $(A_k)_k$ telle que $A_k \in X$ $(\forall k \in \mathbb{N})$ et $A_k \to A$, la suite $(f(A_k))_k$ est convergente, de limite $f(A)$.

(2) (continuité métrique) Pour tout $\varepsilon > 0$, il existe un $\alpha > 0$ tel que si $M \in X$ est α-proche de A, alors $f(M)$ est ε-proche de $f(A)$.

Démonstration. (2) $\Rightarrow$ (1) : Supposons vraie la continuité $\forall\varepsilon\exists\alpha$ en A. Alors soit $(A_k)_k$ une suite de X telle que $A_k \to A$, et fixons $\varepsilon > 0$. Il existe un α vérifiant la propriété : si $M \in X$ est α-proche de A, alors $f(M)$ est ε-proche de $f(A)$. Mais pour k assez grand on a $d(A_k, A) < \alpha$, donc en fait aussi $d(f(A_k), f(A)) < \varepsilon$. Ceci montre que $d(f(A_k), f(A)) \to 0$, soit $f(A_k) \to f(A)$.

$\neg(2) \Rightarrow \neg(1)$: Supposons fausse la continuité $\forall\varepsilon\exists\alpha$ en A. Donc $\exists\varepsilon_0 > 0$ tel que $\forall\alpha > 0$, il existe un point $M \in X$ tel que $d(M, A) < \alpha$ et cependant $|f(M) - f(A)| \geq \alpha$. Nous appliquons ceci pour tout α de la forme $\alpha = \dfrac{1}{1+k}$, cela nous fournit une suite $(M_k)_k$ de points de X tels que $M_k \to A$ mais $|f(M_k) - f(A)| \geq \varepsilon_0$, donc $f(M_k)$ ne tend pas vers $f(A)$. $\square$

Définition 2.2. Soit $f : X \subset \mathbb{R}^m \to \mathbb{R}$ une fonction définie sur une partie X de $\mathbb{R}^m$. Soit A un point de X.

On dit que f *est continue en* A si l'une ou l'autre des propriétés de continuité en A et vérifiée (continuité séquentielle, ou continuité métrique).

On dit que $f : X \to \mathbb{R}$ est *continue (sur X)* si elle est continue en tout point $A \in X$.

Enfin si $Y \subset X$ on dira que f est *continue sur* Y si la restriction $f_{|Y} : Y \to \mathbb{R}$ est continue.

Corollaire 2.3. *Munissons $\mathbb{R}^m$ d'une norme $\|\cdot\|$ et de la distance d associée.*

Considérons une fonction $f : X \subset \mathbb{R}^m \to \mathbb{R}$. Alors f est continue $\iff$ la propriété suivante a lieu :

$$\forall A \in X,\ \forall\varepsilon > 0,\ \exists\alpha > 0 \text{ tel que si } M \in X \text{ vérifie } d(M, A) < \alpha, \text{ alors } |f(M) - f(A)| < \varepsilon.$$

Exemple 2.4 (continuité des applications coordonnées)**.**

Les deux applications $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont continues de $\mathbb{R}^2$ dans $\mathbb{R}$.

Les trois applications $(x, y, z) \mapsto x$, $(x, y, z) \mapsto y$ et $(x, y, z) \mapsto z$ sont continues de $\mathbb{R}^3$ dans $\mathbb{R}$.

Plus généralement les m applications coordonnées $p_i : (x_1, \cdots, x_m) \mapsto x_i$ sont continues de $\mathbb{R}^m$ dans $\mathbb{R}$.

Un réservoir inépuisable de fonctions continues sera donné par la classe des fonctions Lipschitziennes, voir section suivante.

2.2. Fonctions Lipschitziennes.

Définition 2.5 (fonctions Lipschitziennes)**.** Considérons une fonction $f : X \subset \mathbb{R}^m \to \mathbb{R}$. Soit $L \geq 0$ une constante réelle. Supposons que $\|\cdot\|$ est une norme fixée sur $\mathbb{R}^m$.

Nous dirons que f est *K-Lipschitzienne (pour $\|\cdot\|$)* si pour $A, B \in X$ on a

$$|f(A) - f(B)| \leq K \|A - B\|$$

Nous dirons que f est *Lipschitzienne (pour $\|\cdot\|$)* s'il existe un réel $K \geq 0$ tel que f est K-Lipschitzienne pour $\|\cdot\|$. Les réels K tels que f est K-Lipschitzienne sont appelés les *constantes de Lipschitz de f (pour $\|\cdot\|$)*.

Par exemple on peut remarquer que les fonctions coordonnées $p_i : \mathbb{R}^m \to \mathbb{R}$ sont 1-Lipschitziennes pour les normes $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$. Voir l'exercice 3.4 pour d'autres calculs de constantes de Lipschitz.

La quantité $|f(A) - f(B)|$ s'appelle *l'accroissement (absolu) de f entre A et B*. On peut donc retenir la définition de Lipschitzien comme suit :

f est Lipschitzienne si l'accroissement de f entre deux points est majoré par une fonction linéaire de la distance entre les points (la même fonction linéaire $d \mapsto Kd$ pour toutes les paires de points !)

Lemme 2.6. *Si $f : X \subset \mathbb{R}^m \to \mathbb{R}$ est Lipschitzienne pour une norme alors f est Lipschitzienne pour toute norme.* A partir de maintenant nous dirons donc que f est Lipschitzienne sans préciser pour quelle norme.

Bien sûr la constante de Lipschitz L dépend, elle, de la norme considérée.

Lemme 2.7. *Si $f : X \subset \mathbb{R}^m \to \mathbb{R}$ est Lipschitzienne alors f est continue.*

Démonstration. Supposons f K-Lipschitzienne sur X. Pour $A \in X$ et $\varepsilon > 0$ fixé on pose $\alpha = \dfrac{\varepsilon}{K}$: alors il est vrai que si B est α-proche de A, nous pouvons être sûr que $f(B)$ est ε-proche de $f(A)$. En effet $|f(A) - f(B)| \leq Kd(A, B) < K\dfrac{\varepsilon}{K} = \varepsilon$. □

Pour aller vite : les fonctions Lipschitziennes sont continues car leurs accroissements sont majorées par une fonction linéaire de la distance, or quel que soit la constante K on a $\lim\limits_{d \to 0} Kd = 0$. Mais nous sentons qu'une fonction pourrait être continue simplement parce que ses accroissements sont majorés par une fonction $\phi(d)$ de la distance, avec $\lim\limits_{d \to 0} \phi(d) = 0$ - pas besoin que ϕ soit linéaire ! Voici un exemple très simple de fonction $\phi : d \mapsto \phi(d)$ telle que $\lim\limits_{d \to 0} \phi(d) = 0$ et pourtant ϕ n'est majoré par aucune fonction linéaire lorsque d est proche de 0 : $\phi(d) = \sqrt{d}$.

Cela nous suggère de considérer $f(x, y) = \sqrt{|x| + |y|}$: cette fonction est continue sur $\mathbb{R}^2$, elle est même Lipschitzienne sur $\mathbb{R}^2 \setminus B_{d_1}\big((0, 0), r\big)$ pour tout $r > 0$, mais elle n'est Lipschitzienne sur $B_{d_1}\big((0, 0), r\big)$ pour aucun $r > 0$. Le problème c'est que

$$\forall x > 0, \frac{|f(x, 0) - f(0, 0)|}{d_1\big((x, 0); (0, 0)\big)} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}}, \text{ borné par aucune constante } K \text{ pour } x \to 0$$

Corollaire 2.8.

(1) *Toute norme $N : \mathbb{R}^m \to \mathbb{R}$ est Lipschitzienne donc continue.*
(2) *Toute application linéaire $\phi : \mathbb{R}^m \to \mathbb{R}$ est Lipschitzienne donc continue.*
(3) *Toute application affine $\psi : \mathbb{R}^m \to \mathbb{R}$ est Lipschitzienne donc continue.*

Démonstration. (1) Par la minoration triangulaire on a $|d(0, u) - d(0, v)| \leq d(u, v)$, soit $|N(u) - N(v)| \leq N(u - v)$. Ceci veut dire que $u \mapsto N(u)$ est 1-Lipschitzienne de $(\mathbb{R}^m, N)$ vers $(\mathbb{R}, ||)$.

(2) Si $\phi(x_1, \ldots, x_m) = a_1x_1 + \cdots + a_mx_m$ alors pour $u = (x_1, \ldots, x_m)$ et $v = (y_1, \ldots, y_m)$, on a

$$\begin{aligned} |\phi(u) - \phi(v)| = |a_1(x_1 - y_1) + \cdots + a_m(x_m - y_m)| &\leq |a_1||x_1 - y_1| + \cdots + |a_m||x_m - y_m| \\ &\leq \max\big(|a_i|\big)[|x_1 - y_1| + \cdot + |x_m - y_m|] = K\|u - v\|_1 \end{aligned}$$

avec $K = \max\big(|a_i|\big)$. Donc ϕ est Lipschitzienne.

(3) Si $\psi(x_1, \ldots, x_m) = a_1x_1 + \cdots + a_mx_m + b$ alors posons $\phi(x_1, \ldots, x_m) = a_1x_1 + \cdots + a_mx_m$. Comme la constante b se simplifie, $\psi(u) - \psi(v) = \phi(u) - \phi(v)$, et donc l'accroissement de ψ est tout aussi linéaire en la distance que l'accroissement de ϕ !

□

2.3. **Continuité et opérations.**

Proposition 2.9 (continuité et opérations algébriques)**.** *Soient $f, g : X \subset \mathbb{R}^m \to \mathbb{R}$ deux fonctions continues.*

(1) (continuité de la somme) La fonction $f + g : X \to \mathbb{R}$ est continue.

(2) (continuité du produit) La fonction $fg : X \to \mathbb{R}$ est continue.

(3) (continuité du quotient) Si g ne s'annule pas sur X alors la fonction $\frac{f}{g} : X \to \mathbb{R}$ est continue.

Corollaire 2.10. *(1) Tout polynôme $f : \mathbb{R}^2 \to \mathbb{R}$ est continu.*

(2) Toute fonction $f(x, y) = \dfrac{P(x, y)}{Q(x, y)}$ quotient de deux polynômes est continue sur son ensemble de définition.

Proposition 2.11 (composée $\mathbb{R}^2 \to \mathbb{R} \to \mathbb{R}$)**.** *Soit $g : X \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables. Soit $\phi : D \subset \mathbb{R} \to \mathbb{R}$ une fonction d'une variable avec $g(X) \subset D$. Si g et ϕ sont continues, alors $\phi \circ g$ est continue.*

Exemple 2.12. La fonction $(x, y) \mapsto \sqrt{x^2 + y^2 + 1}$ est continue sur $\mathbb{R}^2$, comme composée du polynôme $P(x, y) = x^2 + y^2 + 1$ et de la fonction $t \mapsto \sqrt{t}$.

3. Continuité et topologie.

Comme dans le cas des fonctions $\mathbb{R} \to \mathbb{R}$, on peut caractériser la continuité de $f : X \subset \mathbb{R}^m \to \mathbb{R}$ en utilisant la topologie. Nous le faisons dans le cas où X, le domaine de définition de f, est simple (voir S5 pour le cas d'un X arbitraire) :

Proposition 3.1. *(1) Soit $O \subset \mathbb{R}^m$ un ouvert, et soit $f : O \to \mathbb{R}$. Alors f est continue $\iff$ pour tout ouvert $U \subset \mathbb{R}$, l'image réciproque $f^{-1}(U)$ est un ouvert de $\mathbb{R}^m$.*

(2) Soit $F \subset \mathbb{R}^m$ un fermé, et soit $f : F \to \mathbb{R}$. Alors f est continue $\iff$ pour tout fermé $G \subset \mathbb{R}$, l'image réciproque $f^{-1}(G)$ est un fermé de $\mathbb{R}^m$.

(3) Soit $f : \mathbb{R}^m \to \mathbb{R}$. Alors f est continue $\iff$ pour tout ouvert $U \subset \mathbb{R}$, l'image réciproque $f^{-1}(U)$ est un ouvert de $\mathbb{R}^m$ $\iff$ pour tout fermé $G \subset \mathbb{R}$, l'image réciproque $f^{-1}(G)$ est un fermé de $\mathbb{R}^m$.

Démonstration. (1) Supposons f continue et soit $U \subset \mathbb{R}$ un ouvert. Pour montrer que $f^{-1}(U)$ est ouvert soit $u \in f^{-1}(U)$: montrons qu'il y a toute une boule ouverte de centre u contenue dans $f^{-1}(U)$. Par hypothèse $f(u) \in U$. Mais U est un ouvert de $\mathbb{R}$, donc $\exists \varepsilon > 0$ tel que $]f(u) - \varepsilon; f(u) + \varepsilon[\subset U$. Comme f est continue, il existe alors $\alpha' > 0$ tel que pour tout $v \in B(u, \alpha')$, et $v \in O$, on a $|f(v) - f(u)| < \varepsilon$. Comme O est ouvert il existe un α_u tel que $B(u, \alpha_u) \subset O$. Donc en posant $\alpha = \min(\alpha', \alpha_0)$, nous avons : pour tout $v \in B(u, \alpha)$, $f(v) \in]f(u) - \varepsilon; f(u) + \varepsilon[$, et pour finir $f(v) \in U$. Ainsi nous avons trouvé un $\alpha > 0$ tel que si v est α-proche de u, alors $f(v)$ est dans U - soit v appartient à $f^{-1}(U)$: $B(u, \alpha) \subset f^{-1}(U)$. Ainsi $f^{-1}(U)$ est voisinage de u.

Réciproquement soit $f : O \to \mathbb{R}$ une fonction telle que pour tout ouvert $U \subset \mathbb{R}$, l'ensemble $f^{-1}(U)$ est un ouvert de $\mathbb{R}^2$. Pour $u \in O$ et $\varepsilon > 0$ fixé posons $U =]f(u) - \varepsilon; f(u) + \varepsilon[$: c'est un ouvert. Donc $f^{-1}(U)$ est ouvert. Or $f^{-1}(U) = \{v \in O$ tels que $f(v)$ est ε-proche de $f(u)\}$, et $u \in f^{-1}(U)$, donc il existe un $\alpha > 0$ tel que $B(u, \alpha) \subset f^{-1}(U)$, ce qui nous donne : pour tout $v \in B(u, \alpha)$, d'abord $v \in O$, ensuite $f(v)$ est ε-proche de $f(u)$. C'est la continuité de f en u.

(2) Supposons $f : F \to \mathbb{R}$ continue et soit $G \subset \mathbb{R}$ un fermé. Montrons que $f^{-1}(G)$ est séquentiellement fermé. Soit donc $(v_k)_k$ une suite de $f^{-1}(G)$ - c'est à dire telle que $f(v_k) \in G$, avec $v_k \to u$ dans $\mathbb{R}^m$. Comme F est fermé, on a en fait $u \in F$. Alors $v_k \to u$ dans F, et f est séquentiellement continue sur F, donc $f(v_k) \to f(u)$. Mais $f(v_k)$ est dans G, qui est fermé, donc G doit contenir la limite $f(u)$. C'est dire que $u \in f^{-1}(G)$, ce qui conclut.

Réciproquement supposons que f n'est pas continue, et démontrons qu'il existe un fermé $G_0 \subset \mathbb{R}$ tel que l'image réciproque $f^{-1}(G_0)$ n'est pas séquentiellement fermé.

Par hypothèse il existe $u_0 \in F$ et $\varepsilon_0 > 0$ tel que pour tout $\alpha > 0$, il existe un $v \in F \cap B(u_0, \alpha)$ avec $|f(v) - f(u_0)| \geq \varepsilon_0$. Discrétisons : pour tout entier $n \in \mathbb{N}$, il existe un $v_n \in F \cap B(u_0, \dfrac{1}{1+n})$ avec $|f(v_n) - f(u_0)| \geq \varepsilon_0$. Donc $v_n \to u_0$ mais $f(v_n)$ reste dans le complémentaire de $]f(u_0) - \varepsilon_0; f(u_0) + \varepsilon_0[$. Posons donc $G_0 = \mathbb{R} \backslash]f(u_0) - \varepsilon_0; f(u_0) + \varepsilon_0[$: nous allons vérifier que $f^{-1}(G_0)$ n'est pas séquentiellement fermé. En effet $(v_n)_n$ est une suite de F telle que $f(v_n)$ reste dans le complémentaire de $]f(u_0) - \varepsilon_0; f(u_0) + \varepsilon_0[$, donc $f(v_n) \in G_0$, donc $v_n \in f^{-1}(G_0)$. On sait que $v_n \to u_0$, mais bien sûr $f(u_0) \notin G_0$, donc $u_0 \notin f^{-1}(G_0)$.

(3) Evident compte tenu de (1) et (2), puisque $\mathbb{R}^m$est à la fois ouvert et fermé.

□

Définition 3.2. Soit $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables définie sur une partie X, et soit $c \in \mathbb{R}$ fixé. Alors *l'ensemble de niveau c pour f* est la partie, notée $\mathcal{L}_c(f)$, et définie par

$$\mathcal{L}_c(f) = \{(x, y) \in X, f(x, y) = c\} \text{ (donc } \mathcal{L}_c(f) = f^{-1}(\{c\})).$$

On parle souvent aussi de *ligne de niveau c pour f*.

Pour $c = 0$, on parle souvent de l'ensemble des zéros de f au lieu d'ensemble de niveau 0 de f.

On peut bien sûr définir les ensembles de niveau $\mathcal{L}_c(f)$ pour une fonction d'un nombre de variables quelconque. Mais l'ensemble $\mathcal{L}_c(f)$ n'est véritablement une *ligne* que pour f une fonction de deux variables. Intuitivement : les deux variables x, y étant indépendantes l'une de l'autre, si on impose une contrainte $f(x, y) = c$, cela supprime un degré de liberté au point (x, y), qui doit donc se déplacer sur un espace de dimension 1 - une courbe.

Exemple 3.3 (ensembles de niveaux pour quelques types de fonction)**.**

(1) Soit $f : \mathbb{R}^m \to \mathbb{R}$ une fonction constante. Alors $\mathcal{L}_c(f) = \emptyset$ si c n'est pas la valeur prise par f, et $\mathcal{L}_c(f) = \mathbb{R}^m$ si c est la valeur prise par f.

(2) Soit $f : (x, y) \mapsto 2x - 3y + 5$. Alors $\mathcal{L}_0(f)$ est l'ensemble des $(x, y) \in \mathbb{R}^2$ tels que $2x - 3y + 5 = 0$, c'est donc la droite (affine) d'équation $2x - 3y = -5$. Plus généralement si $f(x, y) = ax + by - c$ avec $(a, b) \neq (0, 0)$ et $c \in \mathbb{R}$, on constate que l'ensemble des zéros $\mathcal{L}_0(f)$ est la droite d'équation $ax + by = c$ (droite de vecteur normal $(a, b) \neq (0, 0)$).

(3) Soit $f(x, y) = x^2 + y^2$ (polynôme de degré 2 sur $\mathbb{R}^2$). Pour $c < 0$ on voit que $\mathcal{L}_c(f) = \emptyset$. Pour $c = 0$ on a $\mathcal{L}_c(f) = \{(0, 0)\}$. Enfin si $c > 0$ alors $\mathcal{L}_c(f)$ est le cercle de $\mathbb{R}^2$ de centre $(0, 0)$ et de rayon $\sqrt{c}$.

Lemme 3.4. *Soit $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables définie sur une partie X.*

(1) Si $c \neq c'$ alors $\mathcal{L}_c(f) \cap \mathcal{L}_{c'}(f) = \emptyset$.

(2) On a $\mathcal{L}_c(f) \neq \emptyset$ si et seulement si $f(x, y)$ prend la valeur c en un certain point $A = (a; b)$ de X.

Proposition 3.5. *Considérons une fonction de deux variables $f : \mathbb{R}^2 \to \mathbb{R}$, ou plus généralement une fonction $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ définie sur une partie fermée X.*

Si f est continue alors pour toute valeur c la ligne de niveau $\mathcal{L}_c(f)$ est fermée.

Démonstration. En fait $u \in \mathcal{L}_c(f) \iff f(u) = c$, donc $\mathcal{L}_c(f) = f^{-1}(\{c\})$. Comme f est continue et le singleton $\{c\}$ est fermé, cela permet de conclure. □

Corollaire 3.6.

(1) Pour toute fonction continue $f : \mathbb{R}^2 \to \mathbb{R}$ l'ensemble des zéros de f est un fermé de $\mathbb{R}^2$.

(2) Soient $f, g : \mathbb{R}^2 \to \mathbb{R}$ deux fonctions continues. Supposons que $D \subset \mathbb{R}^2$ est une partie dense telle que $\forall u \in D$ on a $f(u) = g(u)$, alors en fait on a $f(u) = g(u)$ pour tout $u \in \mathbb{R}^m$.

Démonstration. Le premier énoncé vient de ce que l'ensemble des zéros est un cas particulier d'ensemble de niveau. Pour démontrer le deuxième énoncé considérons $h = f - g$, et remarquons que l'ensemble des zéros de h est précisément $\{u \in \mathbb{R}^m, f(u) = g(u)\}$. Comme f, g sont continues leur différence h l'est aussi. Donc l'ensemble de ses zéros est fermé. Mais cet ensemble $\{u \in \mathbb{R}^m, f(u) = g(u)\}$ contient la partie dense D : en tant que fermé contenant une partie dense il égale $\mathbb{R}^m$ entier. □

Théorème 3.7 (une fonction continue sur un rectangle fermé est bornée et atteint ses bornes)**.**

Soit $f : [a; b] \times [c; d] \to \mathbb{R}$ une fonction continue. Alors

(1) f est bornée sur $[a; b] \times [c; d]$, i.e. $f\big([a; b] \times [c; d]\big)$ est borné.

(2) f atteint sa borne supérieure et sa borne inférieure sur $[a; b] \times [c; d]$, i.e. il existe (x_1, y_1) et $(x_2, y_2) \in [a; b] \times [c; d]$ tels que $\forall (x, y) \in [a; b] \times [c; d]$, on a $f(x_1, y_1) \leq f(x, y) \leq f(x_2, y_2)$.

