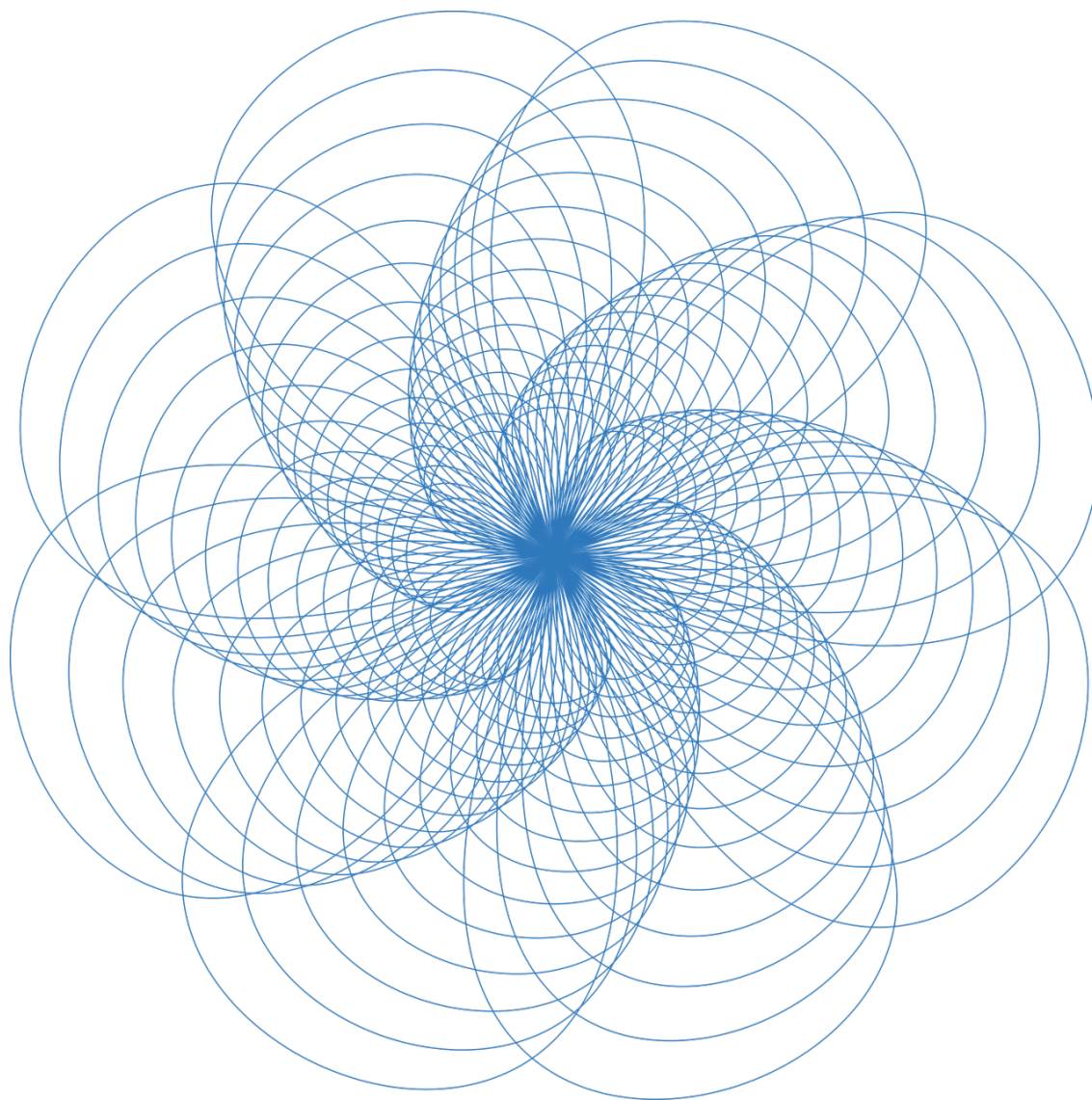


MATHÉMATIQUES

LICENCE



UNIVERSITÉ JEAN-FRANÇOIS CHAMPOLLION
ANNÉE 2022 - 2025

TABLE DES MATIÈRES

VII — FONCTIONS VECTORIELLES

Dans ce chapitre nous étudions les propriétés analytiques et de régularité d'un ensemble plus général de fonctions, les **fonctions d'une variable à valeurs vectorielles** et plus précisément, on se concentrera principalement sur le cas des fonctions de \mathbb{R} dans \mathbb{R}^n muni de sa structure euclidienne si nécessaire.