Démonstration. On raisonne par "dichotomie en dimension 2" : on va découper le rectangle en petits sous-rectangles et analyser la situation sur chacun des petits sous-rectangles. Plus précisément, à tout rectangle fermé $R = [s; t] \times [u; v]$ on associe quatre rectangles définis comme suit :

$$R^{--} = [s; \frac{s+t}{2}] \times [u; \frac{u+v}{2}] \ ; \ R^{+-} = [\frac{s+t}{2}; t] \times [u; \frac{u+v}{2}] \ ;$$
$$R^{-+} = [s; \frac{s+t}{2}] \times [\frac{u+v}{2}; v] \ ; \ R^{++} = [\frac{s+t}{2}; t] \times [\frac{u+v}{2}; v].$$

Il est clair qu'on a une décomposition de la forme $R = R^{--} \cup R^{-+} \cup R^{+-} \cup R^{++}$. Voir Figure 2.

On montre d'abord que f est bornée sur $[a; b] \times [c; d]$ en raisonnant par l'absurde : supposons f non bornée.

On définit alors une suite de rectangles emboîtés : posons d'abord $R_0 = [a; b] \times [c; d]$. Puis comme f n'est pas bornée sur R_0, elle ne peut pas être bornée sur chacun des quatre sous-rectangles $R_0^{--}, R_0^{+-}, R_0^{-+}, R_0^{++}$. Appelons R_1 l'un des quatre sous-rectangles $R_0^{--}, R_0^{+-}, R_0^{-+}, R_0^{++}$ sur lequel f n'est pas bornée. Reprenons l'argument de découpage sur R_1 : comme f n'est pas bornée sur R_1, la fonction f n'est pas bornée non plus sur au moins l'un des quatre sous-rectangles $R_1^{--}, R_1^{+-}, R_1^{-+}, R_1^{++}$ - appelons R_2 l'un de ces sous-rectangles.

Si nous continuons ainsi nous voyons qu'il existe une suite décroissante de rectangles $R_0 \supset R_1 \supset R_2 \supset \cdots \supset R_n \supset$ telle que

(1) $R_0 = [a; b] \times [c; d]$

(2) $\mathrm{diam}(R_n) = \frac{1}{2}\mathrm{diam}([a; b] \times [c; d])$, donc $\mathrm{diam}(R_n) \to 0$

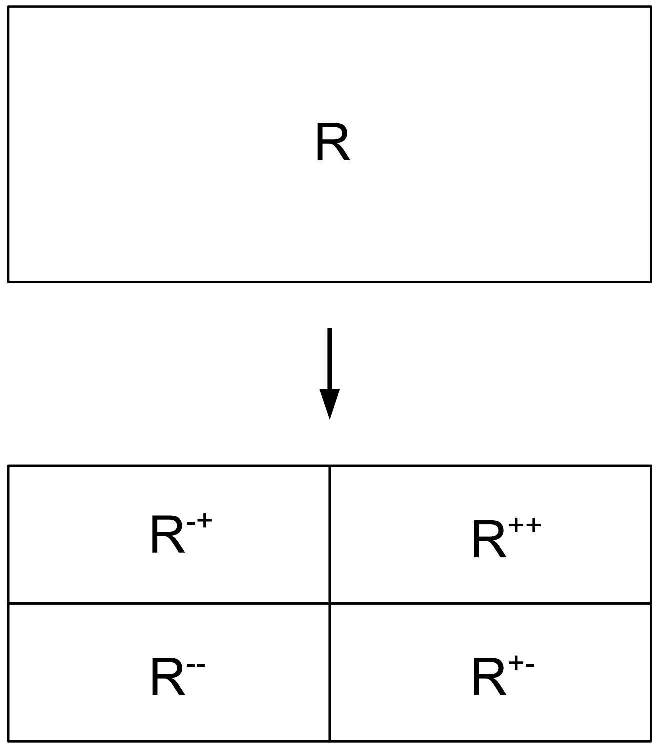


FIGURE 2. Le rectangle R se subdivise en $R = R^{--} \cup R^{-+} \cup R^{+-} \cup R^{++}$.

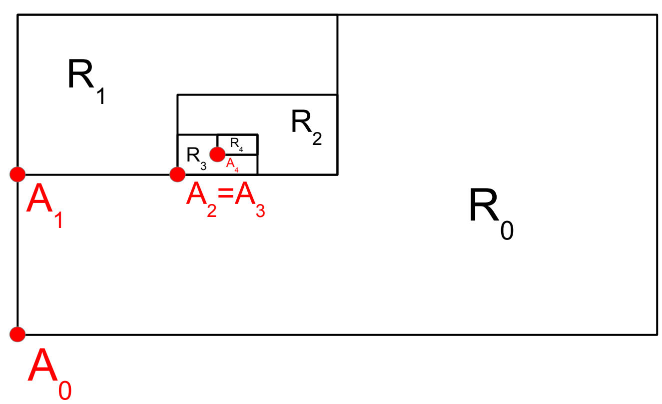


FIGURE 3. Les cinq premiers rectangles emboîtés. Le point A_n désigne le coin inférieur gauche.

(3) f n'est pas bornée sur R_n

Soit alors $A_n = (x_n, y_n)$ le coin inférieur gauche du rectangle R_n. Voir Figure 3.

Comme $R_{n+1} \subset R_n$ les deux suites $(x_n)_n$ et $(y_n)_n$ sont croissantes. Ce sont d'autre part des suites bornées ($x_n \leq b$ et $y_n \leq d$). Donc ces suites sont convergentes : soit $\bar{x} = \lim x_n$ et $\bar{y} = \lim y_n$.

On a $(x_n, y_n) \to (\bar{x}, \bar{y})$. Le rectangle $[a; b] \times [c; d]$ est fermé, donc $(\bar{x}, \bar{y}) \in [a; b] \times [c; d]$. Comme f est continue sur $[a; b] \times [c; d]$, il existe $\varepsilon_1 > 0$ tel que pour $(x, y) \in B((\bar{x}, \bar{y}), \varepsilon_1)$ on a $|f(x, y) - f(\bar{x}), f(\bar{y})| \leq 1$. En particulier f est bornée sur la boule $B((\bar{x}, \bar{y}), \varepsilon_1)$. Or pour n assez grand on a à la fois $d\big((x_n, y_n), (\bar{x}, \bar{y})\big) < \frac{1}{2}\varepsilon_1$ et $\operatorname{diam}(R_n) < \frac{1}{2}\varepsilon_1$. Mais par inégalité triangulaire ceci implique que $R_n \subset B((\bar{x}, \bar{y}), \varepsilon_1)$. Il y a alors contradiction avec le fait que sur R_n, la fonction f n'est pas bornée.

Montrons maintenant que la borne supérieure est atteinte (le raisonnement est bien entendu identique pour la borne inférieure).

Désormais nous savons que f est bornée, donc majorée sur le rectangle $[a; b] \times [c; d]$, posons $M = \sup_{[a;b]\times[c;d]}(f)$. Nous observons alors la propriété suivante : si une fonction g est majorée sur un rectangle R, alors :

(1) g est majorée sur chacun des sous-rectangles $R^{--}, R^{+-}, R^{-+}, R^{++}$

(2) $\sup_R(g) \geq \sup_{R^{--}}(g)$, $\sup_R(g) \geq \sup_{R^{+-}}(g)$, $\sup_R(g) \geq \sup_{R^{-+}}(g)$, $\sup_R(g) \geq \sup_{R^{++}}(g)$

(3) l'une au moins des inégalités précédentes est une égalité : g a le même sup sur R que sur l'un des quatre sous-rectangles.

Nous pouvons commencer à construire une suite de rectangles emboîtés comme suit. Posons $R_0 = [a; b] \times [c; d]$. Puis parmi les quatre sous-rectangles $R_0^{--}, R_0^{+-}, R_0^{-+}, R_0^{++}$, il y en a au moins un sur lequel le sup de f est M : notons R_1 ce rectangle. Ensuite considérons les quatre sous-rectangles $R_1^{--}, R_1^{+-}, R_1^{-+}, R_1^{++}$, il y en a au moins un sur lequel le sup de f est $\sup_{R_1}(f)$, soit M : notons R_2 ce rectangle. Par récurrence nous construisons ainsi une suite décroissante de rectangles $R_0 \supset R_1 \supset R_2 \supset \cdots \supset R_n \supset$ telle que

(1) $R_0 = [a; b] \times [c; d]$

(2) $\operatorname{diam}(R_n) = \frac{1}{2}\operatorname{diam}([a; b] \times [c; d])$, donc $\operatorname{diam}(R_n) \to 0$

(3) $\sup_{R_n}(f) = M$

Soit (à nouveau) (x_n, y_n) le coin inférieur gauche du rectangle R_n. Comme $R_{n+1} \subset R_n$ les deux suites $(x_n)_n$ et $(y_n)_n$ sont croissantes. Ce sont d'autre part des suites bornées ($x_n \leq b$ et $y_n \leq d$). Donc ces suites sont convergentes : soit $\bar{x} = \lim x_n$ et $\bar{y} = \lim y_n$. Nous allons démontrer que $f(\bar{x}, \bar{y}) = M$.

Puisque $\sup_{R_n}(f) = M$ il existe $(x_n', y_n') \in R_n$ tel que $f(x_n', y_n') \geq M - \frac{1}{1+n}$. Comme de toute façon $f(x_n', y_n') \leq M$ nous voyons que $f(x_n', y_n') \to M$. Or

$$d\big((x_n', y_n'), (\bar{x}, \bar{y})\big) \leq d\big((x_n', y_n'), (x_n, y_n)\big) + d\big((x_n, y_n), (\bar{x}, \bar{y})\big) \leq \operatorname{diam}(R_n) + d\big((x_n, y_n), (\bar{x}, \bar{y})\big)$$

Donc $(x_n', y_n') \to (\bar{x}, \bar{y})$. Par continuité de f nous obtenons finalement $f(\bar{x}, \bar{y}) = \lim_{n\to\infty} f(x_n', y_n')$, donc $f(\bar{x}, \bar{y}) = M$. $\square$

L'argument précédent s'adapte à des cadres plus généraux. Par exemple on peut le généraliser sans difficulté à un énoncé sur $\mathbb{R}^m$ (utiliser des pavés $[a_1; b_1] \times \cdots [a_m; b_m]$ au lieu de rectangles). Et on peut aussi obtenir le même énoncé avec $f : F \to \mathbb{R}$ où $F \subset \mathbb{R}^2$ est une partie fermée et bornée, pas forcément un rectangle (inclure dès le départ la partie bornée dans un rectangle, raisonner à chaque fois sur l'intersection de F avec les subdivisions du rectangle de départ).

Quatrième partie 4. **Fonctions de deux variables à valeur réelle : calcul différentiel.**

1. Etude d'une fonction de deux variables le long des droites parallèles aux axes.

1.1. **Applications partielles.** Pour étudier les fonctions de deux variables un outil fondamental est une autre sorte de composition que $\phi \circ g$, qui permet de se ramener à l'étude de fonctions d'une seule variable.

Définition 1.1 (courbes paramétrées de $\mathbb{R}^2$)**.** Une fonction $C : D \subset \mathbb{R} \to \mathbb{R}^2$ revient à la donnée de ses deux composantes, soit $C(t) = \big(x(t), y(t)\big)$, où chacune des deux fonctions $t \mapsto x(t)$ et $t \mapsto y(t)$ est définie sur D, à valeurs réelles. De manière imagée, on pense à $C(t)$ comme à un **point mobile** dans le plan $\mathbb{R}^2$, et on appelle *courbe paramétrée* une telle fonction C.

La fonction C est dite *continue (dérivable, de classe $\mathcal{C}^1$...)* si chacune des deux fonctions $t \mapsto x(t)$ et $t \mapsto y(t)$ est continue (dérivable, de classe $\mathcal{C}^1$...).

Lemme 1.2. *On fixe une norme sur $\mathbb{R}^2$ et on note d la distance associée. Alors on a l'équivalence entre les trois propriétés suivantes :*

(1) $C : D \subset \mathbb{R} \to \mathbb{R}^2$ est continue.

(2) Pour toute suite $(t_n)_n$ d'éléments de D telle que $t_n \to t$ on a $C(t_n) \to C(t)$.

(3) Pour tout $t \in D$ et pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel si un réel $s \in C$ vérifie $|s-t| < \alpha$ alors $d(C(s), C(t)) < \varepsilon$.

Démonstration. Posons $C(t) = \big(x(t), y(t)\big)$. Alors $C : D \subset \mathbb{R} \to \mathbb{R}^2$ est continue $\iff$ (1_x) : $t \mapsto x(t)$ est continue et (1_y) : $t \mapsto y(t)$ est continue. On a $C(t_n) \to C(t) \iff x(t_n) \to x(t)$ et $y(t_n) \to y(t)$: donc (2) est vraie $\iff$ (1_x) et (1_y) sont vraies. De même par équivalence de la norme avec $\|\cdot\|_\infty$, il existe $K > 0$ tel que : d'une part $d(C(s), C(t)) < \varepsilon \Rightarrow \big(|x(s) - x(t)| < K\varepsilon$ et $|y(s) - y(t)| < K\varepsilon\big)$ (donc (3) $\Rightarrow$ (1_x) et (1_y)), et d'autre part $\big(|x(s) - x(t)| < \varepsilon$ et $|y(s) - y(t)| < \varepsilon\big)$ $\Rightarrow d(C(s), C(t)) < K\varepsilon$ (donc (3) $\Leftarrow$ (1_x) et (1_y)). □

Etudier une fonction de deux variables $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ *le long d'une courbe paramétrée* $C : D \subset \mathbb{R} \to \mathbb{R}^2$, c'est par définition *étudier la composée* $t \mapsto f(C(t))$. L'intérêt principal, c'est que la composée $t \mapsto f(C(t))$ est une fonction d'une variable (pour laquelle nous disposons de plein d'outils d'analyse !).

Proposition 1.3 (composée $\mathbb{R} \to \mathbb{R}^2 \to \mathbb{R}$)**.** *Soit $C : D \subset \mathbb{R} \to \mathbb{R}^2$ une fonction d'une variable à valeur dans $\mathbb{R}^2$. Soit $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables avec $C(D) \subset X$.*

Si C et f sont continues, alors $f \circ C$ est continue.

Démonstration. Si $t_n \to t$ dans $\mathbb{R}$, alors, C étant continue, $C(t_n) \to C(t)$ dans $\mathbb{R}^2$. Puis f étant continue, $f(C(t_n)) \to f(C(t))$ dans $\mathbb{R}$. Ainsi $f \circ C$ est (séquentiellement) continue. □

Théorème 1.4 (théorème des valeurs intermédiaires sur $\mathbb{R}^2$)**.** *Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue. Alors $f(\mathbb{R}^2)$ est un intervalle de $\mathbb{R}$. Et pour tout rectangle $[a;b] \times [c;d]$ l'image directe $f([a;b] \times [c;d])$ est un intervalle fermé borné.*

Démonstration. Pour prouver que $f(\mathbb{R}^2)$ est un intervalle il suffit de prouver vérifier la caractérisation "stable par sous-segments". Soient donc $t, t' \in f(\mathbb{R}^2)$ avec $t \leq t'$: nous voulons montrer que $[t;t'] \subset f(\mathbb{R}^2)$. D'abord il existe $u, u' \in \mathbb{R}^2$ tels que $t = f(u), t' = f(u')$. Nous pouvons ensuite considérer l'application continue $C : [0;1] \to \mathbb{R}$ définie par $C(s) = su' + (1-s)u$. Alors C est continue,

$C(0) = u$ et $C(1) = u'$. Par composition, $\phi = f \circ C$ est continue sur $[0;1]$, avec $\phi(0) = f(u) = t$ et $\phi(1) = f(u') = t'$. Par le TVI appliqué à ϕ, nous savons que $[t;t'] \subset \phi([0;1]) = f\big(C([0;1])\big)$, et donc a fortiori $[t;t'] \subset f(\mathbb{R}^2)$.

C'est exactement le même argument qui démontre que pour $t, t' \in f([a;b] \times [c;d])$ avec $t \leq t'$, on a $[t;t'] \subset f([a;b] \times [c;d])$. Donc $f([a;b] \times [c;d])$ est un intervalle. De plus d'après le théorème de Weierstrass nous savons que $f([a;b] \times [c;d])$ est borné, possède un maximum et un minimum. Donc $f([a;b] \times [c;d])$ est un intervalle fermé borné. $\square$

Dans ce cours nous nous contentons d'étudier les fonctions de deux variables le long de droites paramétrées :

Définition 1.5 (droite paramétrée)**.** Soit $A = (a, b) \in \mathbb{R}^2$ un point fixé. Soit $u = (\alpha, \beta) \in \mathbb{R}^2$ un vecteur fixé, $u \neq (0, 0)$.

La *droite paramétrée passant par A et dirigée par u* est l'application $U : \mathbb{R} \to \mathbb{R}^2$ définie par $U(t) = (a + t\alpha, b + t\beta)$.

On remarque que $U(0) = A$, et que l'image directe $U(\mathbb{R})$ est bien la droite affine passant par A, de vecteur directeur u. Une équation cartésienne de cette droite est donnée par $(-\beta)(x-a)+\alpha(y-b) = 0$.

Par exemple pour $A = (2, 1)$ et $u = (1, -1)$ on obtient $U(t) = (2 - t, 1 + t)$. Lorsque t varie dans $\mathbb{R}$ le point $U(t)$ décrit la droite d'équation $x + y = 3$ dans $\mathbb{R}^2$.

(Si $u = (0, 0)$ l'application $U(t) = (a + t\alpha, b + t\beta)$ est bien définie mais elle est constante : le point $U(t)$ reste à la position A, donc on ne peut pas parler de *droite* paramétrée...)

Définition 1.6 (applications partielles)**.**

Soit $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables définie sur une partie X qui est voisinage d'un point $A = (a, b) \in \mathbb{R}^2$.

La fonction $x \mapsto f(x, b)$ est une fonction numérique de la variable x, appelée *application partielle de f par rapport à x en A (ou en $y = b$)* (on parle aussi d'application partielle *rapport à la première variable*). Elle est définie au voisinage de $x = a$.

La fonction $y \mapsto f(a, y)$ est une fonction numérique de la variable y, appelée *application partielle de f par rapport à y en A (ou en $x = a$)* (on parle aussi d'application partielle *rapport à la deuxième variable*). Elle est définie au voisinage de $y = b$.

Il faut noter que l'application $x \mapsto f(x, b)$ est la composée $f \circ U$, où $U : x \mapsto (x, b)$ est la droite paramétrée de $\mathbb{R}^2$ qui passe par A en $x = a$ et qui est dirigée par le vecteur $(1, 0)$ (l'image $U(\mathbb{R})$ est donc la droite horizontale qui passe par A). Donc étudier une application partielle de f par rapport à la première variable, c'est étudier f le long d'une droite paramétrée horizontale. De même : étudier une application partielle de f par rapport à la deuxième variable, c'est étudier f le long d'une droite paramétrée verticale.

Il y a une infinité d'application partielles de f par rapport à x : car il y en a une pour chaque choix de b. Idem pour les applications partielles par rapport à y : une pour chaque valeur de a.

L'application partielle $x \mapsto f(x, b)$ est définie en x précisément si (x, b) est dans le domaine de définition X de f. De même l'application partielle $y \mapsto f(a, y)$ est définie en y précisément si (a, y) est dans le domaine de définition X de f.

Par exemple pour $f(x, y) = x \ln(1 + xy^2)$ (étudiée sur l'ouvert $O = \{(x, y) \in \mathbb{R}^2, xy^2 > -1\}$), les applications partielles sont :

(1) Pour $b \in \mathbb{R}$ fixé : $x \mapsto x \ln(1 + b^2 x)$, le domaine de définition est $]-\frac{1}{b^2}; +\infty[$ si $b \neq 0$, $\mathbb{R}$ si $b = 0$;

(2) Pour $a \in \mathbb{R}$ fixé : $y \mapsto a \ln(1+ay^2)$, le domaine de définition est $\mathbb{R}$ si $a \geq 0$, et $]-\sqrt{-\frac{1}{a}}; +\sqrt{-\frac{1}{a}}[$ si $a < 0$.

Chaque fonction de deux variables $f(x, y)$ donne ainsi naissance à deux familles de fonctions d'une seule variable, ses applications partielles. En théorie : connaître la fonction de deux variables $(x, y) \mapsto f(x, y)$ revient exactement à connaître toutes ses applications partielles. Par exemple la surface représentative de la fonction de deux variables est constituée de la famille des courbes représentatives des applications partielles.

Pour étudier une fonction de deux variables il est donc légitime de commencer par étudier ses applications partielles, qui sont des fonctions d'une seule variable. Or pour étudier une fonction d'une variable (usuelle) on s'empresse de la dériver !

1.2. **Dérivées partielles et gradient.** Nous pouvons maintenant introduire une notion élémentaire de dérivation d'une fonction de deux variables :

Définition 1.7 (dérivées partielles et gradient ponctuel)**.**

Soit $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables définie sur une partie X qui est voisinage d'un point $A = (a, b) \in \mathbb{R}^2$.

A ficher

L'application partielle $x \mapsto f(x, b)$ est une fonction numérique de la variable x définie au voisinage de $x = a$. Si $x \mapsto f(x, b)$ est dérivable en $x = a$, on note ce nombre dérivé $\frac{\partial f}{\partial x}(a, b)$ (ou $\frac{\partial f}{\partial x}(A)$), et on l'appelle la *dérivée partielle de f par rapport à x en A*.

L'application partielle $y \mapsto f(a, y)$ est une fonction numérique de la variable y définie au voisinage de $y = b$. Si $y \mapsto f(a, y)$ est dérivable en $y = b$, on note ce nombre dérivé $\frac{\partial f}{\partial y}(a, b)$ (ou $\frac{\partial f}{\partial y}(A)$), et on l'appelle la *dérivée partielle de f par rapport à y en A*.

Nous dirons que f est *partiellement dérivable en A* lorsque f admet en A une dérivée partielle par rapport à x et par rapport à y. On regroupe alors ces deux dérivées partielles en un vecteur, appelé le *gradient* de f en A, que l'on note

$$\nabla_A f = (\frac{\partial f}{\partial x}(A), \frac{\partial f}{\partial y}(A)).$$

Il faut noter qu'une "dérivée partielle" est juste une dérivée familière : la dérivée d'une fonction d'une seule variable (x ou y selon le cas). Il y a deux calculs de dérivées, donc chacune de ces dérivées n'est qu'un renseignement partiel de dérivation.

Le qualificatif *partiel* se comprend encore mieux si on pense qu'au lieu de dériver $f(x, y)$ le long de la droite horizontale ou de la droite verticale passant par A, on aurait pu chercher à dériver $f(x, y)$ le long d'une droite quelconque passant par A, et dirigée par un vecteur unitaire $u_0 = (\cos(\theta_0), \sin(\theta_0))$. Autrement dit on pourrait dériver l'application d'une variable $t \mapsto f(a+t\cos(\theta_0), b+t\sin(\theta_0))$ (quand la dérivée existe elle est notée $\frac{\partial f}{\partial u_0}(A)$). Il y a a priori une infinité de telles dérivées directionnelles, puisque le vecteur u_0 peut prendre une infinité de directions ! Les dérivées $\frac{\partial f}{\partial x}(A)$, $\frac{\partial f}{\partial y}(A)$ correspondent respectivement aux cas $\theta_0 = 0$ et $\theta_0 = \frac{\pi}{2}$.

En fait dans les bons cas (lorsque f est *différentiable*) les deux dérivées partielles $\dfrac{\partial f}{\partial x}(A), \dfrac{\partial f}{\partial y}(A)$ sont suffisantes pour calculer toutes les dérivées de f le long de n'importe quelle courbe. En particulier on peut montrer que dans ce cas :

$$\frac{\partial f}{\partial u_0}(A) = \cos(\theta_0)\frac{\partial f}{\partial x}(A) + \sin(\theta_0)\frac{\partial f}{\partial y}(A)$$

Exemple 1.8. Reprenons la fonction $f(x,y) = x\ln(1+xy^2)$. Soit $A = (a,b)$ dans l'ouvert de définition $O = \{(x,y) \in \mathbb{R}^2, xy^2 > -1\}$. Pour y fixé égal à b, l'application partielle par rapport à x est $x \mapsto x\ln(1+b^2x)$, qui est de la forme $x\ln(u(x))$, avec $u(x)$ dérivable. Donc cette application partielle est dérivable sur son domaine de définition, de dérivée

$$\frac{d}{dx}[x \mapsto x\ln(1+b^2x)] = \ln(1+b^2x) + x\frac{b^2}{1+b^2x}.$$

Ainsi $\dfrac{\partial f}{\partial x}(a,b) = \ln(1+ab^2) + \dfrac{ab^2}{1+ab^2}$. Et pour x fixé égal à a, l'application partielle par rapport à y est $y \mapsto a\ln(1+ay^2)$, qui est de la forme $a\ln(v(y))$, avec $v(y)$ dérivable. Donc cette application partielle est dérivable sur son domaine de définition, de dérivée

$$\frac{d}{dy}[y \mapsto a\ln(1+ay^2)] = a\frac{2ay}{1+ay^2}.$$

Ainsi $\dfrac{\partial f}{\partial y}(a,b) = \dfrac{2a^2b}{1+ab^2}$.

Très vite, quand on est habitué, on parvient à calculer $\dfrac{\partial f}{\partial x}$ en considérant que dans la formule de $f(x,y)$, le nombre y ne varie pas, il n'y a plus qu'une variable (x). Il faut un temps de pratique pour arriver à bloquer (geler) mentalement la variable y, et laisser courir x seulement. Tous les moyens sont bons pour se souvenir de dériver par rapport à x, et pas par rapport à y ! Par exemple on pourrait imaginer une couleur réservée à la variable active, une autre pour la variable gelée :

$$\frac{\partial f}{\partial x} = \frac{d}{dx}[x\ln(1+xy^2)] = (x)'\ln(1+xy^2)+x\big(\ln(1+xy^2)\big)' = \ln(1+xy^2)+x\frac{(1+xy^2)'}{1+xy^2} = \ln(1+xy^2)+x\frac{y^2}{1+xy^2}$$

Quand on veut calculer $\dfrac{\partial f}{\partial y}$, on imagine x bloqué, et seule y libre de varier. Si on est perdu, on peut toujours revenir à l'artifice de notation $x \mapsto f(x,b)$ avec lequel il est clair que y est bloqué en b, et x est la variable.

Définition 1.9 (fonctions dérivées partielles, champ de gradient)**.**

Soit $f : O \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables définie sur un ouvert O.

Supposons que f admet des dérivées partielles en tout $(x,y) \in O$ dans les deux directions : nous dirons alors que f *est partiellement dérivable sur* O. Nous obtenons alors deux nouvelles fonctions de deux variables sur O :

$$\frac{\partial f}{\partial x} : (x,y) \mapsto \frac{\partial f}{\partial x}(x,y) \text{ et } \frac{\partial f}{\partial y} : (x,y) \mapsto \frac{\partial f}{\partial y}(x,y).$$

On obtient aussi, en considérant le gradient en tout point de O, une fonction $\nabla f : O \to \mathbb{R}^2$ (gradient de f) qui à (x,y) associe $\nabla_{(x,y)} f = (\dfrac{\partial f}{\partial x}(x,y), \dfrac{\partial f}{\partial y}(x,y))$. Le gradient est appelé un *champ de vecteurs* : si on trace en tout point A de O le vecteur gradient $\nabla_A f$ en mettant l'origine du gradient au point A, cela évoque des tiges dans un champs...

Exemple 1.10. Considérons $f(x,y) = 1 + 5xy^2 + \dfrac{1}{1+3x^2+y^4}$ définie sur $\mathbb{R}^2$.

Considérons l'application $x \mapsto f(x,y) = 1 + 5xy^2 + \dfrac{1}{1+3x^2+y^4}$ (x variable, y fixé). Alors $\dfrac{d}{dx}[x \mapsto 1+5xy^2+\dfrac{1}{1+3x^2+y^4}] = 5y^2 - 6x\dfrac{1}{(1+3x^2+y^4)^2}$, donc $\dfrac{\partial f}{\partial x}(x,y) = 5y^2 - 6x\dfrac{1}{(1+3x^2+y^4)^2}$.

De même $\dfrac{\partial f}{\partial y}(x,y) = \dfrac{d}{dy}[y \mapsto 1 + 5xy^2 + \dfrac{1}{1+3x^2+y^4}] = 10xy - 4y^3\dfrac{1}{(1+3x^2+y^4)^2}$.

Nous obtenons $\nabla_{(x,y)} f = \big(5y^2 - 6x\dfrac{1}{(1+3x^2+y^4)^2}; 10xy - 4y^3\dfrac{1}{(1+3x^2+y^4)^2}\big)$.

1.3. **Règles de calcul pour les dérivées partielles.**

Proposition 1.11 (dérivées partielles des sommes et des produits)**.**

Soit $O \subset \mathbb{R}^2$ un ouvert. Soient $f : O \to \mathbb{R}$ et $g : O \to \mathbb{R}$ deux fonctions numériques définies sur O. On suppose f et g partiellement dérivables sur O.

Alors les fonction $f+g$ et fg sont partiellement dérivables sur O et on a :

$$\frac{\partial (f+g)}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}\ ,\ \frac{\partial (f+g)}{\partial y} = \frac{\partial f}{\partial y} + \frac{\partial g}{\partial y}\ ,\quad \textit{soit } \nabla(f+g) = \nabla f + \nabla g$$

A ficher

$$\textit{et : } \frac{\partial (fg)}{\partial x} = g\frac{\partial f}{\partial x} + f\frac{\partial g}{\partial x}\ ,\ \frac{\partial (fg)}{\partial y} = g\frac{\partial f}{\partial y} + f\frac{\partial g}{\partial y}\ ,\quad \textit{soit } \nabla(fg) = g\nabla f + f\nabla g.$$

Démonstration. Pour $(x_0,y_0) \in O$ fixé, nous calculons $\dfrac{\partial f+g}{\partial x}(x_0,y_0)$. Il s'agit de dériver l'application partielle de $f+g$ par rapport à la première variable en $x = x_0$. Or cette application partielle est $\phi : x \mapsto (f+g)(x,y_0)$, donc $\phi(x) = f(x,y_0) + g(x,y_0)$ et il suffit d'appliquer la formule de dérivation d'une somme pour obtenir $\phi'(x_0) = \dfrac{\partial f}{\partial x}(x_0,y_0) + \dfrac{\partial g}{\partial x}(x_0,y_0)$. Soit $\dfrac{\partial (f+g)}{\partial x}(x_0,y_0) = \dfrac{\partial f}{\partial x}(x_0,y_0) + \dfrac{\partial g}{\partial x}(x_0,y_0)$. Ce pour tout $(x_0,y_0) \in O$, donc $\dfrac{\partial (f+g)}{\partial x} = \dfrac{\partial f}{\partial x} + \dfrac{\partial g}{\partial x}$. De la même façon $\dfrac{\partial (f+g)}{\partial y} = \dfrac{\partial f}{\partial y} + \dfrac{\partial g}{\partial y}$ et donc

$$\nabla(f+g) = (\frac{\partial (f+g)}{\partial x}; \frac{\partial (f+g)}{\partial y}) = (\frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}; \frac{\partial f}{\partial y} + \frac{\partial g}{\partial y}) = (\frac{\partial f}{\partial x}; \frac{\partial f}{\partial y}) + (\frac{\partial g}{\partial x}; \frac{\partial g}{\partial y}) = \nabla f + \nabla g.$$

Le cas du produit se traite de la même manière. Calculons $\dfrac{\partial (fg)}{\partial x}(x_0,y_0)$. Il s'agit de dériver l'application partielle de fg par rapport à la première variable en $x = x_0$. Or cette application partielle est $\phi : x \mapsto (fg)(x,y_0)$, donc $\phi(x) = f(x,y_0)g(x,y_0)$ et il suffit d'appliquer la formule de dérivation d'un produit (règle de Leibniz) pour obtenir $\phi'(x_0) = \dfrac{\partial f}{\partial x}(x_0,y_0) \times g(x_0,y_0) + f(x_0,y_0)\dfrac{\partial g}{\partial x}(x_0,y_0)$. Soit $\dfrac{\partial (fg)}{\partial x}(x_0,y_0) = g(x_0,y_0)\dfrac{\partial f}{\partial x}(x_0,y_0) + f(x_0,y_0)\dfrac{\partial g}{\partial x}(x_0,y_0)$. Ce pour tout $(x_0,y_0) \in O$, donc $\dfrac{\partial (fg)}{\partial x} = g\dfrac{\partial f}{\partial x} + f\dfrac{\partial g}{\partial x}$. On a la même formule pour la dérivée partielle par rapport à la deuxième variable, et on en déduit alors la formule pour $\nabla(fg)$. $\square$

Proposition 1.12 (dérivées partielles des composées $\mathbb{R}^2 \to \mathbb{R} \to \mathbb{R}$)**.** *Soit $O \subset \mathbb{R}^2$ un ouvert. Soit $f : O \to \mathbb{R}$ une fonction numérique définie sur O. On suppose f partiellement dérivable sur O. Soit $\phi : U \to \mathbb{R}$ une fonction définie sur un ouvert $U \subset \mathbb{R}$, avec $f(O) \subset U$, et ϕ est dérivable sur U.*

Alors la composée $(x, y) \mapsto \phi(f(x, y))$ (définie sur O) est partiellement dérivable sur O et on a

A faire

$$\frac{\partial(\phi \circ f)}{\partial x} = \frac{\partial f}{\partial x} \times (\phi' \circ f), \frac{\partial(\phi \circ f)}{\partial y} = \frac{\partial f}{\partial y} \times (\phi' \circ f) \textit{ soit } \nabla(\phi \circ f) = (\phi' \circ f) \times \nabla f$$

Démonstration. Par hypothèse, f est définie sur O, ϕ est définie sur U, et $f(O) \subset U$, donc l'application composée $(x, y) \mapsto \phi\big(f(x, y)\big)$ est bien définie pour tout $(x, y) \in O$.

Pour étudier les dérivées partielles de la fonction $\phi \circ f$, nous devons considérer ses applications partielles. L'application partielle par rapport à la première variable en un point $(x_0, y_0) \in O$ est $x \mapsto \phi\big(f(x, y_0)\big)$. (Elle est définie sur $J_{y_0} = \{x \in \mathbb{R} \text{ tels que } f(x, y_0) \in U\}$, ce qui forme un ouvert de $\mathbb{R}$ (contenant x_0). Le caractère ouvert de J_{y_0} vient de ce que $x \mapsto f(x, y_0)$ est continue comme composée de $f(x, y)$ continue avec la droite paramétrée $x \mapsto (x, y_0)$, et d'autre part J_{y_0} est l'image réciproque de l'ouvert U par la fonction continue $x \mapsto f(x, y_0)$.) Cette application partielle de $\phi \circ f$ est la composée de deux fonctions d'une variable : ϕ, et l'application partielle $x \mapsto f(x, y_0)$, donc elle est dérivable par dérivation des fonctions composées. De plus la formule pour sa dérivée en (x_0, y_0) est

$$\frac{d[\phi\big(f(x, y_0)\big)]}{dx}(x_0, y_0) = \frac{d[f(x, y_0)]}{dx}(x = x_0)\frac{d[\phi]}{dt}\big(t = f(x_0, y_0)\big) = \frac{\partial f}{\partial x}(x_0, y_0)\phi'\big(f(x_0, y_0)\big),$$

ce qui prouve la formule $\dfrac{\partial(\phi \circ f)}{\partial x} = \dfrac{\partial f}{\partial x} \times (\phi' \circ f)$. De même on a $\dfrac{\partial(\phi \circ f)}{\partial y} = \dfrac{\partial f}{\partial y} \times (\phi' \circ f)$, et la formule du gradient s'en déduit. □

Théorème 1.13 (régularité des fonctions usuelles (I))**.** *Soit $O \subset \mathbb{R}^2$ un ouvert. Soit $f : O \to \mathbb{R}$ une fonction* usuelle*, c'est à dire obtenue à partir d'un nombre fini d'opérations algébriques (sommes, produits) et de compositions avec des fonctions $\phi : I \to \mathbb{R}$ avec ϕ une fraction rationnelle, une fonction trigonométrique, l'exponentielle ou le logarithme. Alors :*

A faire

! important !

C^0 des fct usuelles et dérivées C^0

(1) f est continue sur O ;

(2) f est partiellement dérivable sur O ;

(3) De plus les dérivées partielles $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont encore des fonctions usuelles (et à ce titre sont continues).

Par exemple les polynômes sont des fonctions usuelles au sens du théorème : il s'obtiennent à l'aide de fonctions constantes ou des formes coordonnées par un nombre fini d'opération de sommes et de produits. Autre exemple de fonction usuelle :

$$f(x, y) = \frac{1 - x + 2y}{3 + x + 4y} - 5\ln(1 + x^2 + y^3) + \mathrm{e}^{x + y\cos(xy)}\sin(x + y^2).$$

Informellement, une fonction usuelle est une formule $f(x, y)$ où n'apparaissent que les variables x et y, des constantes réelles, et puis des fonctions usuelles d'une variable : $\sin(t), \cos(t), \dfrac{1}{t}, \ln(t), \mathrm{e}^t$. Attention dans cet énoncé $\sqrt{t} : \mathbb{R}^+ \to \mathbb{R}$ n'est pas considérée comme une fonction usuelle, mais la restriction de $\sqrt{t}$ à $\mathbb{R}^{+*}$ est usuelle car on a $\sqrt{t} = \mathrm{e}^{\frac{1}{2}\ln(x)}$ sur $\mathbb{R}^{+*}$.

Démonstration. Le théorème est vrai pour les fonctions de base que sont les formes affines.

Supposons que f_1, f_2 sont deux fonctions usuelles sur O. Supposons que nous savons déjà que f_1 et f_2 vérifient les conclusions du théorème (par exemple : f_1 et f_2 sont des formes affines). Considérons la somme $f = f_1 + f_2$. Alors f est une fonction usuelle. Elle est continue et partiellement dérivable sur O comme somme de deux fonctions continues et partiellement dérivables sur O. De plus $\dfrac{\partial f}{\partial x} = \dfrac{\partial f_1}{\partial x} + \dfrac{\partial f_2}{\partial x}$. Or nous savons déjà que $\dfrac{\partial f_1}{\partial x}$ et $\dfrac{\partial f_2}{\partial x}$ sont des fonctions usuelles. Alors leur somme $\dfrac{\partial f}{\partial x}$ est encore une fonction usuelle. Considérons maintenant le produit $g = f_1 f_2$. Alors g est une fonction usuelle. Elle est continue et partiellement dérivable sur O comme produit de deux fonctions continues et partiellement dérivables sur O. De plus $\dfrac{\partial g}{\partial x} = \dfrac{\partial f_1}{\partial x} f_2 + \dfrac{\partial f_2}{\partial x} f_1$. Or nous savons déjà que $\dfrac{\partial f_1}{\partial x}$ et $\dfrac{\partial f_2}{\partial x}$ sont des fonctions usuelles (aisi que f_1 et f_2 bien sûr). Donc $\dfrac{\partial g}{\partial x}$ est encore une fonction usuelle.

Enfin supposons que h est une fonction usuelle sur O, et que nous savons déjà que h vérifie les conclusions du théorème. Soit alors $\phi : U \to \mathbb{R}$ une fonction de base sur un ouvert U : fraction rationnelle, cos, sin, ln ou exp. Et supposons que $h(O) \subset U$, de sorte que $f = \phi \circ h$ est définie sur O. Alors f est continue et partiellement dérivable sur O, et de plus $\nabla f = \phi' \circ h \times \nabla h$. Or $\phi' : \mathbb{R} \to \mathbb{R}$ est encore une fonction de base, donc $\phi' \circ h$ est une fonction usuelle. Et par hypothèse $\dfrac{\partial h}{\partial x}$ et $\dfrac{\partial h}{\partial y}$ sont usuelles, donc en fait pour finir $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont usuelles. $\square$

Définition 1.14 (fonctions de classe $\mathcal{C}^0, \mathcal{C}^1, \mathcal{C}^k, \mathcal{C}^\infty$)**.**

Soit $O \subset \mathbb{R}^2$ un ouvert. Soit $f : O \to \mathbb{R}$ une fonction.

(1) On dit que f est de classe $\mathcal{C}^0$ sur O si f est continue sur O

(2) On dit que f est de classe $\mathcal{C}^1$ sur O si f est continue sur O, f est partiellement dérivable sur O, et les deux fonctions $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ sont continues sur O.

(3) Plus généralement, si $k \in \mathbb{N}$, on dit que f est de classe $\mathcal{C}^{k+1}$ sur O si f est continue et partiellement dérivable sur O, et si les deux fonctions $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ sont de classe $\mathcal{C}^k$ sur O.

(4) On dit que f est de classe $\mathcal{C}^\infty$ sur O si f est de classe $\mathcal{C}^k$ sur O pour tout entier $k \geq 0$.

Par exemple d'après le Théorème 1.13 les fonctions usuelles $f : O \to \mathbb{R}$ sont toutes de classe $\mathcal{C}^\infty$.

2. Points critique et principe de Fermat.

Définition 2.1 (point critique, point régulier)**.** Soit $O \subset \mathbb{R}^2$ un ouvert et soit $f : O \to \mathbb{R}$ une fonction partiellement dérivable en un point $A \in O$.

On dit que A est un *point critique de f* lorsque $\nabla_A f = (0,0)$, autrement dit lorsque $\dfrac{\partial f}{\partial x}(A) = \dfrac{\partial f}{\partial y}(A) = 0$. Lorsque f est partiellement dérivable sur O, on pourra noter $PC(f)$ l'ensemble des points critiques de f dans O.

On dit que A est un *point régulier de f* lorsqu'il n'est pas critique, c'est à dire $\nabla_A f \neq (0,0)$, soit encore lorsque $\dfrac{\partial f}{\partial x}(A) \neq 0$ ou $\dfrac{\partial f}{\partial y}(A) \neq 0$.

Exemple 2.2. Soit $f(x,y) = 2x^2 - 3xy + 4y^2 - x - 5y + 6$. Alors $\dfrac{\partial f}{\partial x}(x,y) = 4x - 3y - 1$ et $\dfrac{\partial f}{\partial y}(x,y) = -3x + 8y - 5$. Pour trouver les éventuels points critiques de f il faut trouver les (x,y) qui vérifient à la fois $\dfrac{\partial f}{\partial x}(x,y) = 0$, et $\dfrac{\partial f}{\partial y}(x,y) = 0$. C'est à dire ici : les (x,y) tels que $4x - 3y - 1 = 0$ et $-3x + 8y - 5 = 0$. Le premier lieu est ici une droite affine, ainsi que le deuxième lieu. Il existe un unique (x,y) tel que $4x - 3y - 1 = 0$ et $-3x + 8y - 5 = 0$ (c'est à dire à l'intersection des deux droites), c'est $(1,1)$. L'unique point critique de $f(x,y)$ est $(1,1)$, $PC(f) = \{(1,1)\}$.

On peut extraire de l'exemple ci-dessus un fait général : l'ensemble des points critiques de $f(x,y)$ s'obtient comme intersection des deux ensembles de niveau $\mathcal{V}$ et $\mathcal{H}$. Ici $\mathcal{V}$ est l'ensemble des zéros de la fonction $(x,y) \mapsto \dfrac{\partial f}{\partial x}(x,y)$ - c'est la courbe formée des points de O où le gradient est vertical puisque sa composante horizontale est nulle. Et $\mathcal{H}$ est l'ensemble des zéros de la fonction $(x,y) \mapsto \dfrac{\partial f}{\partial y}(x,y)$ - c'est la courbe formée des points de O où le gradient est horizontal puisque sa composante verticale est nulle.

Dans l'exemple chacune des deux courbes était particulièrement simple : une droite ! En général on a affaire à deux courbes sécantes en un certain nombre de points isolés.

Définition 2.3 (extremum global)**.** Soit $f : X \to \mathbb{R}$ une fonction définie sur une partie $X \subset \mathbb{R}^2$, et soit $A \in X$ un point. On dit que :

(1) *f présente un minimum global en A* si $\forall M \in X$ on a $f(M) \geq f(A)$ - donc $f(A)$ est la valeur minimale de f sur X entier

(2) *f présente un maximum global en A* si $\forall M \in X$ on a $f(M) \leq f(A)$ - donc $f(A)$ est la valeur maximale de f sur X entier

(3) *f présente un extremum global en A* si f présente en A un minimum ou un maximum global.

(4) *f présente un minimum global strict en A* si $\forall M \in X, M \neq A$ on a $f(M) > f(A)$

(5) *f présente un maximum global strict en A* si $\forall M \in X$ on a $f(M) < f(A)$

(6) *f présente un extremum global strict en A* si f présente en A un minimum ou un maximum global strict.

On remarque que si f présente un minimum strict en A, alors f présente un minimum en A (pour $M = A$ l'inégalité $f(M) \geq f(A)$ est évidente) En général, la valeur minimale peut être atteinte en deux points distincts A, A' de X ; mais pas si f présente un minimum strict en A.