Définition:

Soit $A \subset \mathbb{R}$ une partie de \mathbb{R} et $(f_1, f_2, \dots, f_n) \in \mathcal{F}(A, \mathbb{R})$ des fonctions réelles, alors on appelle **fonction vectorielle** les fonctions de la forme:

$$\begin{aligned} f : A &\longrightarrow \mathbb{R}^n \\ t &\longmapsto (f_1(t), f_2(t), \dots, f_n(t)) \end{aligned}$$

Dans le cas où $n = 2$ et si

Limites:

On caractérise alors la **limite d'une fonction vectorielle** f en un point $a \in A$, par la limite par composante, ie on a:

$$\lim_{x \rightarrow a} f(t) = l \iff \forall i \in \llbracket 1 ; n \rrbracket ; \lim_{x \rightarrow a} f_i(t) = l_i$$

En particulier on remarque alors la simplicité de cette généralisation, les limites se calculent simplement composantes par composantes. On verra par la suite que cette définition permet bien d'étendre toutes les notions analytiques classiques.

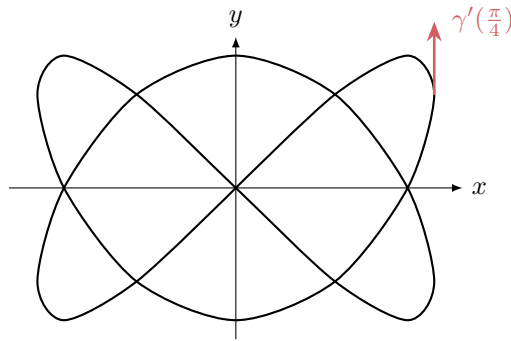
Propriétés de régularité:

La limite étant caractérisée composantes par composantes, la notion de continuité s'étend naturellement composantes par composantes. Par ailleurs, on étend aussi la notion de dérivabilité et on dit qu'une fonction vectorielle f est **dérivable en a** si et seulement si la limite suivante existe:

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

Ce qui équivaut alors à l'existence des limites correspondantes pour les composantes, et f est donc dérivable en a (par extension de classe \mathcal{C}^k) si chacune de ses composantes sont dérivables en a (ou de classe \mathcal{C}^k en a) et on note alors f' sa **dérivée** qui s'obtient simplement en dérivant composante par composante.

Géométriquement $f'(t)$ représente le **vecteur tangent** (aussi appelé vecteur vitesse) à la courbe au temps t . Dans \mathbb{R}^2 par exemple:



Propriétés non-conservées:

Néanmoins pour le cas de fonctions vectorielles, certaines propriétés ne sont pas conservées. En particulier, le **théorème de Rolle** n'est plus valable ainsi que le **théorème des accroissements finis**.¹ Néanmoins, l'inégalité des accroissements finis reste valide et s'interprète facilement cinématiquement:

Un point mobile qui se déplace à une vitesse instantanée inférieure à K sur un temps T , se trouve au maximum à une distance KT de son point de départ.

Dérivées de composées :

Soit f une fonction vectorielle dérivable, on peut alors se demander comment, étant donnée une application linéaire L , se comporte la dérivée de la composée $L \circ f$ et on montre alors à partir de la définition que:

$$(L(f))' = L(f')$$

On veut alors généraliser, et on considère deux fonctions vectorielles f, g dérivables, et une application bilinéaire B , alors de la même manière on peut montrer que:

$$B(f, g)' = B(f', g) + B(f, g')$$

On peut alors y voir un analogue de la **règle du produit** pour la dérivation classique, en particulier si $B = \langle \cdot | \cdot \rangle$, on a:

$$\langle f | g \rangle' = \langle f' | g \rangle + \langle f | g' \rangle$$

Finalement, si on a une famille f_1, \dots, f_n de fonctions vectorielles dérivables et une application multilinéaire M , alors:

$$M(f_1, \dots, f_n)' = M(f_1', \dots, f_n) + \dots + M(f_1, \dots, f_n')$$

En particulier si $M = \det$, et pour les fonctions f, g, h on a:

$$\det(f, g, h)' = \det(f', g, h) + \det(f, g', h) + \det(f, g, h')$$

Intégration:

On peut alors raisonnablement penser que l'intégration sur un segment $[a; b]$ se généralisera de la même manière, et c'est le cas ! Si toutes les fonctions (f_i) sont Riemann-intégrables sur le segment, alors on dira que la fonction f est Riemann-intégrable et on définit:

$$\int_a^b f(t) dt = \left(\int_a^b f_1(t) dt, \dots, \int_a^b f_n(t) dt \right)$$

En particulier cela donne lieu à une interprétation cinématique :

Si f est une fonction qui donne le vecteur vitesse d'un point mobile à un temps t , alors son intégrale est le vecteur position associé.

Points remarquables:

On peut par la suite définir plus d'outils analytiques sur cet ensemble de fonctions, mais pour cela il faut définir la notion d'arcs paramétré qui nécessite les définitions suivantes qui se ne sont que terminologie:

- On appelle **point régulier** un point $f(t)$ tel que $f'(t) \neq 0$
- On appelle **point critique** un point $f(t)$ tel que $f'(t) = 0$

On appelle **point multiple** un point P de E tel qu'il existe (t_0, \dots, t_k) deux à deux distincts tels que $\forall i \in \llbracket 1; k \rrbracket f(t_i) = P$ et on appelle alors **multiplicité** l'entier k .

¹Evident après un dessin.

VII — FONCTIONS À PLUSIEURS VARIABLES

Dans ce chapitre nous étudions les propriétés analytiques et de régularité d'un ensemble plus général de fonctions, les **fonctions d'un espace vectoriel normé dans un autre** et plus précisément, on se concentrera principalement sur le cas des fonctions de \mathbb{R}^n dans \mathbb{R}^p .

Soit \mathcal{D} une partie de \mathbb{R}^n et $(x_1, \dots, x_n) \in \mathcal{D}$, on considère alors la fonction:

$$\begin{aligned} f : \mathcal{D} \subseteq \mathbb{R}^n &\longrightarrow \mathbb{R}^p \\ (x_1, \dots, x_n) &\longmapsto (f_1(x), \dots, f_p(x)) \end{aligned}$$

On dira alors que la fonction f_i est la i -ème application **composante** de f .

Définitions:

On peut alors définir quelques notions de vocabulaire:

- Si la fonction est à valeur scalaire, on dira que c'est un **champ de scalaires**.
- Si la fonction est à valeur vectorielles, on dira que c'est un **champ de vecteurs**.

Représentations:

On cherche alors à représenter de telles fonctions, on sait que dans le cas élémentaire des fonctions à une seule variable, on peut représenter le **graphe** d'une fonction f donné par:

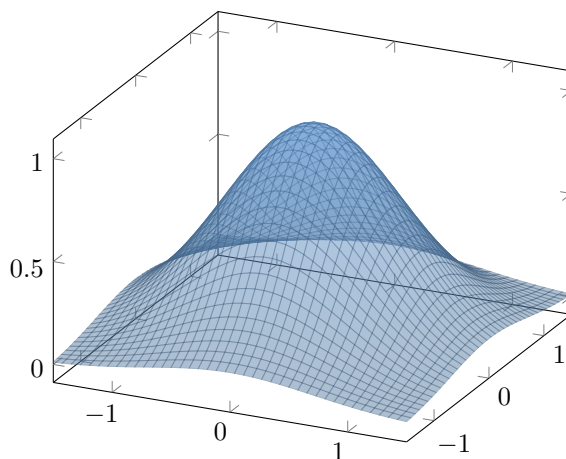
$$\mathcal{G}_f := \{(x, f(x)) \in \mathbb{R}^2 ; x \in \mathbb{R}\}$$