Exemple 2.4. Reprenons l'exemple de $f(x,y) = 2x^2 - 3xy + 4y^2 - x - 5y + 6$. Commençons par transformer l'expression de $f(x,y)$. On a

$$2x^2 - 3xy + 4y^2 - x - 5y + 6 = 2(x^2 - \frac{1}{4}x) - 3xy + 4y^2 - 5y + 6 = 2(x^2 - \frac{3y+1}{2}x) + 4y^2 - 5y + 6 =$$

$$2(x - \frac{3y+1}{4})^2 - \frac{(3y+1)^2}{8} + 4y^2 - 5y + 6 = \frac{(4x-3y-1)^2}{8} + \frac{23}{8}y^2 - \frac{23}{4}y + \frac{47}{8}$$

et finalement $f(x,y) = \dfrac{(4x-3y-1)^2}{8} + \dfrac{23}{8}(y-1)^2 + 3$. Donc $f(x,y) \geq 3$, avec égalité si et seulement si les deux termes carrés sont nuls, c'est à dire $y = 1$, et $x = 1$. Ainsi f présente un minimum global strict en $(1,1)$.

D'après le théorème de Weierstrass, si $f : [a;b] \times [c;d] \to \mathbb{R}$ est une fonction continue, alors f présente sur le rectangle fermé $[a;b] \times [c;d]$ un maximum global et un minimum global.

Cependant nous devons en général étudier des fonctions $f : X \subset \mathbb{R}^2 \to \mathbb{R}$ définies sur des parties X non bornées. Par exemple, lorsque $f : \mathbb{R}^2 \to \mathbb{R}$ est une fonction continue sur $\mathbb{R}^2$, nous savons que $f(\mathbb{R}^2)$ est un intervalle de $\mathbb{R}$, et, souvent, on a en fait $f(\mathbb{R}^2) = \mathbb{R}$. Dans ce cas f ne présente d'extremum *global* en aucun point.

Mais il y a une version *locale* de la notion d'extremum, qui se rencontre même pour des fonctions sans extremum global :

Définition 2.5 (extremum local)**.** Soit $f : X \to \mathbb{R}$ une fonction définie sur une partie $X \subset \mathbb{R}^2$, et soit $A \in X$ un point. On suppose $\mathbb{R}^2$ muni d'une norme et de la distance associée. On dit que :

(1) *f présente un minimum local en A* si $\exists r > 0, \forall M \in X \cap B(A,r)$ on a $f(M) \geq f(A)$ - donc $f(A)$ est la valeur minimale de f sur $X \cap B(A,r)$

(2) *f présente un maximum local en A* si $\exists r > 0, \forall M \in X \cap B(A,r)$ on a $f(M) \leq f(A)$ - donc $f(A)$ est la valeur maximale de f sur $X \cap B(A,r)$

(3) *f présente un extremum local en A* si f présente en A un minimum ou un maximum local.

(4) *f présente un minimum local strict en A* si $\exists r > 0, \forall M \in X \cap B(A,r), M \neq A$ on a $f(M) > f(A)$

(5) *f présente un maximum local strict en A* si $\exists r > 0, \forall M \in X \cap B(A,r)$ on a $f(M) < f(A)$

(6) *f présente un extremum local strict en A* si f présente en A un minimum ou un maximum local strict.

Par exemple $f(x,y) = (x^2+y^2)\cos(x+y)$ présente un minimum local strict en $(0,0)$, mais aucun extremum global. En effet pour (x,y) suffisamment proche de $(0,0)$ on a $x+y$ proche de 0, donc $\cos(x+y) \geq \dfrac{1}{2}$, et $f(x,y) \geq \dfrac{1}{2}(x^2+y^2)$, en particulier $f(x,y) \geq 0$ (et même > 0 si $(x,y) \neq (0,0)$. En revanche le signe de $f(x,y)$ change complètement si (x,y) est loin de $(0,0)$. Par exemple si x est fixé et si on pose $y = \pi - x$, alors $x + y = \pi$, donc $\cos(x+y) = -1$ et donc $f(x,y) < 0$, ce qui prouve déjà que $f(0,0)$ n'est pas un minimum de $f(\mathbb{R}^2)$. En fait en considérant $f(x,x) = 2x^2\cos(2x)$, on voit que $f(\mathbb{R}^2) = \mathbb{R}$, donc f ne présente aucun extremum global.

Théorème 2.6 (Principe de Fermat)**.**

Soit O un ouvert et soit $f : O \to \mathbb{R}$ une fonction partiellement dérivable sur O.

Si $f(x,y)$ présente un extremum local en $A \in O$, alors A est un point critique de f.

La condition O ouvert est essentielle dans cet énoncé !

Démonstration. (Nous donnons l'argument dans le cas où l'extremum en A est un minimum.) Nous utiliserons la norme $\|\cdot\|_\infty$.

Ecrivons $A = (a,b)$. Puisque $A \in O$, avec O ouvert, il existe $r_1 > 0$ tel que $B_{d_\infty}(A,r_1) \subset O$. Puisque $f(x,y)$ présente un extremum local en A il existe $r_2 > 0$ tel que $\forall M \in B_{d_\infty}(A,r_2) \cap O$

on a $f(M) \geq f(A)$. Pour $r = \min(r_1, r_2)$ on a $r > 0$, $B_{d_\infty}(A, r) \subset O$ et $f(M) \geq f(A)$ pour tout $M \in B_{d_\infty}(A, r)$. Pour tout $x \in]a - r; a + r[$ on a $(x, b) \in B_{d_\infty}(A, r)$, donc $f(x, b) \geq f(a, b)$. La fonction d'une variable $x \mapsto f(x, b)$ est définie et dérivable sur $]a - r; a + r[$, et elle admet un minimum en $x = a$, donc sa dérivée s'annule en a. C'est dire que $\dfrac{\partial f}{\partial x}(a, b) = 0$. Par le même argument $\dfrac{\partial f}{\partial y}(a, b) = 0$. Donc (a, b) est un point critique de f. □

3. Approximation affine, différentielle.

3.1. **Brefs rappels sur les développements limités d'ordre 1 pour les fonctions numériques réelles.** Soit $f :]a; b[\to \mathbb{R}$ une fonction. Et soit $x_0 \in]a; b[$.

Alors, en x_0, f est dérivable *équivaut* à f admet un développement limité d'ordre 1 - c'est à dire qu'il existe des réels a, b et une fonction $\varepsilon :]a; b[\to \mathbb{R}$ tels que $f(x) = a + b(x - x_0) + (x - x_0)\varepsilon(x)$, avec $\lim\limits_{x \to x_0} \varepsilon(x) = 0$.

De plus, lorsqu'il existe une écriture $f(x) = a + b(x - x_0) + (x - x_0)\varepsilon(x)$ (avec $\lim\limits_{x \to x_0} \varepsilon(x) = 0$), alors cette écriture est unique : $a = f(x_0)$, $b = f'(x_0)$ et pour finir forcément on a la formule $\varepsilon(x) = \dfrac{f(x) - f(x_0)}{x - x_0} - f'(x_0)$ (pour $x \neq x_0$). La partie affine du développement limité est donc en fait égale à $x \mapsto f(x_0) + f'(x_0)(x - x_0)$, dont le graphe est la droite d'équation $y = f(x_0) + f'(x_0)(x - x_0)$, c'est à dire la tangente à la courbe représentative de f. On peut retrouver la tangente (et donc $f'(x_0)$) à partir de l'approximation affine !

Pour une fonction de deux variables f il n'est pas possible de considérer un taux d'accroissement $\dfrac{f(x, y) - f(x_0, y_0)}{(x, y) - (x_0, y_0)}$ (division d'un réel par un vecteur !). On va donc plutôt développer une théorie des développements limités d'ordre 1 généralisant celle connue en une variable : alors avoir un tel DL sera équivalent à la bonne généralisation de dérivable, qui s'appelle *différentiable* pour les fonctions de deux variables. Une partie essentielle de la formule de Taylor, c'est de contrôler le reste de manière très précise : la fameuse fonction $\varepsilon(x)$ tend vers 0 quand x tend vers x_0. Notre premier travail c'est donc de définir correctement des limites de fonctions $\varepsilon(x, y)$...

3.2. **Limites épointées, fonctions évanescentes (ou $o(1)$).**

Déjà dans la définition des dérivées d'une fonction d'une variable on était obligé de considérer des limites en un point où le taux d'accroissement n'est pas défini. En plusieurs variables, nous allons introduire les deux notions de limite qui interviennent, clarifier les relations entre ces deux limites, et faire le lien avec la continuité.

Définition 3.1 (limites et limites épointées)**.** On munit $\mathbb{R}^m$ de l'une de ses normes et de la distance d associée.

Soit $O \subset \mathbb{R}^m$ un ouvert et soit $v_0 \in O$ un point fixé. On note O' l'ensemble $O \setminus \{v_0\}$ (ouvert *épointé* - on a enlevé à l'ouvert O le point v_0 - noter que $O' = O \cap \big(\mathbb{R}^m \setminus \{v_0\}\big)$ est ouvert comme intersection de deux ouverts).

Soit f une fonction telle que $f(u)$ est défini pour tout $u \in O'$, à valeur réelle. Soit enfin $\ell \in \mathbb{R}$ un réel.

(1) Nous dirons que *la limite épointée de f en v_0 est ℓ* (noté $\lim\limits_{u \to v_0'} f(u) = \ell$) lorsque :

$$\forall \varepsilon > 0, \exists \alpha > 0 \text{ tel que si } u \in O' \text{ et } d(u, v_0) < \alpha, \text{ alors } |f(u) - \ell| < \varepsilon$$

(2) Supposons maintenant que $f(u)$ est défini pour tout $u \in O$. Nous dirons que *la limite de f en v_0 est ℓ* (noté $\lim\limits_{u \to v_0} f(u) = \ell$) lorsque :

$$\forall \varepsilon > 0, \exists \alpha > 0 \text{ tel que si } u \in O \text{ et } d(u, v_0) < \alpha, \text{ alors } |f(u) - \ell| < \varepsilon$$

Remarque 3.2 (indépendance par rapport à la norme choisie)**.** En utilisant l'équivalence des normes on voit que l'existence des limites dans la définition ci-dessus ne dépend pas de la norme considérée.

Remarque 3.3 (en dimension $m = 1$, lien avec les limites à gauche et à droite)**.** Soit $]a; b[$ un intervalle ouvert de $\mathbb{R}$, soit $x_0 \in]a; b[$ et soit $f :]a; b[\setminus \{x_0\} \subset \mathbb{R} \to \mathbb{R}$ une fonction. Alors

$$\lim_{x \to x_0'} f(x) = \ell \iff \lim_{x \to x_0^+} f(x) = \ell \text{ et } \lim_{x \to x_0^-} f(x) = \ell$$

Lemme 3.4 (unicité de ℓ)**.** *Les limites (ou les limites épointées), quand elles existent, sont uniques.*

Démonstration. Faisons-le pour les limites épointées (sinon remplacer O' par O dans l'argument !). Si $\lim\limits_{u \to v_0'} f(u) = \ell$ et $\lim\limits_{u \to v_0'} f(u) = \ell'$, fixons $\varepsilon > 0$: il existe alors $\alpha > 0$ et $\alpha' > 0$ tels que pour tout $u \in B(v_0, \alpha) \cap O'$ on a $|f(u) - \ell| < \frac{1}{2}\varepsilon$, et pour tout $u \in B(v_0, \alpha') \cap O'$ on a $|f(u) - \ell'| < \frac{1}{2}\varepsilon$. Comme O est voisinage de v_0, si on pose $\alpha_0 = \min(\alpha, \alpha')$, alors il y a une infinité de points u dans $B(v_0, \alpha_0) \cap O'$, et donc pour un tel u : $|\ell - \ell'| \le |\ell - f(u)| + |f(u) - \ell'| < \varepsilon$. Ainsi $\ell = \ell'$. $\square$

Lemme 3.5 (version séquentielle)**.** *Soit $O \subset \mathbb{R}^m$ un ouvert et soit $v_0 \in O$ un point fixé. Soit $f : O' \to \mathbb{R}$ une fonction.*

(1) On a $\lim\limits_{u \to v_0'} f(x, y) = \ell \iff$ pour toute suite $(u_n)_n$ de O' tendant vers v_0, la suite numérique $f(u_n)_n$ tend vers ℓ.

(2) Supposons maintenant que $f(u)$ est défini pour tout $u \in O$, y compris $u = v_0$. Alors $\lim\limits_{u \to v_0} f(x, y) = \ell \iff$ pour toute suite $(u_n)_n$ de O tendant vers v_0, la suite numérique $f(u_n)_n$ tend vers ℓ.

Une suite $(u_n)_n$ de O' est simplement une suite de O telle que pour tout $n \in \mathbb{N}$ on a $u_n \neq v_0$.

Démonstration. Faisons l'argument pour les limites épointées (sinon remplacer O' par O !). Supposons $\lim\limits_{u \to v_0'} f(u) = \ell$ et soit $(u_n)_n$ une suite de O' tendant vers v_0. Si $\varepsilon > 0$ est fixé il existe $\alpha > 0$ tel que pour tout $u \in B(v_0, \alpha) \cap O'$ on a $|f(u) - \ell| < \varepsilon$. Comme $u_n \to v_0$ il existe $N \in \mathbb{N}$ tel que pour $n \ge N$ on a $|u_n - v_0| < \alpha$, donc (comme on sait que $u_n \in O'$) on a $|f(u_n) - \ell| < \varepsilon$. On a montré que $f(u_n)_n$ tend vers ℓ.

Supposons maintenant que la limite épointée de f en v_0 n'existe pas ou n'est pas égale à ℓ : alors $\exists \varepsilon_0 > 0$ tel que pour tout $\alpha > 0$; il existe un point $u_\alpha \in O'$ tel que $|u_\alpha - v_0| < \alpha$ et pourtant $|f(u_\alpha) - \ell| \ge \varepsilon_0$. En prenant pour α les nombres $\frac{1}{1+n}$ on obtient pour chaque entier n un point $u_n \in O'$ tel que $|u_n - v_0| < \frac{1}{1+n}$ et pourtant $|f(u_n) - \ell| \ge \varepsilon_0$. Alors $(u_n)_n$ est une suite de O' qui tend vers v_0 et $(f(u_n))_n$ ne tend pas vers ℓ. $\square$

Corollaire 3.6 (limites et opérations). *Soit $O \subset \mathbb{R}^m$ un ouvert et soit $v_0 \in O$ un point fixé.*

(1) *Soient $f : O' \to \mathbb{R}$ et $g : O' \to \mathbb{R}$ deux fonctions telles que $\lim\limits_{u \to v_0'} f(u) = \ell$ et $\lim\limits_{u \to v_0'} g(u) = k$, alors $\lim\limits_{u \to v_0'} (f+g)(u) = \ell + k$ et $\lim\limits_{u \to v_0'} (fg)(u) = \ell k$.*

(2) *Sot $f : O' \to \mathbb{R}$ une fonction telle que $\lim\limits_{u \to v_0'} f(u) = \ell$ et soit $\phi : U \to \mathbb{R}$ une fonction définie sur un ouvert U contenant ℓ et $f(O')$, avec ϕ continue en ℓ. Alors $\phi \circ f$ est définie sur O' et $\lim\limits_{u \to v_0'} (\phi \circ f)(u) = \phi(\ell)$*

On a les mêmes propriétés si on considère des limites (non épointées) en v_0.

Lemme 3.7 (relations entre les différentes notions). *Soit $O \subset \mathbb{R}^m$ un ouvert et soit $v_0 \in O$ un point fixé. Soit $f : O \to \mathbb{R}$ une fonction.*

(1) *Si $\lim\limits_{u \to v_0} f(u) = \ell$ alors $\ell = f(v_0)$ et de plus $\lim\limits_{u \to v_0'} f(u) = \ell$ $(= f(v_0))$.*

(2) *$\lim\limits_{u \to v_0} f(u) = f(v_0) \iff f$ est continue en v_0.*

(3) *Si $\lim\limits_{u \to v_0'} f(u) = \ell$ avec $\ell \neq f(v_0)$ alors $f(u)$ n'a pas de limite quand $u \to v_0$. Si $\lim\limits_{u \to v_0'} f(u) = f(v_0)$ alors $\lim\limits_{u \to v_0} f(u) = f(v_0)$ et f est continue en v_0.*

Nous voyons que la limite usuelle n'est pas très innovante (elle correspond à la continuité), c'est en fait la limite épointée qui apporte des renseignements, mais elle est un peu plus délicate à manier à cause de la subtilité "$u \neq v_0$".

Démonstration. (1) Considérons la suite constante $u_n = v_0$: elle tend vers v_0, donc nécessairement $f(u_n)$ tend vers ℓ. D'où $\ell = f(v_0)$. La condition qui assure $\lim\limits_{u \to v_0} f(u) = \ell$ est plus forte que celle qui donne $\lim\limits_{u \to v_0'} f(u) = \ell$ $(= f(v_0))$, puisque dans la limite épointée il y a une condition supplémentaire $u \neq v_0$.

(2) Les deux définitions sont identiques (en prenant la définition $\forall \varepsilon \exists \alpha$ pour la continuité en x_0).

(3) La première affirmation découle de (1). Pour justifier la deuxième affirmation : si pour tout $u \in O'$ qui est α-proche de v_0, on a $f(u)$ à distance $< \varepsilon$ de $f(v_0)$, cela reste vrai pour $u = v_0$!

$\square$

Corollaire 3.8 (prolongement par continuité). *Soit $O \subset \mathbb{R}^m$ un ouvert et soit $v_0 \in O$ un point fixé. Soit $g : O' \to \mathbb{R}$ une fonction. Alors*

(1) *ou bien la limite épointée de $g(u)$ quand u tend vers v_0 n'existe pas et dans ce cas il n'existe aucune fonction continue $f : O \to \mathbb{R}$ telle que pour tout $u \in O, u \neq v_0$ on a $f(u) = g(u)$*

(2) *ou bien il existe un réel ℓ tel que $\lim\limits_{u \to v_0'} g(u) = \ell$ et dans ce cas il existe une unique fonction continue $f : O \to \mathbb{R}$ telle que pour tout $u \in O, u \neq v_0$ on a $f(u) = g(u)$. Cette fonction f est définie en v_0 par $f(v_0) = \ell$.*

Avec les limites épointées, on peut envisager une limite de $f(x,y)$ même quand $f(x_0,y_0)$ n'est pas défini - ou alors $f(x_0,y_0)$ est défini, mais on ne veut pas tenir compte de cette valeur. Par exemple on peut étudier $f(x,y) = \dfrac{\sin(xy)}{xy}$ (définie sur $O = \{(x,y) \in \mathbb{R}^2, xy \neq 0\}$). Elle n'est définie sur aucune des deux droites $x = 0$ ou $y = 0$, mais en utilisant $\lim_{t \to 0'} \frac{\sin(t)}{t} = 1$ on démontre que

$$\lim_{(x,y)\to(x_0,0)'} f(x,y) = \lim_{(x,y)\to(0,y_0)'} f(x,y) = \lim_{(x,y)\to(0,0)'} f(x,y) = 1.$$

Définition 3.9 (évanescent en (x_0, y_0))**.** Soit $O \subset \mathbb{R}^2$ un ouvert et soit $(x_0, y_0) \in O$ un point fixé. Soit $h : O \to \mathbb{R}$ une fonction : on dit que h est *évanescente en* (x_0, y_0) si $\lim_{(x,y)\to(x_0,y_0)'} h(x) = 0$. On écrit parfois $h \underset{(x_0,y_0)}{=} o(1)$ (lu : “h est petit o de 1 en (x_0, y_0)”).

On remarque que la propriété h est évanescente en (x_0, y_0) est indépendante de la valeur de $h(x_0, y_0)$.

Proposition 3.10. *Soit $O \subset \mathbb{R}^2$ un ouvert et soit $(x_0, y_0) \in O$ un point fixé. Soit $f : O \to \mathbb{R}$ une fonction. Alors f est continue en (x_0, y_0) $\iff$ la fonction $x \mapsto f(x, y) - f(x_0, y_0)$ est évanescente en (x_0, y_0).*

En particulier si $h : O \to \mathbb{R}$ est une fonction telle que $h(x_0, y_0) = 0$, alors il revient au même de demander que h soit évanescente en (x_0, y_0) ou que h soit continue en (x_0, y_0).

3.3. Approximation affine d'une fonction de deux variables, différentielle.

Définition 3.11. Soit O un ouvert de $\mathbb{R}^2$, soit $(x_0, y_0) \in O$ fixé, et soit $f : O \to \mathbb{R}$ une fonction. Soit d'autre part $A : (x, y) \mapsto C(x - x_0) + C'(y - y_0) + D$ une fonction affine. Les propriétés suivantes sont équivalentes et indépendante de la norme $\| \cdot \|$ considérée sur $\mathbb{R}^2$:

(1) Il existe une factorisation de $f(x, y) - A(x, y)$ sous la forme $f(x, y) - A(x, y) = \|(x - x_0, y - y_0)\|\varepsilon(x, y)$ (pour tout $(x, y) \in O$), où $\varepsilon : O \to \mathbb{R}$ est une fonction évanescente en (x_0, y_0).

(2) Il existe une factorisation de $f(x, y) - A(x, y)$ sous la forme $f(x, y) - A(x, y) = \|(x - x_0, y - y_0)\|\varepsilon(x)$ (pour tout $(x, y) \in O$), où $\varepsilon : O \to \mathbb{R}$ est une fonction continue en (x_0, y_0) avec $\varepsilon(x_0, y_0) = 0$.

(3) La fonction $h : O \to \mathbb{R}$ définie par $h(x, y) = \dfrac{f(x, y) - A(x, y)}{\|(x - x_0, y - y_0)\|}$ pour $(x, y) \in O, (x, y) \neq (x_0, y_0)$, et $h(x_0, y_0)$ arbitraire, est évanescente en (x_0, y_0) :

$$\lim_{(x,y)\to(x_0,y_0),(x,y)\neq(x_0,y_0)} \frac{f(x, y) - A(x, y)}{\|(x - x_0, y - y_0)\|} = 0.$$

Lorsque ces propriétés sont vérifiées, on dit que $A : (x, y) \mapsto C(x - x_0) + C'(y - y_0) + D$ est une *approximation affine de f en* (x_0, y_0).

L'intérêt de mettre une fonction affine $A(x, y)$ sous la forme $C(x - x_0) + C'(y - y_0) + D$ plutôt que $Cx + C'y + D'$ c'est que sous cette forme la constante D vaut en fait $D = A(x_0, y_0)$.

Exemple 3.12. (1) N'importe quelle application affine $f : (x, y) \mapsto ax + by + c$ admet une approximation affine en (x_0, y_0) quelconque : elle-même ! Si nous posons $A(x, y) = ax + by + c$, nous avons aussi $A(x, y) = a(x - x_0) + b(y - y_0) + c'$ (avec $c' = ax_0 + by_0 + c = A(x_0, y_0)$), et $f(x, y) - A(x, y) = \ldots 0$.

(2) Si $f(x, y) = 2 - 3x + 5y + x^2 + 4xy - y^3$ alors on peut voir l'approximation affine de $f(x, y)$ sur son écriture selon les puissances croissantes. On pose $A(x, y) = 2 - 3x + 5y$. Puis on contrôle chacun des termes de degré ≥ 2 comme suit :

(a) $|x^2| = |x||x| \leq \|(x, y)\|_\infty \varepsilon_1(x, y)$ avec $\varepsilon_1(x, y) = |x|$, évanescent en $(0, 0)$.

(b) $|4xy| = |x||4y| \leq \|(x,y)\|_\infty \varepsilon_2(x,y)$ avec $\varepsilon_2(x,y) = |4y|$, évanescent en $(0,0)$.

(c) $|-y^3| = |y||-y^2| \leq \|(x,y)\|_\infty \varepsilon_3(x,y)$ avec $\varepsilon_2(x,y) = y^2$, évanescent en $(0,0)$.

Pour $(x,y) \neq (0,0)$ on a alors

$$\frac{|f(x,y) - A(x,y)|}{\|(x,y)\|_\infty} = \frac{|x^2 + 4xy - y^3|}{\|(x,y)\|_\infty} \leq \varepsilon_1(x,y) + \varepsilon_2(x,y) + \varepsilon_3(x,y),$$

évanescent en $(0,0)$.

Proposition 3.13 (approximation affine $\Rightarrow$ continue et partiellement dérivable)**.**

Soit O un ouvert de $\mathbb{R}^2$, soit $(x_0,y_0) \in O$ fixé, et soit $f : O \to \mathbb{R}$ une fonction.

Supposons que f admet une approximation affine $A : (x,y) \mapsto C(x-x_0) + C'(y-y_0) + D$. Alors

(1) $D = f(x_0,y_0)$

(2) f est partiellement dérivable en (x_0,y_0) et $C = \frac{\partial f}{\partial x}(x_0,y_0), C' = \frac{\partial f}{\partial y}(x_0,y_0)$

(3) f est continue en (x_0,y_0)

En particulier l'approximation affine (si elle existe !) est unique, sa partie linéaire est

$$(h,k) \mapsto h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0),$$

cette forme linéaire est appelée la différentielle de f en (x_0,y_0)*, notée* $D_{(x_0,y_0)}f$*.*

Une fonction qui admet une approximation affine en un point (x_0,y_0) est dite différentiable en (x_0,y_0)*. Une fonction $f : O \subset \mathbb{R}^2 \to \mathbb{R}$ qui est différentiable en tout point de O est dite* différentiable (sur O)

On remarque que la différentielle est le produit scalaire par le gradient.

A la différence du cas de la dérivabilité sur $\mathbb{R}$, l'énoncé dans la proposition n'est qu'une implication. La réciproque est fausse, comme le montre la fonction $f(x,y) = \frac{xy}{x^2+y^2}$ pour $(x,y) \neq (0,0), f(0,0) = 0$: nous avons vu en TD que cette fonction n'est même pas continue en $(0,0)$, pourtant elle est partiellement dérivable sur $\mathbb{R}^2$ (même en $(0,0)$, où les dérivées partielles sont nulles).

Démonstration. Par hypothèse il existe une fonction $\varepsilon : O \to \mathbb{R}$ évanescente en (x_0,y_0) telle que

$$\forall (x,y) \in O, f(x,y) = D + C(x-x_0) + C'(y-y_0) + \|(x-x_0, y-y_0)\|\varepsilon(x,y)$$

En évaluant en $(x,y) = (x_0,y_0)$ nous obtenons $f(x_0,y_0) = D + 0 + 0 + 0$. En fixant $y = y_0$ et en laissant varier x (proche de x_0 pour que $(x,y_0) \in O$ et donc l'égalité est valable) nous obtenons :

$$f(x,y_0) = D + C(x-x_0) + 0 + \|(x-x_0,0)\|\varepsilon(x,y_0),$$

soit

$$f(x,y_0) = f(x_0,y_0) + C(x-x_0) + 0 + |x-x_0|\|(1,0)\|\varepsilon(x,y_0)$$

Alors pour $x \neq x_0$ on a $\frac{f(x,y_0) - f(x_0,y_0)}{x-x_0} = C + \varepsilon'(x)$ avec $\varepsilon'(x) = \pm\|(1,0)\|\varepsilon(x,y_0)$, qui tend vers 0 quand x tend vers 0. Nous en déduisons que l'application partielle $x \mapsto f(x,y_0)$ est dérivable en x_0, et que $\frac{\partial f}{\partial x}(x_0,y_0) = C$. Un argument analogue montre que l'application partielle $y \mapsto f(x_0,y)$

est dérivable en y_0, et que $\dfrac{\partial f}{\partial x}(x_0, y_0) = C'$. Pour finir démontrons que f est continue en (x_0, y_0). Nous avons :

$$|f(x,y) - f(x_0,y_0)| = |f(x,y) - D| \leq |C||x - x_0| + |C'||y - y_0| + \|(x - x_0, y - y_0)\||\varepsilon(x,y)|$$

D'une part il existe $K > 0$ tel que $|x - x_0| \leq K\|(x - x_0, y - y_0)\|$ et $|y - y_0| \leq K\|(x - x_0, y - y_0)\|$. D'autre part comme $\varepsilon(x,y) \to 0$ quand $(x,y) \to (x_0,y_0)$ il existe $M_1 \geq 0$ tel que $|\varepsilon(x,y)| \leq M_1$ pour tout (x,y) suffisamment proche de (x_0,y_0). En posant $M = \max(K|C|, K|C'|, M_1)$ nous obtenons la majoration suivante, valable pour tout (x,y) suffisamment proche de (x_0,y_0) :

$$|f(x,y) - f(x_0,y_0)| \leq M\|(x - x_0, y - y_0)\|$$

Cela assure visiblement la continuité de f en (x_0,y_0). □

On peut retenir que l'existence d'une approximation affine est une propriété très forte : le fait qu'une fonction de deux variables tende vers 0 quand $(x,y) \to (x_0,y_0)$ est beaucoup plus fort que les versions partielles de cette propriété. La différentiabilité est beaucoup plus forte que la dérivabilité partielle.

Définition 3.14 (développement de Taylor-Young en un point)**.** Soit O un ouvert de $\mathbb{R}^2$, soit $(x_0,y_0) \in O$ fixé, et soit $f : O \to \mathbb{R}$ une fonction différentiable en (x_0,y_0). Alors par définition on peut écrire :

$$f(x,y) = f(x_0,y_0) + (x - x_0)\frac{\partial f}{\partial x}(x_0,y_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0,y_0) + \|(x - x_0, y - y_0)\|\varepsilon(x,y) \text{ avec } \varepsilon(x,y) \underset{(x_0,y_0)}{=} o(1)$$

ce développement de $f(x,y)$ en une somme hiérarchisée de termes s'appelle le développement de Taylor-Young de $f(x,y)$ au point (x_0,y_0).

Définition 3.15 (plan tangent à la surface représentative)**.** Soit O un ouvert de $\mathbb{R}^2$, soit $(x_0,y_0) \in O$ fixé, et soit $f : O \to \mathbb{R}$ une fonction différentiable en (x_0,y_0). On appelle *plan tangent à la surface représentative de* f le plan $P \subset \mathbb{R}^3$ d'équation $z = f(x_0,y_0) + (x - x_0)\dfrac{\partial f}{\partial x}(x_0,y_0) + (y - y_0)\dfrac{\partial f}{\partial y}(x_0,y_0)$. On note que P est le plan passant par $(x_0, y_0, f(x_0,y_0))$, de vecteur normal $(\dfrac{\partial f}{\partial x}(x_0,y_0), \dfrac{\partial f}{\partial y}(x_0,y_0), -1)$.

Par exemple le plan tangent en un maximum ou un minimum (local) (x_0,y_0) est le plan horizontal d'équation $z = f(x_0,y_0)$ (c'est vrai plus généralement en tout point critique).

Théorème 3.16 (classe $\mathcal{C}^1$)**.** *Supposons que $O \subset \mathbb{R}$ est ouvert et que $f : O \to \mathbb{R}$ est une fonction partiellement dérivable telle que $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continues sur O. Alors f admet une approximation affine en tout point de O - f est différentiable sur O.*

Démonstration. Fixons un point $(x_0,y_0) \in O$. D'après la Proposition 3.13 nous devons montrer que la fonction affine $A(x,y) = f(x_0,y_0) + (x - x_0)\dfrac{\partial f}{\partial x}(x_0,y_0) + (y - y_0)\dfrac{\partial f}{\partial y}(x_0,y_0)$ est l'approximation affine de f en (x_0,y_0).

Pour cela nous devons prouver que $f(x,y) - A(x,y)$ est de la forme $\|(x - x_0, y - y_0)\|\varepsilon(x,y)$ avec $\varepsilon(x,y)$ évanescente en (x_0,y_0). A cet effet, nous allons estimer l'accroissement de f entre (x_0,y_0) et (x,y). Pour simplifier les expressions nous posons $h = x - x_0$ et $k = y - y_0$, de sorte que

$(x,y) = (x_0+h, y_0+k)$. Nous supposons que $|h|, |k| < r$ avec r tel que $B_{d_\infty}\big((x_0,y_0), r\big) \subset O$ (possible puisque O est ouvert). Alors :

$$f(x_0+h, y_0+k) - f(x_0,y_0) = \big(f(x_0+h,y_0+k) - f(x_0+h,y_0)\big) + \big(f(x_0+h,y_0) - f(x_0,y_0)\big)$$

Le deuxième terme $f(x_0+h,y_0)-f(x_0,y_0)$ est l'accroissement de l'application partielle $x \mapsto f(x,y_0)$ entre x_0 et x_0+h (bien définie puisque $(x_0+s, y_0) \in O$ pour tout s tel que $|s| < r$). Or f est partiellement dérivable, donc l'application partielle $x \mapsto f(x,y_0)$ est dérivable et nous pouvons lui appliquer le théorème des accroissements finis : il existe $c \in]x_0; x_0+h[$ (ou $c \in]x_0+h; x_0[$ si $h<0$, ou $c = x_0$ si $h=0$) tel que

$$f(x_0+h,y_0) - f(x_0,y_0) = (x_0+h-x_0)\frac{\partial f}{\partial x}(c,y_0), \text{ soit } f(x_0+h,y_0)-f(x_0,y_0) = h\frac{\partial f}{\partial x}(c,y_0)$$

Comme $\dfrac{\partial f}{\partial x}$ est continue en (x_0,y_0), la quantité $\varepsilon_1(h) = \dfrac{\partial f}{\partial x}(c,y_0) - \dfrac{\partial f}{\partial x}(x_0,y_0)$ tend vers 0 quand $h \to 0$. Nous avons déjà $f(x_0+h,y_0) - f(x_0,y_0) = h\dfrac{\partial f}{\partial x}(x_0,y_0) + h\varepsilon_1(h)$.

Le terme $f(x_0+h,y_0+k) - f(x_0+h,y_0)$ est l'accroissement de l'application partielle $y \mapsto f(x_0+h, y_0)$ entre y_0 et y_0+k (bien définie puisque $(x_0+h, y_0+t) \in O$ pour tout t tel que $|t| < r$). Cette application partielle est dérivable par hypothèse et nous lui appliquons le théorème des accroissements finis : il existe $d \in]y_0; y_0+k[$ (ou $d \in]y_0+k; y_0[$ si $h<0$, ou $d = y_0$ si $k=0$) tel que

$$f(x_0+h,y_0+k)-f(x_0+h,y_0) = (y_0+k-y_0)\frac{\partial f}{\partial y}(x_0+h,d), \text{ soit } f(x_0+h,y_0+k)-f(x_0+h,y_0) = k\frac{\partial f}{\partial x}(x_0+h,d)$$

Comme $\dfrac{\partial f}{\partial y}$ est continue en (x_0,y_0), la quantité $\varepsilon_2(h,k) = \dfrac{\partial f}{\partial y}(x_0+h,d) - \dfrac{\partial f}{\partial y}(x_0,y_0)$ tend vers 0 quand $(h,k) \to (0,0)$. Nous avons maintenant $f(x_0+h,y_0+k) - f(x_0+h,y_0) = k\dfrac{\partial f}{\partial y}(x_0,y_0) + k\varepsilon_2(h,k)$. Alors :

$$f(x_0+h,y_0+k) - f(x_0,y_0) = k\frac{\partial f}{\partial y}(x_0,y_0) + k\varepsilon_2(h,k) + h\frac{\partial f}{\partial x}(x_0,y_0) + h\varepsilon_1(h), \text{ soit}$$
$$f(x,y) - A(x,y) = h\varepsilon_1(h) + k\varepsilon_2(h,k)$$

On écrit $h\varepsilon_1(h) + k\varepsilon_2(h,k) = \|(h,k)\|_\infty\, \varepsilon(h,k)$ avec $\varepsilon(h,k) = \dfrac{h\varepsilon_1(h)+k\varepsilon_2(h,k)}{\|(h,k)\|_\infty}$ si $(h,k) \neq (0,0)$ et $\varepsilon(0,0) = 0$. Comme

$$|h\varepsilon_1(h) + k\varepsilon_2(h,k)| \leq |h||\varepsilon_1(h)| + |k||\varepsilon_2(h,k)| \leq \max(|h|,|k|)[|\varepsilon_1(h) + \varepsilon_2(h,k)|]$$

on a $|\varepsilon(h,k)| \leq |\varepsilon_1(h) + \varepsilon_2(h,k)|$, donc ε est évanescent en $(0,0)$. □

3.4. **Opérations sur les fonctions différentiables.**

Proposition 3.17 (opérations algébriques)**.**

Soient $f, g : O \subset \mathbb{R}^2 \to \mathbb{R}$ deux fonctions définies sur un ouvert $O \subset \mathbb{R}^2$. Soit $(x_0,y_0) \in O$ un point fixé. Supposons que f et g sont différentiables en (x_0,y_0).

(1) La somme $f+g$ est différentiable en (x_0,y_0) et $D_{(x_0,y_0)}(f+g) = D_{(x_0,y_0)}(f) + D_{(x_0,y_0)}(g)$.

(2) Le produit fg est différentiable en (x_0,y_0) et $D_{(x_0,y_0)}(fg) = g(x_0,y_0)D_{(x_0,y_0)}(f) + f(x_0,y_0)D_{(x_0,y_0)}(g)$.

En particulier la somme et le produit de deux fonctions différentiables sur l'ouvert O est une fonction différentiable sur O.

Proposition 3.18 (compositions)**.** *Soient f une fonction définies sur un ouvert $O \subset \mathbb{R}^2$. Soit $(x_0, y_0) \in O$ un point fixé. Supposons que f est différentiable en (x_0, y_0).*

(1) Pour toute fonction $\phi : I \subset\!\!\rightarrow \mathbb{R}$ telle que $f(O) \subset I$ et ϕ est dérivable en $t_0 = f(x_0, y_0)$, la composée $\phi \circ f$ est différentiable en (x_0, y_0) et $D_{(x_0,y_0)}(\phi \circ f) = \phi'(t_0) D_{(x_0,y_0)}(f)$.

(2) Pour toute courbe paramétrée $C : I \subset\!\!\rightarrow \mathbb{R}^2$, $C(t) = (x(t), y(t))$, telle que la courbe image $C(I)$ est contenue dans O, et pour tout $t_0 \in I$ tel que $C(t_0) = (x_0, y_0)$, avec C dérivable en t_0, la composée $h = f \circ C$ est dérivable en t_0 et de plus

$$h'(t_0) = x'(t_0)\frac{\partial f}{\partial x}(x_0, y_0) + y'(t_0)\frac{\partial f}{\partial y}(x_0, y_0), \textit{ ou encore } h'(t_0) = \nabla_{(x_0,y_0)} f \cdot \vec{V}(t_0)$$

(dans cette dernière formule $\vec{V}(t_0)$ désigne le vecteur $(x'(t_0), y'(t_0))$, vitesse *de la courbe C au temps t_0).*

Corollaire 3.19. *Toute fonction usuelle $f : O \subset \mathbb{R}^2 \to \mathbb{R}$ est différentiable sur O.*

Démonstration. Les formes affines sont différentiables. De plus être différentiable est une propriété stable par les opérations algébriques et compositions (Propositions 3.17 et 3.18). Donc les fonctions usuelles sont différentiables, puisqu'elles s'obtiennent par une suite finies d'opérations en partant des formes affines. $\square$

Démonstration de la stabilité de "différentiable" par rapport aux différentes opérations.

Soient f, g, ϕ, C comme dans les Propositions 3.17 et 3.18. On va obtenir la conclusion en utilisant dans chaque cas les développements de Taylor adaptés.

Puisque f est différentiable en (x_0, y_0) elle y admet par définition une approximation affine, c'est à dire qu'il existe une forme affine $(x, y) \mapsto A(x, y)$ et une fonction $\varepsilon_f(x, y)$ évanescente en (x_0, y_0) telles que

$$\forall (x, y) \in O, f(x, y) = A(x, y) + \|(x - x_0, y - y_0)\| \varepsilon_f(x, y)$$

De même il existe une forme affine $(x, y) \mapsto B(x, y)$ et une fonction $\varepsilon_g(x, y)$ évanescente en (x_0, y_0) telles que

$$\forall (x, y) \in O, g(x, y) = B(x, y) + \|(x - x_0, y - y_0)\| \varepsilon_g(x, y)$$

En ajoutant les deux développements de Taylor nous obtenons

$$\forall (x, y) \in O, f(x, y) + g(x, y) = A(x, y) + B(x, y) + \|(x - x_0, y - y_0)\| \varepsilon_f(x, y) + \|(x - x_0, y - y_0)\| \varepsilon_g(x, y)$$

L'application $S : (x, y) \mapsto A(x, y) + B(x, y)$ est la somme de deux formes affines : c'est bien une forme affine. Ensuite le "reste" se factorise

$$\|(x - x_0, y - y_0)\| \varepsilon_f(x, y) + \|(x - x_0, y - y_0)\| \varepsilon_g(x, y) = \|(x - x_0, y - y_0)\| (\varepsilon_f(x, y) + \varepsilon_g(x, y))$$

Or la fonction $\varepsilon : (x, y) \mapsto \varepsilon_f(x, y) + \varepsilon_g(x, y)$ est définie sur O, et de plus sa limite épointée en (x_0, y_0) vaut $0 + 0 = 0$, elle est bien évanescente en (x_0, y_0). Tout ceci prouve que

$$(f + g)(x, y) = S(x, y) + \|(x - x_0, y - y_0)\| \varepsilon(x, y), \text{ avec } S \text{ affine et } \varepsilon \text{ évanescente en } (x_0, y_0),$$

donc $f + g$ admet la fonction S comme approximation affine en (x_0, y_0). Donc $f + g$ est différentiable en (x_0, y_0), et de plus $D_{(x_0,y_0)}(f + g)$ est la partie linéaire de $A + B$. Mais en fait nous savons que $A(x, y) = f(x_0, y_0) + D_{(x_0,y_0)} f(x, y)$ et $B(x, y) = g(x_0, y_0) + D_{(x_0,y_0)} g(x, y)$, donc la partie linéaire de $A + B$ est bien $D_{(x_0,y_0)}(f) + D_{(x_0,y_0)}(g)$.

Si nous multiplions les deux développements de Taylor de f et g nous obtenons $(\forall (x, y) \in O)$:

$$(A \times B)(x, y) + \|(x - x_0, y - y_0)\| (B \times \varepsilon_f)(x, y) + \|(x - x_0, y - y_0)\| (A \times \varepsilon_g)(x, y) + \|(x - x_0, y - y_0)\|^2 (\varepsilon_f \times \varepsilon_g)(x, y)$$

Les termes après $A \times B$ seront intégrés au futur reste, mais traitons d'abord la partie $(A \times B)(x, y)$ (un "simple" polynôme de degré 2 !). Nous savons que

$$A(x, y) = f(x_0, y_0) + L_f(x - x_0, y - y_0) \text{ et } B(x, y) = g(x_0, y_0) + L_g(x - x_0, y - y_0)$$

avec $L_f(h, k) = D_{(x_0,y_0)} f(h, k) = h\dfrac{\partial f}{\partial x}(x_0, y_0) + k\dfrac{\partial f}{\partial y}(x_0, y_0)$ et $L_g(h, k) = D_{(x_0,y_0)} g(h, k) = h\dfrac{\partial g}{\partial x}(x_0, y_0) + k\dfrac{\partial g}{\partial y}(x_0, y_0)$. Donc on obtient :

$$(A \times B)(x, y) = f(x_0, y_0)g(x_0, y_0) + f(x_0, y_0)L_g(x - x_0, y - y_0) + g(x_0, y_0)L_f(x - x_0, y - y_0) + Q(x, y)$$

$$\text{avec } Q(x, y) = L_f(x - x_0, y - y_0) \times L_g(x - x_0, y - y_0)$$

Voyons comment contrôler $|Q(x, y)|$ lorsque $(x, y) \to (x_0, y_0)$. D'abord :

$$|L_f(h, k)| = |h\frac{\partial f}{\partial x}(x_0, y_0) + k\frac{\partial f}{\partial y}(x_0, y_0)| \leq \max(|h|, |k|)(|\frac{\partial f}{\partial x}(x_0, y_0)| + |\frac{\partial f}{\partial y}(x_0, y_0)|)$$

Alors

$$|Q(x, y)| = |L_f(h, k)L_g(h, k)| \leq \|(h, k)\|_\infty \alpha(x, y) \text{ avec } \alpha(x, y) = C_f C_g \|(h, k)\|_\infty$$

en posant $C_f = |\dfrac{\partial f}{\partial x}(x_0, y_0)| + |\dfrac{\partial f}{\partial y}(x_0, y_0)|$ et $C_g = |\dfrac{\partial g}{\partial x}(x_0, y_0)| + |\dfrac{\partial g}{\partial y}(x_0, y_0)|$. Nous en déduisons que pour $(x, y) \neq (x_0, y_0)$ on a

$$|\frac{Q(x, y)}{\|(x - x_0, y - y_0)\|_\infty}| \leq \alpha(x, y), \text{ donc } \lim_{(x,y)\to(x_0,y_0)'} \frac{Q(x, y)}{\|(x - x_0, y - y_0)\|_\infty} = 0$$

Pour finir nous obtenons un développement de la forme

$$f(x, y)g(x, y) = P(x, y) + [Q(x, y) + \|(h, k)\|(B \times \varepsilon_f)(x, y) + \|(h, k)\|(A \times \varepsilon_g)(x, y) + \|(h, k)\|^2(\varepsilon_f \times \varepsilon_g)(x, y)]$$

où $P(x, y) = f(x_0, y_0)g(x_0, y_0) + f(x_0, y_0)L_g(x - x_0, y - y_0) + g(x_0, y_0)L_f(x - x_0, y - y_0)$ est une forme affine, et le terme entre crochets, une fois divisé par $\|(h, k)\|$, tend vers 0 quand $(x, y) \to (x_0, y_0)'$. $\square$

4. Etude le long d'un segment quelconque, accroissements finis et tangente aux courbes de niveaux.

Théorème 4.1 (dérivée le long d'un segment)**.**

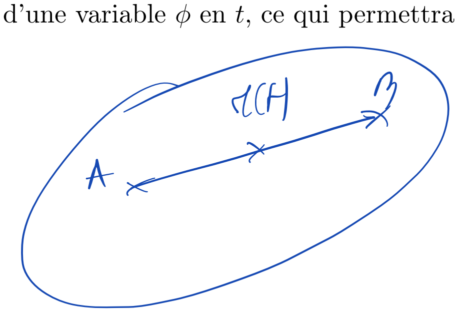
Soit $O \subset \mathbb{R}^2$ un ouvert et soit $f : O \to \mathbb{R}$ une fonction différentiable sur O.