Dans le cas d'une fonction de plusieurs variables, on peut alors généraliser cette définition par:

$$\mathcal{G}_f := \{(x, f(x)) ; x \in \mathbb{R}^n\}$$

De façon générale, le graphe d'une fonction de \mathbb{R}^n dans \mathbb{R}^p est un objet à n dimensions dans l'espace à $(n+p)$ dimensions (les points sont caractérisés par n paramètres, en termes savants on dira qu'il s'agit d'une sous-variété de $\mathbb{R}^{(n+p)}$ de dimension n).

Si on considère par exemple une fonction de \mathbb{R}^2 dans \mathbb{R} , c'est alors une fonction qui à chaque point du plan associe une valeur numérique, qu'on peut alors représenter comme **une hauteur** et on obtient alors **une surface**, par exemple:



Graphe de la fonction $f : (x, y) \mapsto \exp(-x^2 - y^2)$

Dans le cas général néanmoins, il est plus difficile de représenter ce type de fonctions mais on peut par exemple s'inspirer des fonctions $\mathbb{C} \rightarrow \mathbb{C}$ et utiliser la **couleur** comme 4-ème dimension de l'image. On peut aussi de manière heuristique représenter simplement le domaine de départ et/ou d'arrivée pour étudier le comportement de la fonction.

Cas des champs de vecteurs:

De manière générale, le caractère multidimensionnel de l'espace d'arrivée importe peu dans tout le chapitre qui va suivre, les notions analytiques étant définies à partir d'une limite, elles passeront aux composantes comme les notions évoquées au chapitre précédent.

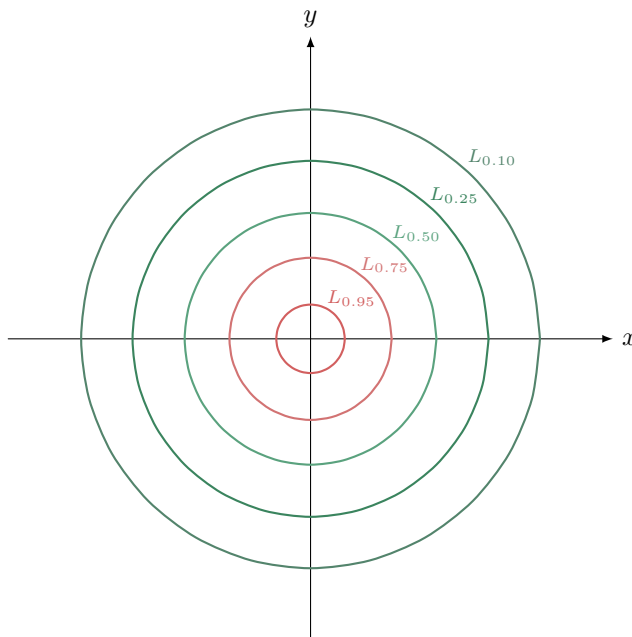
Dans toute la suite on considèrera alors souvent le cas épuré de champs de scalaires et il est attendu d'adapter les résultats élémentaires aux cas à valeurs vectorielles.

Lignes de niveaux:

On considère une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, alors on appelle **ligne de niveau** associé à la valeur λ l'ensemble suivant:

$$L_\lambda := \left\{ (x, y) \in \mathbb{R}^2 ; f(x, y) = \lambda \right\}$$

C'est une courbe plane qui représente l'ensemble des points du plan où la fonction prends une valeur fixée, tracer plusieurs lignes de niveaux permet alors d'obtenir des informations sur la fonction, par exemple:



Lignes de niveau de la fonction $f : (x, y) \mapsto \exp(-x^2 - y^2)$

Changements de variables:

L'étude des fonctions à plusieurs variables va nous amener à étudier les **changements de variables** que nous allons définir:

On appelle changement de variable tout \mathcal{C}^k difféomorphisme entre deux espaces.