Supposons que $A_0 = (x_0, y_0)$ et $A_1 = (x_1, y_1)$ sont deux points de O tels que le segment $[A_0A_1]$ est entièrement contenu dans O. Pour $t \in [0; 1]$ posons alors $x_t = tx_1 + (1 - t)x_0$ et $y_t = ty_1 + (1 - t)y_0$, ainsi le point $A_t = (x_t, y_t)$ est dans $[A_0A_1]$.

Considérons alors la fonction $\phi : [0; 1] \to \mathbb{R}$ définie par $\phi(t) = f(x_t, y_t)$. D'abord ϕ est dérivable sur $[0; 1]$, ensuite on a

$$\boxed{\forall t \in [0; 1], \ \phi'(t) = \langle \nabla_{A_t} f, \overrightarrow{A_0A_1} \rangle}$$

Démonstration. Soit $t \in [0; 1]$ fixé, on va obtenir un développement limité d'ordre 1 pour la fonction d'une variable ϕ en t, ce qui permettra de conclure.



$f : O \longrightarrow \mathbb{R}$

$g : t \longmapsto f(M(t))$

g'

On a $\phi(t) = f(A_t)$, et $A_t \in [A_0A_1] \subset O$, donc f est différentiable en A_t, nous pouvons donc écrire son développement de Taylor en A_t :

$$f(x,y) = f(x_t,y_t) + (x-x_t)\frac{\partial f}{\partial x}(x_t,y_t) + (y-y_t)\frac{\partial f}{\partial y}(x_t,y_t) + \|(x-x_t, y-y_t)\|\varepsilon(x,y)$$

avec $\varepsilon(x,y) \underset{(x_t,y_t)}{=} o(1)$. Utilisons ce développement pour exprimer $\phi(t+h)$ (h proche de 0 et $t+h \in [0;1]$). Nous obtenons :

$$\phi(t+h) = f(x_{t+h}, y_{t+h}) = f(x_t,y_t)+(x_{t+h}-x_t)\frac{\partial f}{\partial x}(x_t,y_t)+(y_{t+h}-y_t)\frac{\partial f}{\partial y}(x_t,y_t)+\|(x_{t+h}-x_t, y_{t+h}-y_t)\|\varepsilon(x_{t+h},y_{t+h})$$

En fait $x_{t+h} = x_t + h(x_1 - x_0)$ et $y_{t+h} = y_t + h(y_1 - y_0)$, donc l'expression devient :

$$\phi(t+h) = \phi(t) + h(x_1-x_0)\frac{\partial f}{\partial x}(x_t,y_t) + h(y_1-y_0)\frac{\partial f}{\partial y}(x_t,y_t) + |h|\|\overrightarrow{A_0A_1}\|\varepsilon(x_{t+h},y_{t+h})$$

Alors pour $h \neq 0$ nous obtenons

$$\frac{\phi(t+h)-\phi(t)}{h} = (x_1-x_0)\frac{\partial f}{\partial x}(x_t,y_t)+(y_1-y_0)\frac{\partial f}{\partial y}(x_t,y_t)+\bar{\varepsilon}(h) \text{ avec } \bar{\varepsilon}(h) = \frac{|h|}{h}\|\overrightarrow{A_0A_1}\|\varepsilon(x_{t+h},y_{t+h})$$

Lorsque $h \to 0'$ on a $\frac{|h|}{h}\|\overrightarrow{A_0A_1}\| = \pm\|\overrightarrow{A_0A_1}\|$, borné, et $(x_{t+h}, y_{t+h}) \to (x_t, y_t)$, donc $\varepsilon(x_{t+h}, y_{t+h}) \to 0$. Nous concluons

$$\lim_{h\to 0'} \frac{\phi(t+h)-\phi(t)}{h} = (x_1-x_0)\frac{\partial f}{\partial x}(x_t,y_t) + (y_1-y_0)\frac{\partial f}{\partial y}(x_t,y_t), \text{ soit en effet } \phi'(t) = \langle \nabla_{A_t} f, \overrightarrow{A_0A_1}\rangle$$

□

Attention, pour que l'énoncé (quantitatif) ci-dessous soit correct, il faut vraiment utiliser la norme euclidienne !

Théorème 4.2 (inégalité des accroissements finis)**.**

Soit $O \subset \mathbb{R}^2$ un ouvert et soit $f : O \to \mathbb{R}$ une fonction différentiable sur O.

Supposons que $A_0 = (x_0, y_0)$ et $A_1 = (x_1, y_1)$ sont deux points de O tels que le segment $[A_0A_1]$ est entièrement contenu dans O. Supposons de plus que K est un réel ≥ 0 tel que pour tout point $M \in [A_0A_1]$, on a $\|\nabla_M f\|_2 \leq K$.

Alors $|f(A_1) - f(A_0)| \leq K\|\overrightarrow{A_0A_1}\|_2$.

Démonstration. On pose $\phi(t) = f(tx_1 + (1-t)x_0, ty_1 + (1-t)y_0)$ et on sait d'après le Théorème 4.1 que ϕ est dérivable sur $[0;1]$. D'après le théorème des accroissements finis (en une variable) nous savons qu'il existe $c \in]0;1[$ tel que $\phi(1) - \phi(0) = (1-0)\phi'(c)$.

Cela nous donne $f(A_1) - f(A_0) = \phi'(c) = \langle \nabla_{A_c} f, \overrightarrow{A_0A_1}\rangle$, donc par Cauchy-Schwarz

$$|f(A_1) - f(A_0)| \leq \|\nabla_{A_c} f\|_2\|\overrightarrow{A_0A_1}\|_2 \leq K\|\overrightarrow{A_0A_1}\|_2.$$

□

Corollaire 4.3 (gradient borné $\Rightarrow$ Lipschitzienne)**.** *Soit $f : O \to \mathbb{R}$ une fonction différentiable avec $O = \mathbb{R}^2$, ou plus généralement $O =]a;b[\times]c;d[$. On suppose qu'il existe un réel $K \geq 0$ avec la propriété : pour tout $M \in O$ on a $\|\nabla_M f\|_2 \leq K$. Alors f est K-Lipschitzienne sur O.*

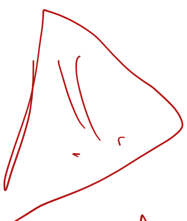
Démonstration. Pour les deux types d'ouverts O considérés on a : si $A_0, A_1 \in O$ alors $[A_0 A_1] \subset O$ et en tout point $M \in [A_0 A_1]$ on a $\|\nabla_M f\|_2 \leq K$, donc on peut appliquer le Théorème 4.2. $\square$

Corollaire 4.4. *Soit $f : O \to \mathbb{R}$ une fonction différentiable de classe $\mathcal{C}^1$ sur l'ouvert O. Alors f est Lipschitzienne sur tout rectangle ouvert $]a; b[\times]c; d[$ tel que le rectangle fermé $[a; b] \times [c; d]$ est contenu dans O.*

Démonstration. Comme la fonction f est supposées de classe $\mathcal{C}^1$, ses deux dérivées partielles sont continues sur O. Soit $[a; b] \times [c; d]$ un rectangle fermé contenu dans O, alors par le théorème de Weierstrass chacune des deux dérivées partielles est bornée sur $[a; b] \times [c; d]$. Donc $\|\nabla_M f\|_2$ est bornée pour $M \in [a; b] \times [c; d]$ et a fortiori pour $M \in]a; b[\times]c; d[$. Il n'y a plus qu'à appliquer le Corollaire 4.3. $\square$

Proposition 4.5 (direction du gradient et sens de variation)**.** *Soit $O \subset \mathbb{R}^2$ un ouvert et soit $f : O \to \mathbb{R}$ une fonction différentiable sur O.*

Supposons que $A_0 = (x_0, y_0)$ et $A_1 = (x_1, y_1)$ sont deux points de O tels que le segment $[A_0 A_1]$ est entièrement contenu dans O.



(1) *Si k est un réel ≥ 0 tel que pour tout point $M \in [A_0 A_1]$, on a $\langle \nabla_M f, \overrightarrow{A_0 A_1} \rangle \geq k$, alors $f(A_1) - f(A_0) \geq k$.*

(2) *Si k est un réel ≥ 0 tel que pour tout point $M \in [A_0 A_1]$, on a $\langle \nabla_M f, \overrightarrow{A_0 A_1} \rangle \leq -k$, alors $f(A_1) - f(A_0) \leq -k$. En particulier $f(A_1) \leq f(A_0)$.*

Produit scalaire

(3) *Si pour tout point $M \in [A_0 A_1]$, on a $\langle \nabla_M f, \overrightarrow{A_0 A_1} \rangle = 0$ alors pour tout point $M \in [A_0 A_1]$ on a $f(M) = f(A_0)$. Et réciproquement : si $f(M) = f(A_0)$ pour tout point $M \in [A_0 A_1]$ alors $\langle \nabla_M f, \overrightarrow{A_0 A_1} \rangle = 0$ (le gradient est perpendiculaire au segment en tout point de ce segment).*

En particulier ($k = 0$) : si le gradient $\nabla_M f$ fait un angle aigu avec $\overrightarrow{A_1 A_2}$ en tout point $M \in [A_1 A_2]$ alors $f(A_1) \geq f(A_0)$, alors que si le gradient $\nabla_M f$ fait un angle obtus avec $\overrightarrow{A_1 A_2}$ en tout point $M \in [A_1 A_2]$ alors $f(A_1) \leq f(A_0)$.

Démonstration. On pose $\phi(t) = f(A_0 + t\overrightarrow{A_0 A_1})$ (pour $t \in [0; 1]$), et on applique le théorème 4.1. On obtient que $\phi'(t) = \langle \nabla_{A_t} f, \overrightarrow{A_0 A_1} \rangle$ (avec $A_t = A_0 + t\overrightarrow{A_0 A_1}$).

(1) Par hypothèse on a donc $\phi'(t) \geq k$ pour tout $t \in [0; 1]$. Or d'après le TAF on a $\phi(1) - \phi(0) = \phi'(c)(1 - 0)$ pour un certain réel $c \in]0; 1[$, ce qui nous donne $\phi(1) - \phi(0) = \phi'(c) \geq k$. D'où l'inégalité demandée. (On aurait de la même manière $f(A_t) - f(A_0) \geq kt$.)

(2) Ici par hypothèse on a $\phi'(t) \leq -k$ pour tout $t \in [0; 1]$. Or d'après le TAF on a $\phi(1) - \phi(0) = \phi'(c)(1 - 0)$ pour un certain réel $c \in]0; 1[$, ce qui nous donne $\phi(1) - \phi(0) = \phi'(c) \leq -k$. D'où à nouveau l'inégalité demandée.

(3) Il suffit de remarquer que ϕ est constante sur $[0; 1]$ si et seulement si le gradient de f en tout point du segment est perpendiculaire au segment $[A_0 A_1]$. Or par la formule de dérivation le long d'un segment on a $\phi'(t) = \langle \nabla_{A_t} f, \overrightarrow{A_0 A_1} \rangle$, et ϕ constante $\iff$ $\phi'(t) = 0$ sur $[0; 1]$ $\iff$ $\langle \nabla_{A_t} f, \overrightarrow{A_0 A_1} \rangle = 0$ sur $[0; 1]$.

$\square$

Pour conclure le chapitre on fixe une fonction $f : O \subset \mathbb{R}^2 \to \mathbb{R}$ définie et différentiable sur l'ouvert O. On considère un point $(x_0, y_0) \in O$ et on suppose que $\nabla_{(x_0, y_0)} f \neq (0, 0)$ - donc que (x_0, y_0) est un point **régulier**.

On considère l'unique courbe de niveau de f qui passe par (x_0, y_0) : rappelons qu'il s'agit de $\mathcal{L}_c(f)$ avec $c = f(x_0, y_0)$. On cherche à décrire $\mathcal{L}_c(f)$ au voisinage du point (x_0, y_0), c'est à dire à décrire l'intersection de $\mathcal{L}_c(f)$ avec une petite boule de centre (x_0, y_0).

Nous ne pouvons pas cette année démontrer l'intégralité du théorème suivant, mais nous pouvons en justifier la dernière affirmation :

Théorème 4.6 (fonction implicite pour décrire la courbe de niveau $\mathcal{L}_c(f)$ et tangente à $\mathcal{L}_c(f)$)**.** *Supposons* $\dfrac{\partial f}{\partial y}(x_0, y_0) \neq 0$*. Alors il existe deux réels* $r, r' > 0$ *et une fonction* $h :]x_0 - r; x_0 + r[\to \mathbb{R}$ *tels que l'équivalence suivante est vraie :*

$$\big((x, y) \in]x_0 - r; x_0 + r[\times]y_0 - r'; y_0 + r'[\text{ et } (x, y) \in \mathcal{L}_c(f)\big) \iff \big(x \in]x_0 - r; x_0 + r[\text{ et } y = h(x)\big)$$

Le graphe de h *est donc un morceau entier de la ligne de niveau passant par* (x_0, y_0).

La tangente au graphe de h *en* (x_0, y_0) *est la droite* Δ *qui admet les trois descriptions équivalentes suivantes :*

(1) Δ *est la droite passant par* (x_0, y_0) *et perpendiculaire à* $\nabla_{(x_0,y_0)} f$

(2) Δ *a pour équation* $(x - x_0)\dfrac{\partial f}{\partial x}(x_0, y_0) + (y - y_0)\dfrac{\partial f}{\partial y}(x_0, y_0) = 0$

(3) Δ *est la droite passant par* (x_0, y_0) *et de pente* $h'(x_0) = -\dfrac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)}$

Si $\dfrac{\partial f}{\partial x}(x_0, y_0) \neq 0$ *on a un énoncé analogue : localement, la courbe de niveau est le graphe d'une fonction* $]y_0 - s; y_0 + s[\to \mathbb{R}$ *dont la tangente en* y_0 *est encore la droite passant par* (x_0, y_0) *et perpendiculaire à* $\nabla_{(x_0,y_0)} f$

Pour h on parle de fonction implicite parce que l'équation $f(x, y) = c$ ne permet pas en général d'exprimer y en fonction de x par une formule explicite - mais le théorème nous dit cependant que h existe !

Nous retenons que la courbe de niveau admet une tangente en (x_0, y_0) et que cette tangente est la la droite passant par (x_0, y_0) et perpendiculaire à $\nabla_{(x_0,y_0)} f$.

Passons comme annoncé à la preuve de la dernière partie du théorème :

Proposition 4.7 (graphe contenu dans une ligne de niveau)**.** *Soit* $h :]x_0 - r; x_0 + r[\to \mathbb{R}$ *une fonction dérivable telle que :*

(1) $h(x_0) = y_0$*, soit* $(x_0, y_0) \in \Gamma(h)$ *;*

(2) pour tout $x \in]x_0 - r; x_0 + r[$ *on a* $f(x, h(x)) = c\,(= f(x_0, y_0))$*, soit* $\Gamma(h) \subset \mathcal{L}_c(f)$ *;*

(3) et $\dfrac{\partial f}{\partial y}(x, h(x)) \neq 0$ *pour tout* $x \in]x_0 - r; x_0 + r[$.

$h'(x)$

Alors $\cancel{f'}(x) = -\dfrac{\frac{\partial f}{\partial x}(x, h(x))}{\frac{\partial f}{\partial y}(x, h(x))}$*, de sorte que l'équation de la tangente à* $\Gamma(h)$ *en* $x = x_0$ *est de l'une des trois formes équivalentes suivantes :*

(1) $y = h(x_0) - \dfrac{\frac{\partial f}{\partial x}(x_0, h(x_0))}{\frac{\partial f}{\partial y}(x_0, h(x_0))}(x - x_0)$

(2) $y = y_0 - \dfrac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)}(x - x_0)$

(3) $\dfrac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \dfrac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0$

Démonstration. Les trois formes de l'équation de la tangente étant clairement équivalentes il suffit de démontrer la formule de la dérivée.

Soit x_1 fixé dans $]x_0 - r; x_0 + r[$, posons $y_1 = h(x_1)$. On veut calculer $h'(x_1)$. Il existe un réel $\alpha > 0$ tel que pour tout réel t de valeur absolue $< \alpha$ on a $x_1 + t \in]x_0 - r; x_0 + r[$. Pour un tel t on doit estimer le taux d'accroissement $\tau(t) = \dfrac{h(x_1+t) - h(x_1)}{t} = \dfrac{h(x_1+t) - y_1}{t}$. Tout ce que l'on sait, c'est que $f(x, h(x)) = c$. Même si cela ne nous définit pas explicitement le nombre $h(x)$ en fonction de x, cela va suffire. On va procéder comme pour la dérivée le long d'un segment : on va utiliser le développement de Taylor de $f(x, y)$ en (x_1, y_1). Il existe une fonction $\varepsilon : O \to \mathbb{R}$ telle que

$$\forall (x,y) \in O, f(x,y) = f(x_1, y_1) + (x - x_1)\frac{\partial f}{\partial x}(x_1, y_1) + (y - y_1)\frac{\partial f}{\partial y}(x_1, y_1) + \|(x - x_1, y - y_1)\|\varepsilon(x, y)$$

avec $\lim\limits_{(x,y)\to(x_1,y_1)} \varepsilon(x, y) = 0$. Appliquons la formule ci-dessus avec $(x, y) = (x_1 + t, h(x_1 + t))$:

$$f(x_1 + t, h(x_1 + t)) = f(x_1, y_1) + t\frac{\partial f}{\partial x}(x_1, y_1) + (h(x_1 + t) - y_1)\frac{\partial f}{\partial y}(x_1, y_1) + \|(t, h(x_1 + t) - y_1)\|\overline{\varepsilon}(t)$$

en posant $\overline{\varepsilon}(t) = \varepsilon(x_1 + t, h(x_1 + t))$. Comme $\Gamma(h) \subset \mathcal{L}_c(f)$ on a $f(x_1 + t, h(x_1 + t)) = f(x_1, y_1) = c$, et donc la relation se simplifie en :

$$0 = t\frac{\partial f}{\partial x}(x_1, y_1) + (h(x_1 + t) - y_1)\frac{\partial f}{\partial y}(x_1, y_1) + \|(t, h(x_1 + t) - y_1)\|\overline{\varepsilon}(t)$$

Par hypothèse $\dfrac{\partial f}{\partial y}(x_1, y_1) \neq 0$ donc pour t tel que $0 < |t| < \alpha$ nous avons

$$\tau(t) = -\frac{\frac{\partial f}{\partial x}(x_1, y_1)}{\frac{\partial f}{\partial y}(x_1, y_1)} + \|(1, \tau(t))\|\frac{|t|}{t}\overline{\varepsilon}(t)$$

La quantité $\dfrac{|t|}{t}$ est bornée sur $] - \alpha; 0[\cup]0; \alpha[$. Par hypothèse h est dérivable donc $\lim\limits_{t\to 0'} \tau(t) = h'(x_1)$, donc $\lim\limits_{t\to 0'} \|(1, \tau(t))\| = \|(1, h'(x_1))\|$, de sorte que la quantité $\|(1, \tau(t))\|$ est également bornée sur $] - \alpha; 0[\cup]0; \alpha[$. Enfin $\lim\limits_{t\to 0'} x_1 + t = x_1$ et (par continuité) $\lim\limits_{t\to 0'} h(x_1 + t) = h(x_1) = y_1$, de sorte que par composition de limites on a $\lim\limits_{t\to 0'} \overline{\varepsilon(t)} = 0$. Pour finir on obtient

$$h'(x_1) = \lim_{t\to 0'} \tau(t) = -\frac{\frac{\partial f}{\partial x}(x_1, y_1)}{\frac{\partial f}{\partial y}(x_1, y_1)}, \text{ ce que l'on voulait.}$$

$\square$

Cinquième partie 5. Introduction aux intégrales doubles

Dans ce chapitre on donne une intuition de ce que sont les intégrales doubles, et quelques moyens de les calculer, mais sans apporter de justifications rigoureuses : le but est ici de sentir comment on mène certains calculs.

Rappelons que le calcul d'une intégrale simple se ramène au calcul d'une primitive. Malheureusement il n'y a pas d'interprétation aussi évidente pour les intégrales doubles, c'est pourquoi on est obligé, pour les définir, de revenir à l'interprétation géométrique de l'intégrale.

Rappelons d'abord l'interprétation géométrique des intégrales simples $\int_a^b f(x)dx$.

Pour une fonction $f : [a; b] \to \mathbb{R}$ positive (mettons : f continue), le nombre $\int_a^b f(x)dx$ représente l'aire de la région $R(f; [a; b])$ du plan comprise entre la base $[a; b] \times \{0\}$ et le graphe $G(f)$.

Par exemple si f est constante $= c$, la région $R(f; [a; b])$ est le rectangle $[a; b] \times [0; c]$, son aire est $c \times (b - a)$, et cela est cohérent avec le calcul par primitives : $\int_a^b cdx = [cx]_a^b = cb - ca$.

1. Définition géométrique et premières propriétés de l'intégrale double.

1.1. Définition comme volume.

Définition 1.1. Soit $D \subset \mathbb{R}^2$ une partie. Pour simplifier supposons D bornée, et fermée. Pour simplifier encore plus nous supposons même que D est l'intérieur d'un contour fermé (bien que ceci soit délicat à définir rigoureusement !), union le contour lui-même pour que D soit fermée.

Soit $f : D \to \mathbb{R}$ une fonction ≥ 0. Pour simplifier supposons f continue sur D.

Nous considérons la région suivante de l'espace de dimension 3 :

$$S(f; D) = \{(x, y, z) \in \mathbb{R}^3 \text{ tels que } (x, y) \in D \text{ et } 0 \leq z \leq f(x, y)\}$$

Donc le solide $S(f; D)$ est la partie de l'espace comprise entre la base $D \times \{0\}$ et le graphe de f.

Alors par définition $\boxed{\text{l'intégrale de } f \text{ sur } D \text{ est le volume du solide } S(f; D)}$, notée

$$\iint_D f(x, y)dxdy.$$

Nous dirons que D est le *domaine d'intégration* de $f(x, y)$.

Lorsque f n'est plus supposée ≥ 0, on considère les deux fonctions positives $f^+(x, y) = \max(f(x, y), 0)$ et $f^-(x, y) = \max(-f(x, y), 0)$, qui vérifient $f(x, y) = f^+(x, y) - f^-(x, y)$. Et on pose

$$\iint_D f(x, y)dxdy = \iint_D f^+(x, y)dxdy - \iint_D f^-(x, y)dxdy$$

(on dit que f^+ est la *partie positive* et f^- la *partie négative* de f, elles sont aussi continues que f ! - utiliser la formule $\max(a, b) = \frac{1}{2}(a + b + |a - b|)$.

Cette définition a un sens physique très concret (on effectue une certaine mesure sur un solide associé à la fonction), mais elle ne permet pas encore des calculs explicites. En effet déterminer le volume d'un solide comme $S(f;D)$ est en général une question redoutable, sauf dans quelques cas particuliers où il existe une formule simple pour le volume. Notons cependant un des très rares calculs immédiats d'après la définition :

Lemme 1.2 (intégrales des fonctions constantes)**.** *Supposons f constante sur son domaine d'intégration D, soit $f(x,y)=c$ pour tout $(x,y)\in D$. Alors*

(1) $S(f;D)$ est un "cylindre" de base $D\times\{0\}$ et de hauteur $|c|$ (au dessus du plan horizontal $z=0$ si $c\geq 0$, au dessous de ce plan si $c<0$), et donc

(2) $\iint_D f(x,y)dxdy = Aire(D)\times c$

Pour parvenir à effectuer des calculs plus généraux, il nous faut d'abord déduire toutes les conséquences de cette définition géométrique. Cela nous permettra ensuite d'obtenir (par passage à la limite) la formule de Fubini, qui rend les calculs faisables.

1.2. **Propriétés découlant de la définition comme volume.**

Proposition 1.3 (l'intégrale est positive et croissante)**.** *Supposons que D est fermé et borné.*

(positivité) *Soit $f:D\to\mathbb{R}$ une fonction continue. Si f est $\underline{\geq 0}$ sur son domaine d'intégration D alors $\iint_D f(x,y)dxdy\geq 0$. (Remarque : si $f\geq 0$ et s'il existe un point (x_0,y_0) tel que $f(x_0,y_0)>0$ et D est voisinage de (x_0,y_0), alors $\iint_D f(x,y)dxdy>0$.)*

(croissance $/D$) *Soit $f:D\to\mathbb{R}$ une fonction ≥ 0 (continue sur son domaine d'intégration D). Si D' est un domaine d'intégration (fermé), D' contenu dans D, alors $\iint_{D'} f(x,y)dxdy\leq\iint_D f(x,y)dxdy$.*

Si de plus $f(x,y)=0$ pour tout $(x,y)\in D\setminus D'$, alors cette inégalité est une égalité :

$$\iint_{D'} f(x,y)dxdy=\iint_D f(x,y)dxdy.$$

(croissance $/f$) *Soit $f,g:D\to\mathbb{R}$ deux fonctions continues.*

Si $\underline{f\leq g \text{ sur } D}$ alors $\iint_D f(x,y)dxdy\leq\iint_D g(x,y)dxdy$.

Démonstration. (1) Lorsque $f\geq 0$ son intégrale est un volume, donc est ≥ 0.

(2) Nous avons $D'\subset D$. Comme $f\geq 0$ sur D, alors $f\geq 0$ sur D'. Donc $\iint_D f(x,y)dxdy=$ Vol$(S(f;D))$ et $\iint_{D'} f(x,y)dxdy=$ Vol$(S(f;D'))$. Comme $D'\subset D$ nous voyons que

$$S(f;D')=\{(x,y,z)\in\mathbb{R}^3,(x,y)\in D' \text{ et } 0\leq z\leq f(x,y)\}\subset\{(x,y,z)\in\mathbb{R}^3,(x,y)\in D \text{ et } 0\leq z\leq f(x,y)\}=S(f;D)).$$

Donc nous obtenons Vol$(S(f;D'))\leq$ Vol$(S(f;D))$, par croissance du volume par rapport à l'inclusion des solides. Et ceci correspond exactement à l'inégalité recherchée entre les intégrales.

Supposons maintenant en plus que $f=0$ sur $D\setminus D'$. Alors nous voyons que $S(f;D)$ est l'union du solide $S(f;D')$ et de la plaque $D\times\{0\}$. Comme la plaque n'a pas d'épaisseur elle est de volume nul. Donc Vol$(S(f;D))=$ Vol$(S(f;D'))$, et nous obtenons l'égalité des intégrales.

(3) Supposons $f \leq g$ sur D. Nous allons démontrer l'inégalité des intégrales doubles dans trois cas particuliers importants, qui suffisent pour la suite. Il nous faudra attendre la fin du chapitre pour obtenir une preuve complète du résultat.

1er cas : $f \geq 0$. Alors $g \geq 0$, on a $\iint_D f(x,y)dxdy = \mathrm{Vol}(S(f;D))$ et $\iint_D g(x,y)dxdy = \mathrm{Vol}(S(g;D))$. Or comme $f \leq g$ nous voyons que $\boxed{S(f;D) \subset S(g;D)}$, donc comme le volume est croissant, nous obtenons $\iint_D f(x,y)dxdy \leq \iint_D g(x,y)dxdy$.

2ème cas : $f \leq 0$ (i.e. $f = -f^-$). Alors $g \leq 0$ et $\iint_D f(x,y)dxdy = -\mathrm{Vol}(S(-f;D))$, $\iint_D g(x,y)dxdy = -\mathrm{Vol}(S(-g;D))$. Comme $g \leq f \leq 0$, on a $0 \leq (-f) \leq (-g)$ donc comme dans le premier cas $\mathrm{Vol}(S(-f;D)) \leq \mathrm{Vol}(S(-g;D))$, ce qui donne l'inégalité recherchée en passant à l'opposé.

3ème cas : L'une des deux fonctions f ou g est de signe constant (ce cas contient les deux autres ... mais il les utilise !). D'abord, supposons que c'est g qui est de signe constant (le raisonnement est semblable lorsque f est de signe constant).

Si $f \geq 0$ sur D alors nous sommes en fait dans le premier cas et nous avons déjà expliqué pourquoi l'inégalité des intégrales est vraie. Si $g \leq 0$ alors nous sommes dans le deuxième cas.

Nous pouvons donc supposer que $g \geq 0$. Si $f \leq 0$ sur D alors $\iint_D f(x,y)dxdy = -\mathrm{Vol}(S(f^-;D)) \leq 0$ tandis que $\iint_D g(x,y)dxdy \geq 0$ (par positivité de l'intégrale), et l'inégalité est vraie.

Il nous reste à comprendre le cas où g est positive et f change de signe sur D.
Soit $D^+ = \{(x,y) \in D, f(x,y) \geq 0\}$. Alors $f^+ = f$ sur D^+, et $f^+ = 0$ sur $D \setminus D^+$.

En utilisant la croissance de l'intégrale par rapport au domaine nous avons $\iint_{D^+} g(x,y)dxdy \leq \iint_D g(x,y)dxdy$. Sur D^+ nous avons $0 \leq f \leq g$ donc d'après le premier cas nous savons que $\iint_{D^+} f(x,y)dxdy \leq \iint_{D^+} g(x,y)dxdy$. Or $\iint_{D^+} f(x,y)dxdy = \mathrm{Vol}(S(f^+;D^+)) = \mathrm{Vol}(S(f^+;D))$. Il vient donc

$$\iint_D f(x,y)dxdy = \mathrm{Vol}(S(f^+;D)) - \mathrm{Vol}(S(f^-;D)) \leq \mathrm{Vol}(S(f^+;D)) = \iint_{D^+} f(x,y)dxdy$$

On conclut en combinant les inégalités obtenues.

□

Remarque 1.4. Pour $f : D \to \mathbb{R}$ une fonction continue sur un domaine d'intégration fermé borné D, notons $D^+ = \{(x,y) \in D, f(x,y) \geq 0\}$ et de même $D^- = \{(x,y) \in D, f(x,y) \leq 0\}$. Nous obtenons

$$\iint_{D^+} f(x,y)dxdy = \iint_D f^+(x,y)dxdy \text{ et } \iint_{D^-} f(x,y)dxdy = \iint_D f^-(x,y)dxdy$$

En effet d'une part $f = f^+$ sur D^+, donc $\iint_{D^+} f(x,y)dxdy = \iint_{D^+} f^+(x,y)dxdy$. Et d'autre part $f^+ \geq 0$ sur D^+ avec $f^+ = 0$ sur $D \setminus D^+$, donc d'après le 2) de la proposition précédente nous savons que $\iint_{D^+} f^+(x,y)dxdy = \iint_D f^+(x,y)dxdy$. Même raisonnement avec D^-, f^-. On peut donc reformuler la définition même de l'intégrale en utilisant les domaines D^+ et D^- plutôt que les fonctions f^+ et f^- :

$$\boxed{\iint_D f(x,y)dxdy = \iint_{D^+} f(x,y)dxdy + \iint_{D^-} f(x,y)dxdy.}$$

C'est une autre formule de découpage qui sera vue à la Proposition 1.7, où l'on demande que l'intersection des sous-domaines d'intégration $D_1 \cap D_2$ soit "petite" (contenue dans une union finie de segments). Ici $D^+ \cap D^- = \{(x,y) \in D, f(x,y) = 0\}$, et cette intersection pourrait contenir tout un disque de rayon > 0 ... Mais, pour compenser, c'est la fonction f qui est nulle sur $D^+ \cap D^-$.

Corollaire 1.5 (encadrement sur D quelconque)**.** *Soit $f : D \to \mathbb{R}$ une fonction continue sur son domaine d'intégration D.*

Supposons que les valeurs $f(x,y)$ soient comprises dans $[m;M]$ (pour $(x,y) \in D$ quelconque). Alors $\mathit{Aire}(D) \times m \leq \iint_D f(x,y)dxdy \leq \mathit{Aire}(D) \times M$.

Démonstration. Comme $f \leq M$ sur D (et que la fonction constante $g(x,y) = M$ ne change pas de signe...) nous savons que $\iint_D f(x,y)dxdy \leq \iint_D g(x,y)dxdy = \text{Aire}(D) \times M$. On obtient de même $\text{Aire}(D) \times m \leq \iint_D f(x,y)dxdy$. □

Corollaire 1.6 (formule de la moyenne sur un rectangle)**.** *Soit $f : [a;b] \times [c;d] \to \mathbb{R}$ une fonction continue sur le rectangle d'intégration $D = [a;b] \times [c;d]$. Alors il existe un point $(x_0, y_0) \in [a;b] \times [c;d]$ tel que $\boxed{\iint_D f(x,y)dxdy = \mathit{Aire}(D) \times f(x_0,y_0)}$.*

Démonstration. Comme f est continue sur le rectangle fermé, nous savons par la combinaison du théorème de Weierstrass et du théorème des valeurs intermédiaires que f prend sur le rectangle $[a;b] \times [c;d]$ toutes les valeurs intermédiaires entre sa valeur minimale m et sa valeur maximale M. D'après l'encadrement des intégrales (Corollaire 1.5) nous savons que $\dfrac{\iint_D f(x,y)dxdy}{\text{Aire}(D)} \in [m;M]$. Donc il existe bien un $(x_0,y_0) \in D$ tel que $f(x_0,y_0) = \dfrac{\iint_D f(x,y)dxdy}{\text{Aire}(D)}$. □

Voici l'analogue pour les intégrales doubles de la relation de Chasles pour les intégrales simples :

Proposition 1.7 (découpage fini du domaine d'intégration)**.** *Soit $f : D \to \mathbb{R}$ une fonction (continue sur son domaine d'intégration D). Supposons que D admette un découpage en deux sous-domaines $D = D_1 \cup D_2$, et que $D_1 \cap D_2$ soit une union finie de segments (ou soit contenue dans une union finie de segments).*

Alors $\boxed{\iint_D f(x,y)dxdy = \iint_{D_1} f(x,y)dxdy + \iint_{D_2} f(x,y)dxdy}$.

Plus généralement si D admet un découpage en un nombre fini de sous-domaines $D = D_1 \cup \cdots \cup D_n$, et que $D_i \cap D_j$ soit (contenue dans) une union finie de segments (pour i différent de j), alors

$$\iint_D f(x,y)dxdy = \iint_{D_1} f(x,y)dxdy + \cdots + \iint_{D_n} f(x,y)dxdy.$$

Démonstration. Le cas d'un découpage général se déduit facilement par récurrence du cas $D = D_1 \cup D_2$, que nous traitons donc uniquement.

On démontre le résultat en considérant d'abord le cas particulier d'une fonction de signe constant.

1er cas : $f \geq 0$. Alors $f \geq 0$ sur D_1 et $f \geq 0$ sur D_2. Nous avons

$$\iint_D f(x,y)dxdy = \mathrm{Vol}(S(f;D)) \ , \ \iint_{D_1} f(x,y)dxdy = \mathrm{Vol}(S(f;D_1)) \ , \ \iint_{D_2} f(x,y)dxdy = \mathrm{Vol}(S(f;D_2))$$

De plus $S(f;D) = S(f;D_1) \cup S(f;D_2)$, et $S(f;D_1) \cap S(f;D_2)$ est (contenue dans) une union de finies plaques verticales d'épaisseur nulle car $D_1 \cap D_2$ est (contenue dans) une union finie de segments. Il s'ensuit que $\mathrm{Vol}(S(f;D)) = \mathrm{Vol}(S(f;D_1)) + \mathrm{Vol}(S(f;D_2))$, ce qui donne l'additivité de l'intégrale dans ce cas.

2ème cas : $f \leq 0$. Donc $f \leq 0$ sur D_1, $f \leq 0$ sur D_2, et nous avons

$$\iint_D f(x,y)dxdy = -\mathrm{Vol}(S(-f;D)) \ , \ \iint_{D_1} f(x,y)dxdy = -\mathrm{Vol}(S(-f;D_1)) \ , \ \iint_{D_2} f(x,y)dxdy = -\mathrm{Vol}(S(-f;D_2))$$

Comme $-f \geq 0$ on peut lui appliquer le raisonnement du premier cas et :

$$\mathrm{Vol}(S(-f;D)) = \iint_D (-f(x,y))dxdy = \iint_{D_1} (-f(x,y))dxdy + \iint_{D_2} (-f(x,y))dxdy = \mathrm{Vol}(S(-f;D_1)) + \mathrm{Vol}(S(-f;D_2))$$

ce qui permet de conclure en passant à l'opposé.

3ème cas : f quelconque. On considère les parties suivantes

$$D^+ = \{(x,y) \in D, f(x,y) \geq 0\}; D^- = \{(x,y) \in D, f(x,y) \leq 0\}; D_1^+ = D_1 \cap D^+; D_1^- = D_1 \cap D^-; D_2^+ = D_2 \cap D^+; D_2^- = D_2 \cap D^-$$

Pour $\varepsilon = \pm$, d'une part nous avons $D^\varepsilon = D_1^\varepsilon \cup D_2^\varepsilon$ avec $D_1^\varepsilon \cap D_2^\varepsilon \subset D_1 \cap D_2$ contenu dans une union finie de segments, donc (en appliquant les cas où f est de signe constant) nous voyons que :

$$\iint_{D^\varepsilon} f(x,y)dxdy = \iint_{D_1^\varepsilon} f(x,y)dxdy + \iint_{D_2^\varepsilon} f(x,y)dxdy$$

D'autre part $\iint_D f(x,y)dxdy = \iint_{D^+} f(x,y)dxdy - \iint_{D^-} f(x,y)dxdy$. Il vient

$$\iint_D f(x,y)dxdy = \iint_{D_1^+} f(x,y)dxdy + \iint_{D_2^+} f(x,y)dxdy - [\iint_{D_1^-} f(x,y)dxdy + \iint_{D_2^-} f(x,y)dxdy]$$

$$= \iint_{D_1^+} f(x,y)dxdy - \iint_{D_1^-} f(x,y)dxdy + \iint_{D_2^+} f(x,y)dxdy - \iint_{D_2^-} f(x,y)dxdy = \iint_{D_1} f(x,y)dxdy + \iint_{D_2} f(x,y)dxdy.$$

□

2. Vers la justification de la notation : sommes de Riemann.

2.1. **Subdivisions de rectangles.**

Définition 2.1. Soit $R = [a;b] \times [c;d] \subset \mathbb{R}^2$ un rectangle fermé (avec $a \leq b$ et $c \leq d$ des réels quelconques). Une *subdivision (finie)* de R est une famille finie $\mathcal{S}(R)$ de rectangles obtenue de la manière suivantes :

(1) L'intervalle $[a;b]$ est décomposé en une union finie $[a;b] = [a_0;a_1] \cup [a_1;a_2] \cup \cdots \cup [a_{p-1};a_p]$ (avec p un entier ≥ 1 et $a = a_0 < a_1 < \cdots < a_{p-1} < a_p = b$)

(2) L'intervalle $[c;d]$ est décomposé en une union finie $[c;d] = [c_0;c_1] \cup [c_1;c_2] \cup \cdots \cup [c_{q-1};c_q]$ (avec q un entier ≥ 1 et $c = c_0 < c_1 < \cdots < c_{q-1} < c_q = d$)

(3) les rectangles de $\mathcal{S}(R)$ sont les rectangles $R_{ij} = [a_i;a_{i+1}] \times [c_j;c_{j+1}]$

Nous noterons $A_{i,j}$ le point (a_i, c_j). Ainsi les quatre sommets de R_{ij} sont $A_{i,j}$, $A_{i+1,j}$, $A_{i+1,j+1}$ et $A_{i,j+1}$. Les points $A_{i,j}$ sont appelés les sommets de subdivision.

La subdivision est *régulière* si $|a_{i+1} - a_i| = \dfrac{b-a}{p}$ (indépendant de i) et $|c_{j+1} - c_j| = \dfrac{d-c}{q}$ (indépendant de j).

Nous utiliserons surtout la subdivision régulière $\mathcal{S}_{2^n}(R)$ où $p = q = 2^n$ (pour un certain entier naturel n). Donc les rectangles de $\mathcal{S}_n(R)$ ont pour dimension $\dfrac{b-a}{2^n}$ et $\dfrac{d-c}{2^n}$. On remarque que les rectangles de la subdivision $\mathcal{S}_{2^{n+1}}(R)$ s'obtiennent en découpant chaque rectangle de $\mathcal{S}_{2^n}(R)$ en quatre, de manière équitable dans les deux directions.

Pour la subdivision $\mathcal{S}_{2^n}(R)$ les sommets de subdivision sont en fait

$$A_{i,j} = (a + \frac{i(b-a)}{2^n}; c + \frac{j(d-c)}{2^n}).$$

Proposition 2.2. *Soit $R = [a;b] \times [c;d] \subset \mathbb{R}^2$ un rectangle fermé. Soit $\mathcal{S}(R) = \{R_{ij} = [a_i;a_{i+1}] \times [c_j;c_{j+1}]\}_{ij}$ une subdivision de R. Soit $f : R \to \mathbb{R}$ une fonction continue.*

(1) $\displaystyle\iint_R f(x,y)dxdy = \sum_{i,j} \iint_{R_{ij}} f(x,y)dxdy$

(2) Pour tout i,j il existe un point $\bar{P}_{i,j} = (\bar{x}_{ij}, \bar{y}_{ij})$ dans R_{ij} tel que

$$\iint_R f(x,y)dxdy = \sum_{i,j} f(\bar{x}_{ij}, \bar{y}_{ij})|a_{i+1} - a_i||c_{j+1} - c_j| = \sum_{i,j} f(\bar{x}_{ij}, \bar{y}_{ij})\delta_x^i \delta_y^j$$

(en notant $|a_{i+1} - a_i| = \delta_x^i$ et $|c_{j+1} - c_j| = \delta_y^j$)

2.2. **Sommes de Riemann.**

Définition 2.3. Soit $R = [a;b] \times [c;d] \subset \mathbb{R}^2$ un rectangle fermé. Soit $f : R \to \mathbb{R}$ une fonction continue.

Une somme de Riemann de f est une somme de la forme

$$\Sigma(f) = \sum_{i,j} f(x_{ij}, y_{ij})|a_{i+1} - a_i||c_{j+1} - c_j| = \sum_{i,j} f(x_{ij}, y_{ij})\mathrm{Aire}(R_{ij})$$

où l'on a choisi une subdivision $\mathcal{S}(R)$ de R et où l'on a choisi, pour chaque rectangle R_{ij} de la subdivision, un point $P_{i,j} = (x_{ij}, y_{ij})$ dans le rectangle R_{ij}.

Définition 2.4 (sommes de Riemann supérieure, inférieure, au coin, adaptée à l'intégrale)**.** Soit $R = [a;b] \times [c;d] \subset \mathbb{R}^2$ un rectangle fermé. Soit $f : R \to \mathbb{R}$ une fonction continue. Dans les quatre cas suivants nous considérons la subdivision $\mathcal{S}_{2^n}(R)$.