En particulier, on verra que le passage en coordonnées polaires défini comme suite est bien un changement de variable et en tant que tel, il préserve les propriétés différentielles:

$$\begin{aligned} \phi : \mathbb{R}_+^* \times]-\pi ; \pi[&\longrightarrow \mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\}) \\ (r, \theta) &\longmapsto (r \cos(\theta), r \sin(\theta)) \end{aligned}$$

VII — LIMITE & CONTINUITÉ

Dans ce chapitre et comme vu dans le chapitre de Topologie, nous étudions le concept de limite dans le cas d'une fonction à plusieurs variables, plus précisément, sauf exceptions, dans le cas d'un champ de scalaire. Dans les cas où les généralités sont triviales, on utilisera sans vergogne l'exemple canonique d'une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. On rappelle qu'une limite d'une telle fonction se caractérise métriquement par la propriété:

$$\forall \varepsilon > 0, \exists \lambda > 0, \forall x \in \mathbb{R}^n, \|x - a\| < \lambda \implies |f(x) - l| < \varepsilon$$

Limites suivant une restriction:

Cette définition correspond bien à l'intuition *topologique* de la notion de limite. En particulier, on peut alors montrer que si f tends vers l en (a, b) , alors en particulier si on considère un **chemin** de la forme:

$$\gamma : t \mapsto (\gamma_1(t), \gamma_2(t))$$

Et si ce chemin est tel qu'il passe par (a, b) en t_0 , alors on a nécessairement:

$$\lim_{(x,y) \rightarrow (a,b)} f(x,y) = l \implies \lim_{t \rightarrow t_0} f(\gamma(t)) = l$$

En particulier cela nous permet de construire une méthode pour démontrer la **non-existence de limite**, ie:

Si il existe deux chemins différents suivant lesquels la fonction admet deux limites différentes, alors elle n'admet pas de limite en ce point.

Erreurs communes:

Cette définition correspond bien à l'intuition *topologique* de la notion de limite. Néanmoins il faut remarquer les propriétés suivantes spécifiques aux fonctions à plusieurs variables:

- Il est possible qu'une fonction admette une limite suivant plusieurs chemins, mais pas de limite.
- Il est même possible qu'une fonction admette une limite suivant toutes les droites, mais pas de limite.
- De manière générale, admettre une limite sur plusieurs restrictions ne permet pas d'affirmer l'existence d'une limite.

Opérations:

Comme dans le cas des fonctions réelles, on peut alors démontrer les propriétés opératoires de la limite, en effet si on considère f, g deux champs scalaires qui admettent deux limites l_1, l_2 en $A \in \mathbb{R}^p$, alors on a les propriétés suivantes:

- Le passage à la limite est **linéaire**.
- Le passage à la limite est **multiplicatif**.
- Le passage à la limite (non-nulle) d'une fonction localement non-nulle est **inversible**.

La première propriété se généralise aux champs vectoriels, mais pas les deux secondes bien évidemment, car les opérations en jeu (effectuées à l'arrivée) ne sont pas définies pour des vecteurs.

Enfin pour deux fonction qui sont composables, on a la propriété fondamentale suivante:

- Le passage à la limite est compatible avec la **composition**.

Théorème d'encadrement :

On peut généraliser le théorème des gendarmes aux fonctions à plusieurs variables, en effet si on considère f, g, h trois fonctions telles que $f(x, y) \leq g(x, y) \leq h(x, y)$ dans un voisinage de (a, b) alors on a:

$$\lim_{(x,y) \rightarrow (a,b)} f(x,y) = \lim_{(x,y) \rightarrow (a,b)} h(x,y) = l \implies \lim_{(x,y) \rightarrow (a,b)} g(x,y) = l$$

Passage en polaire :

Si on considère une fonction à deux variables, on peut exprimer (x, y) en coordonnées polaires, ie considérer la composée $f(x, y) = f(r \cos(\theta), r \sin(\theta))$ et alors f tends vers 0 en $(0, 0)$ si et seulement si il existe une fonction $g(r)$ telle que:

$$|f(r \cos(\theta), r \sin(\theta))| \leq g(r)$$

Avec $g(r)$ qui tends vers 0 quand r tends vers 0. **Attention** la fonction g ne doit plus dépendre de θ .