Pour définir une somme de Riemann il reste à expliquer comment on choisit un point $P_{i,j} \in R_{ij}$:

(1) (sommes de Riemann au coin inférieur gauche) On choisit ici $P_{i,j} = (a_i, c_j) = A_{i,j}$, cela donne donc la somme de Riemann $\Sigma_n^{\swarrow}(f) = \sum_{ij} f(a_i, c_j) \frac{(b-a)}{2^n} \frac{(d-c)}{2^n}$

(2) (sommes de Riemann supérieure) La fonction f est continue sur le rectangle R_{ij}, donc elle y est majorée, et y atteint son maximum en un point $P^+_{i,j}$ (dépend de f !). Cela donne donc la somme de Riemann

$$\Sigma_n^+(f) = \sum_{ij} f(P^+_{i,j}) \frac{(b-a)}{2^n} \frac{(d-c)}{2^n} = \sum_{ij} [\max_{(x,y)\in R_{ij}} f(x,y)] \frac{(b-a)}{2^n} \frac{(d-c)}{2^n}.$$

(3) (sommes de Riemann inférieure) La fonction f est continue sur le rectangle R_{ij}, donc elle y est minorée, et y atteint son minimum en un point $P^-_{i,j}$. Cela donne donc la somme de Riemann

$$\Sigma_n^-(f) = \sum_{ij} f(P^-_{i,j}) \frac{(b-a)}{2^n} \frac{(d-c)}{2^n} = \sum_{ij} [\min_{(x,y)\in R_{ij}} f(x,y)] \frac{(b-a)}{2^n} \frac{(d-c)}{2^n}.$$

(4) D'après la Proposition 2.2 nous voyons qu'il est possible de choisir un point $\bar{P}_{i,j}$ tel que la somme de Riemann correspondante $\Sigma_n^=(f)$ vérifie $\Sigma_n^=(f) = \iint_R f(x,y)dxdy$.

Proposition 2.5. *Soit $R = [a;b] \times [c;d] \subset \mathbb{R}^2$ un rectangle fermé. Soit $f : R \to \mathbb{R}$ une fonction continue.*

(1) Quel que soit $n \in \mathbb{N}$, pour toute somme de Riemann $\Sigma_n(f)$ (utilisant la subdivision $\mathcal{S}_n(R)$ avec un choix arbitraire de point $P_{i,j}$ dans R_{ij}), on a

$$\Sigma_n^-(f) \le \Sigma_n(f) \le \Sigma_n^+(f)$$

(2) Quel que soit $n \in \mathbb{N}$, on a

$$\Sigma_n^-(f) \le \iint_R f(x,y)dxdy \le \Sigma_n^+(f)$$

(3) La suite $\big(\Sigma_n^-(f)\big)_n$ est croissante, la suite $\big(\Sigma_n^+(f)\big)_n$ est décroissante.

(4) En fait $|\Sigma_n^+(f) - \Sigma_n^-(f)| \to 0$.

Les trois suites $\big(\Sigma_n^-(f)\big)_n$, $\big(\Sigma_n^+(f)\big)_n$ et $\big(\Sigma_n(f)\big)_n$ convergent vers $\iint_R f(x,y)dxdy$.

Démonstration. (1) Quel que soit le point $P_{i,j}$ choisi dans R_{ij} on a $f(P^-_{i,j}) \le f(P_{i,j}) \le f(P^+_{i,j})$. En multipliant par Aire(R_{ij}) et en sommant sur i,j on obtient l'encadrement de sommes de Riemann.

(2) L'inégalité annoncée est un cas particulier de la précédente, en prenant comme choix pour $P_{i,j}$ un point $\bar{P}_{i,j}$ tel que $\iint_{R_{ij}} f(x,y)dxdy = f(\bar{P}_{i,j})\text{Aire}(R_{ij})$, puisque dans ce cas on a $\Sigma_n(f) = \iint_R f(x,y)dxdy$.

(3) Découpons un rectangle $R'(\subset R)$ en quatre rectangles : $R' = R'_1 \cup R'_2 \cup R'_3 \cup R'_4$, avec chaque R'_k de dimension moitié de celles de R'. Alors $\min_{(x,y)\in R'} f(x,y) \le \min_{(x,y)\in R'_k} f(x,y)$ (l'une de ces inégalités est une égalité). Donc

$$\big(\min_{(x,y)\in R'} f(x,y)\big)\text{Aire}(R') = \big(\min_{(x,y)\in R'} f(x,y)\big)[\text{Aire}(R'_1) + \text{Aire}(R'_2) + \text{Aire}(R'_3) + \text{Aire}(R'_4)]$$

$\leq \text{Aire}(R'_1)\big(\min_{(x,y)\in R'_1} f(x,y)\big) + \text{Aire}(R'_2)\big(\min_{(x,y)\in R'_2} f(x,y)\big) + \text{Aire}(R'_3)\big(\min_{(x,y)\in R'_3} f(x,y)\big) + \text{Aire}(R'_4)\big(\min_{(x,y)\in R'_4} f(x,y)\big)$

En appliquant cette inégalité sur chaque rectangle R_{ij} de la subdivision $\mathcal{S}_n(R)$, nous obtenons $\Sigma_n^-(f) \leq \Sigma_{n+1}^-(f)$. On obtient de manière analogue $\Sigma_{n+1}^+(f) \leq \Sigma_n^+(f)$

(4) Nous donnons l'argument lorsque f est K-Lipschitzienne sur R (muni de la distance d_1). Notons d'abord que

$$\Sigma_n^+(f) - \Sigma_n^-(f) = \sum_{ij} [\max_{R_{ij}} f - \min_{R_{ij}} f] \text{Aire}(R_{ij}), \text{ d'où } |\Sigma_n^+(f) - \Sigma_n^-(f)| \leq \sum_{ij} |[\max_{R_{ij}} f - \min_{R_{ij}} f]| \text{Aire}(R_{ij})$$

Comme f est continue sur R_{ij} il existe des points $P_{i,j}^+, P_{i,j}^- \in R_{ij}$ tels que $\max_{R_{ij}} f = f(P_{i,j}^+)$ et $\min_{R_{ij}} f = f(P_{i,j}^-)$. Alors

$$|\max_{R_{ij}} f - \min_{R_{ij}} f| = |f(P_{i,j}^+) - f(P_{i,j}^-)| \leq K d_1(P_{i,j}^+, P_{i,j}^-) \leq K \frac{(b-a)+(d-c)}{2^n}$$

En combinant les deux inégalités obtenues

$$|\Sigma_n^+(f) - \Sigma_n^-(f)| \leq \sum_{ij} K \frac{(b-a)+(d-c)}{2^n} \text{Aire}(R_{ij}) = \big(K \frac{(b-a)+(d-c)}{2^n}\big)\big(\sum_{ij} \text{Aire}(R_{ij})\big) = K \frac{(b-a)+(d-c)}{2^n} \text{Aire}(R)$$

Donc clairement $|\Sigma_n^+(f) - \Sigma_n^-(f)| \to 0$.

Maintenant les deux suites $\big(\Sigma_n^-(f)\big)_n$ et $\big(\Sigma_n^+(f)\big)_n$ sont adjacentes, donc elles convergent vers la même limite - que nous notons I. Par le théorème des gendarmes (et le point (2) de cette proposition), il vient alors que la limite commune I est en fait l'intégrale $\iint_R f(x,y)dxdy$. Enfin toute somme de Riemann $\Sigma_n(f)$ converge vers I également, d'après le point (1) de cette proposition (et le théorème des gendarmes).

□

Remarque 2.6 (Sommes de Riemann et intégrales des fonctions d'une variable)**.** Ce que nous avons établi ci-dessus est une généralisation aux fonctions de deux variables d'une approche similaire pour l'intégration en une variable ! En décalquant la preuve ci-dessus on obtiendrait le résultat suivant :

Soit $f : [a;b] \to \mathbb{R}$ une fonction K-Lipschitzienne. Pour chaque entier $n \in \mathbb{N}$ et pour chaque entier i entre 0 et $2^n - 1$, on choisit un réel x_i dans l'intervalle $[a_i; a_{i+1}]$ (avec $a_i = a + \frac{i(b-a)}{2^n}$), et enfin on pose

$$\Sigma_n(f) = \sum_i f(x_i) \frac{b-a}{2^n}$$

Alors $|\Sigma_n(f) - \int_a^b f(x)dx| \leq K \times \frac{b-a}{2^n} \times \text{Long}([a;b]) = \frac{K(b-a)^2}{2^n}$. En particulier la somme de Riemann $\Sigma_n(f)$ converge vers $\int_a^b f(x)dx$.

2.3. **Linéarité de l'intégrale double.**

Théorème 2.7. *Soient f, g deux fonctions continues sur un rectangle fermé $R = [a;b] \times [c;d]$. Soit λ, μ deux réels. Alors*

$$\iint_R (\lambda f + \mu g)(x,y)dxdy = \lambda \iint_R f(x,y)dxdy + \mu \iint_R g(x,y)dxdy.$$

Démonstration. Pour toute subdivision $\mathcal{S}_n(R)$ et tout choix de points $P_{i,j}$ dans R_{ij} on a

$$\Sigma_n(\lambda f+\mu g)=\sum_{ij}\big(\lambda f+\mu g)(P_{i,j})\mathrm{Aire}(R_{ij})=\sum_{ij}\big(\lambda f(P_{i,j})\mathrm{Aire}(R_{ij})+\mu g(P_{i,j})\mathrm{Aire}(R_{ij})\big)$$

$$=\lambda\big(\sum_{ij}f(P_{i,j})\mathrm{Aire}(R_{ij})\big)+\mu\big(\sum_{ij}g(P_{i,j})\mathrm{Aire}(R_{ij})\big)=\lambda\Sigma_n(f)+\mu\Sigma_n(g).$$

Le résultat s'obtient en passant à la limite quand $n\to\infty$ puisque les sommes de Riemann convergent vers l'intégrale. □

3. Fubini et conséquences.

3.1. **Le théorème de Fubini sur un rectangle.**

Théorème 3.1. *Soit $R=[a;b]\times[c;d]\subset\mathbb{R}^2$ un rectangle fermé. Soit $f:R\to\mathbb{R}$ une fonction continue.*

Pour tout $y_0\in[c;d]$ fixé, posons $I_{f/x}(y_0)=\int_a^b f(x,y_0)dx$. Pour tout $x_0\in[a;b]$ fixé, posons $I_{f/y}(x_0)=\int_c^d f(x_0,y)dy$. Ce sont des intégrales partielles de f.

Alors les deux fonctions $I_{f/x}:[c;d]\to\mathbb{R}$ et $I_{f/y}:[a;b]\to\mathbb{R}$ sont continues. De plus

$$\int_c^d I_{f/x}(y)dy=\iint_{[a;b]\times[c;d]}f(x,y)dxdy=\int_a^b I_{f/y}(x)dx,\ \text{ en d'autres termes :}$$

$$\iint_{[a;b]\times[c;d]}f(x,y)dxdy=\int_{y=c}^{y=d}\big(\int_{x=a}^{x=b}f(x,y)dx\big)dy=\int_{x=a}^{x=b}\big(\int_{y=c}^{y=d}f(x,y)dy\big)dx$$

Démonstration. Nous ne traitons ici que le cas particulier où f est K-Lipschitzienne sur R (pour d_1, d_2 ou d_∞) (même si le théorème est vrai pour f continue quelconque).

D'abord la continuité des intégrales partielles $I_{f/x}$ et $I_{f/y}$: nous allons voir que $I_{f/x}$ est $K(b-a)$-Lipschitzienne (analogue pour $I_{f/y}$), donc continue. En effet

$$|I_{f/x}(y_1)-I_{f/x}(y_2)|=|\int_a^b f(x,y_1)dx-\int_a^b f(x,y_2)dx|=|\int_a^b[f(x,y_1)-f(x,y_2)]dx|$$

$$\leq\int_a^b|f(x,y_1)-f(x,y_2)|dx\leq\int_a^b K|y_1-y_2|dx=K(b-a)|y_1-y_2|$$

Pour approcher l'intégrale double, passons par une somme de Riemann, celle où l'on prend la valeur de f au coin inférieur gauche $A_{i,j}$ du rectangle R_{ij}. Nous avons

$$\Sigma_n^{\swarrow}(f)=\sum_{ij}f(A_{ij})\frac{(b-a)(d-c)}{2^n.2^n}=\sum_j\big(\sum_i f(A_{ij})\frac{b-a}{2^n}\big)\frac{d-c}{2^n}$$

L'entier j étant fixé (entre 0 et 2^n-1), on interprète la somme $\sum_i f(A_{ij})\frac{b-a}{2^n}$ sous la forme $\sum_i f(A_{ij})\frac{b-a}{2^n}=\sum_i f(a_i;c_j)\frac{b-a}{2^n}$: c'est une somme de Riemann de l'application partielle $x\mapsto f(x;c_j)$ sur l'intervalle $[a;b]$ (intégration en une variable). Cette somme de Riemann converge vers

$\displaystyle\int_a^b f(x;c_j)dx$, c'est à dire $I_{f/x}(c_j)$, et la valeur absolue de la différence entre l'intégrale et la somme de Riemann est $\leq K(b-a)\times\dfrac{b-a}{2^n}\times\text{Long}([a;b])=\dfrac{K(b-a)^3}{2^n}$.

Nous avons alors

$$|\Sigma_n^{\swarrow}(f)-\sum_j I_{f/x}(c_j)\frac{d-c}{2^n}|=|\sum_j\big(\sum_i f(A_{ij})\frac{b-a}{2^n}-I_{f/x}(c_j)\big)\frac{d-c}{2^n}|$$

$$\leq\sum_j|\sum_i f(A_{ij})\frac{b-a}{2^n}-I_{f/x}(c_j)|\frac{d-c}{2^n}\leq\sum_j\frac{K(b-a)^3}{2^n}\frac{d-c}{2^n}=\frac{K(b-a)^3(d-c)}{2^n}\ (*).$$

Pour conclure l'expression $\displaystyle\sum_j I_{f/x}(c_j)\frac{d-c}{2^n}$ est une somme de Riemann de la fonction Lipschitzienne $I_{f/x}$ sur $[c;d]$, donc cette somme tend vers $\displaystyle\int_c^d I_{f/x}(y)dy$. Tandis que $\Sigma_n^{\swarrow}(f)\to\displaystyle\iint_R f(x,y)dxdy$. D'après $(*)$ les limites sont les mêmes, donc $\displaystyle\iint_R f(x,y)dxdy=\int_c^d I_{f/x}(y)dy$. □

Exemple 3.2.

Soit $D=[0;2]\times[-1;4]$. On cherche à calculer $\displaystyle\iint_D(x^2+y^2)\,dxdy$. Alors d'abord par linéarité de l'intégrale $\displaystyle\iint_D(x^2+y^2)\,dxdy=\iint_D x^2\,dxdy+\iint_D x^2+y^2\,dxdy$. Ensuite par Fubini on a

$$\iint_{[0;2]\times[-1;4]}x^2\,dxdy=\int_{x=0}^{x=2}\big(\int_{y=-1}^{y=4}x^2dy\big)dx=\int_{x=0}^{x=2}5x^2dx=5[\frac{x^3}{3}]_0^2=\frac{40}{3}\text{ et}$$

$$\iint_{[0;2]\times[-1;4]}y^2\,dxdy=\int_{y=-1}^{y=4}\big(\int_{x=0}^{x=2}y^2dx\big)dy=\int_{y=-1}^{y=4}2y^2dx=2[\frac{y^3}{3}]_{y=-1}^{y=4}=\frac{130}{3}.$$

Pour conclure $\displaystyle\iint_D(x^2+y^2)\,dxdy=\frac{40}{3}+\frac{130}{3}=\frac{170}{3}$.

3.2. **Le théorème de Fubini entre deux graphes.**

3.2.1. *Domaine entre deux graphes.*

Définition 3.3. Supposons que $u:[e;f]\to\mathbb{R}$ et $v:[e;f]\to\mathbb{R}$ sont deux fonctions continues telles que $u\leq v$ sur $[e;f]$. Nous pouvons alors considérer deux domaines d'intégration :

$D_{u\leq y\leq v}=\{(x,y)\in\mathbb{R}^2$ tels que $x\in[e;f]$ et $u(x)\leq y\leq v(x)\}$ (en tranches verticales), et

$D_{u\leq x\leq v}=\{(x,y)\in\mathbb{R}^2$ tels que $x\in[e;f]$ et $u(y)\leq x\leq v(y)\}$ (en tranches horizontales).

Exemple 3.4. On peut obtenir beaucoup de domaines familiers sous cette forme !

(1) Le rectangle est à la fois un domaine en tranches verticales et en tranches horizontales.

(2) Un trapèze dont la base et le côté opposé parallèle sont verticaux entre deux abscisses a et b est un domaine en tranches verticales $D_{u\leq y\leq v}$, avec $u:[a;b]\to\mathbb{R}$ et $v:[a;b]\to\mathbb{R}$ deux fonctions affines telles que $u\leq v$ sur $[a;b]$.

Si l'un des côtés verticaux est réduit à un point, on obtient un triangles dont un côté est vertical.

Ainsi le triangle (plein) ABC avec pour sommets $A = (1,1), B = (3,0), C = (3,3)$ est en fait
$$ABC = \{(x,y) \in \mathbb{R}^2 \text{ tels que } x \in [1;3] \text{ et } x + 2y - 3 \leq y \leq x - y\}$$
(le graphe de la fonction affine $u(x) = x + 2y - 3$ est la droite (AB), le graphe de la fonction affine $v(x) = x - y$ est la droite (AC))

Un triangle dont un des côté est horizontal est quant à lui un domaine en tranches horizontales : ainsi le triangle (plein) ABC avec pour sommets $A = (1,1), B = (3,1), C = (4,3)$ est en fait
$$ABC = \{(x,y) \in \mathbb{R}^2 \text{ tels que } y \in [1;3] \text{ et } \frac{1}{2}(3y-1) \leq x \leq \frac{1}{2}(y+5)\}$$

(3) Le disque fermé unité $D = \{(x,y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$ est à la fois un domaine en tranche verticales et un domaine en tranches horizontales :
$$D = \{(x,y) \in \mathbb{R}^2 \text{ tels que } x \in [-1;1] \text{ et } -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}\}$$
$$(= \{(x,y) \in \mathbb{R}^2 \text{ tels que } y \in [-1;1] \text{ et } -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2}\}).$$

3.2.2. *Intégration en tranches.* Soit $f : D \to \mathbb{R}$ une fonction continue sur un domaine d'intégration D qui est fermé et borné. Nous admettrons les faits suivants :

(1) On peut définir des sommes de Riemann pour f sur le domaine D s'il est de la forme $D_{u \leq y \leq v}$ ou $D_{u \leq x \leq v}$ avec u, v de classe $\mathcal{C}^1$.
(2) Ces sommes de Riemann convergent vers $\iint_D f(x,y)dxdy$.
(3) Quand le domaine D est en tranche verticales, on peut organiser la somme de Riemann (qui est une somme double) : on fixe d'abord l'abscisse $x = x_i$, puis on somme les termes de $u(x_i)$ à $v(x_i)$, et ensuite on somme le résultat en faisant varier x_i de a à b.
(4) De même quand le domaine D est en tranche horizontales, on peut organiser la somme de Riemann dans l'autre sens : on fixe d'abord l'ordonnée $y = y_j$, puis on somme les termes de $u(y_j)$ à $v(y_j)$, et ensuite on somme le résultat en faisant varier y_j de c à d.

Pour une somme de Riemann sur un domaine en tranches verticales, la somme partielle en tranche verticale (à $x = x_i$ fixé) est une somme de Riemann de l'application partielle $y \mapsto f(x_i, y)$, multipliée par $\dfrac{b-a}{2^n}$. Donc la somme de Riemann de f sur D est une somme de Riemann de $x \mapsto \int_{u(x)}^{v(x} f(x_i, y)dy$, à une erreur tendant vers 0 près (résultat analogue pour une somme de Riemann sur un domaine en tranches horizontales). Ceci explique pourquoi les deux résultats suivants sont valides :

Théorème 3.5 (Fubini en tranches verticales)**.** *Supposons que* $u : [a;b] \to \mathbb{R}$ *et* $v : [a;b] \to \mathbb{R}$ *sont deux fonctions continues telles que* $u \leq v$ *sur* $[a;b]$.

Soit $D = \{(x,y) \in \mathbb{R}^2$ *tels que* $x \in [a;b]$ *et* $u(x) \leq y \leq v(x)\}$ *et soit* $f : D \to \mathbb{R}$ *une fonction continue.*

Pour tout $x_0 \in [a;b]$ *fixé, posons* $I_{f/y}(x_0) = \int_{u(x_0)}^{v(x_0)} f(x_0, y)dy$.

Alors la fonction $I_{f/y} : [a;b] \to \mathbb{R}$ *est continue. De plus*
$$\iint_D f(x,y)dxdy = \int_a^b I_{f/y}(x)dx, \text{ en d'autres termes :}$$

$$\iint_{a\leq x\leq b, u(x)\leq y\leq v(x)} f(x,y)dxdy = \int_a^b \big(\int_{u(x)}^{v(x)} f(x,y)dy \big) dx$$

Théorème 3.6 (Fubini en tranches horizontales)**.** *Supposons que* $u : [c;d] \to \mathbb{R}$ *et* $v : [c;d] \to \mathbb{R}$ *sont deux fonctions continues telles que* $u \leq v$ *sur* $[c;d]$.

Soit $D = \{(x,y) \in \mathbb{R}^2$ *tels que* $y \in [c;d]$ *et* $u(y) \leq x \leq v(y)\}$ *et soit* $f : D \to \mathbb{R}$ *une fonction continue.*

Pour tout $y_0 \in [c;d]$ *fixé, posons* $I_{f/x}(y_0) = \displaystyle\int_{u(y_0)}^{v(y_0)} f(x,y_0)dx$.

Alors la fonction $I_{f/x} : [c;d] \to \mathbb{R}$ *est continue. De plus*

$$\iint_D f(x,y)dxdy = \int_c^d I_{f/x}(y)dy, \text{ en d'autres termes :}$$

$$\iint_{c\leq y\leq d, u(y)\leq x\leq v(y)} f(x,y)dxdy = \int_c^d \big(\int_{u(y)}^{v(y)} f(x,y)dx \big) dy$$