Continuité :

En tant que conséquence directe de la définition topologique, on dira alors qu'une fonction est **continue** en un point A si et seulement si:

$$\lim_{(x,y) \rightarrow A} f(x, y) = f(A)$$

Et toutes les propriétés opératoires usuelles sont alors aisément démontrées. En particulier, on peut montrer que toutes les fractions rationnelles sont continues sur leur ensemble de définition, et donc tous les polynômes.

VII — DÉRIVÉES DIRECTIONNELLES

Dans ce chapitre nous essaierons de construire un concept de **dérivée** dans le cas d'une fonction à plusieurs variables, plus précisément, sauf exceptions, dans le cas d'un champ de scalaire. Dans les cas où les généralités sont triviales, on utilisera sans vergogne l'exemple canonique d'une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

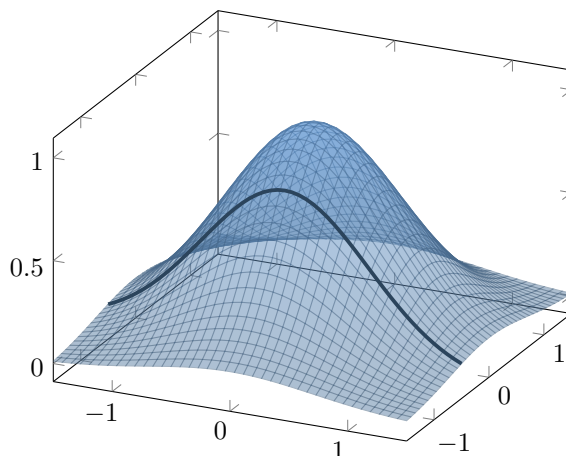
Applications Partielles :

On considère une fonction de $A \subseteq \mathbb{R}^p \rightarrow \mathbb{R}$ et un point $a = (a_1, \dots, a_n)$ **intérieur** à A . Alors il existe un voisinage de a où la fonction **réelle** suivante est bien définie, on l'appelle alors **i-ème application partielle** au point a :

$$\begin{aligned} f_{x_i} : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto f(a_1, \dots, t, \dots, a_n) \end{aligned}$$

Considérer la i-ème application partielle, c'est fixer une direction et considérer la courbe de f suivant cette direction.

Géométriquement, on peut le voir comme l'intersection d'un plan parallèle à l'axe considéré et passant par a , et l'application partielle est alors l'intersection de la courbe et du plan:



Application partielle f_x en $(0, -0.5)$ de $f : (x, y) \mapsto \exp(-x^2 - y^2)$

Dérivées Partielles :

On dira alors qu'une telle fonction **admet une i-ème dérivée partielle** au point a si et seulement si sa i-ème application partielle en a est dérivable en a_i . On se ramène alors à l'étude de la dérivabilité d'une fonction **réelle**, ce qui ne pose pas de problème théorique. Dans ce cas, on notera alors:

$$\frac{\partial f}{\partial x_i}(a) = f'_{x_i}(a_i)$$

Si f admet une i-ème dérivée partielle sur un ouvert \mathcal{U} si et seulement si elle en admet en tout points de \mathcal{U} et on définit alors:

$$\begin{aligned} \frac{\partial f}{\partial x_i} : \mathcal{U} &\longrightarrow \mathbb{R} \\ a &\longmapsto \frac{\partial f}{\partial x_i}(a_i) \end{aligned}$$

On dira alors qu'une fonction à plusieurs variables est de classe \mathcal{C}^k en un point si elle admet toutes ses dérivées partielles en ce point et qu'elles y sont continues.

Propriétés opératoires:

On peut montrer que la dérivation partielle est un **opérateur différentiel** agissant sur un espace de fonctions qui admettent des dérivées partielles, en particulier en utilisant les définitions précédentes et en notant $\mathcal{D}_i(\mathbb{R}^n)$ l'espace des fonctions admettant une i -ème dérivée partielle, alors on peut définir:

$$\begin{aligned}\frac{\partial}{\partial x_i} : \mathcal{D}_i(\mathbb{R}^n) &\longrightarrow \mathcal{F}(\mathcal{U}, \mathbb{R}) \\ f &\longmapsto \frac{\partial f}{\partial x_i}\end{aligned}$$

C'est donc cet opérateur que nous appliquons à une fonction qui admet des dérivées partielles. En particulier, l'intérêt de cette définition est alors de pouvoir affirmer les propriétés suivantes:

- La dérivation partielle est **linéaire**.
- La dérivation partielle suit **la règle du produit**.

La première propriété se généralise aux champs vectoriels, mais pas la seconde bien évidemment, car les opérations en jeu (effectuées à l'arrivée) ne sont pas définies pour des vecteurs.

Exemple: On considère les fonctions suivantes en un point (x, y) :

$$f : (x, y) \longmapsto xy \qquad g : (x, y) \longmapsto x^2y$$

Alors on a¹:

$$\frac{\partial(fg)}{\partial x}(x, y) = \frac{\partial f}{\partial x}(x, y)g(x, y) + \frac{\partial g}{\partial x}(x, y)f(x, y) = yx^2y + 2xyxy = 3x^2y^2$$

On voit bien ici que la dérivée d'un produit n'a de sens que si on est dans le cas de (au moins un) champs scalaires !

Dérivée partielle d'une composée:

Le cas de la composée est plus subtil et on revient temporairement au cas général des deux champs vectoriels suivants dont on supposera qu'ils admettent toutes leurs dérivées partielles partout:

$$g : \mathbb{R}^m \longmapsto \mathbb{R}^n \qquad f : \mathbb{R}^n \longmapsto \mathbb{R}^p$$

On cherche à calculer la dérivée partielle par rapport à la i -ème variable de la fonction $f \circ g$ au point a , on peut alors montrer la formule suivante, appelée **chain rule**:

$$\frac{\partial(f \circ g)}{\partial x_i}(a) = \sum_{k=1}^n \frac{\partial g_k}{\partial x_i}(a) \frac{\partial f}{\partial x_k}(g(a))$$

Exemple: On considère une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 et la fonction g de passage en polaire:

$$g : (r, \theta) \longmapsto (r \cos(\theta), r \sin(\theta))$$

On se propose de calculer la première dérivée partielle de $f \circ g$ au point (r, θ) , on a alors:

$$\begin{aligned}\frac{\partial(f \circ g)}{\partial r}(r, \theta) &= \frac{\partial g_1}{\partial r}(r, \theta) \frac{\partial f}{\partial x}(r \cos(\theta), r \sin(\theta)) + \frac{\partial g_2}{\partial r}(r, \theta) \frac{\partial f}{\partial y}(r \cos(\theta), r \sin(\theta)) \\ &= \cos(\theta) \frac{\partial f}{\partial x}(r \cos(\theta), r \sin(\theta)) + \sin(\theta) \frac{\partial f}{\partial y}(r \cos(\theta), r \sin(\theta))\end{aligned}$$

¹**Attention:** Ici on fait l'abus de notation usuel d'utiliser deux fois la variable x qui ont deux sens différents, l'expression $\frac{\partial fg}{\partial x}$ signifie simplement qu'on dérive par rapport à la première variable (penser $\partial_1(fg)$) et (x, y) représente le point d'évaluation de la dérivée partielle.)

Dérivées partielles d'ordre supérieur :

La i -ème dérivée partielle de f étant à nouveau une fonction de n variables, on peut alors considérer ses dérivées partielles. On définit les dérivées partielles secondes de f (si elles existent), notées:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$$

Et par récurrence les dérivées partielles d'ordre k par:

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a)$$

Et on se pose alors la question de l'importance de **l'ordre de dérivation** et de manière générale, on peut alors montrer que l'ordre d'application de l'opérateur compte, ie que:

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \neq \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$$

Mais le théorème de Schwarz donne alors une condition suffisante pour que l'ordre ne compte pas:

Si les dérivées partielles sont toutes continues en un point, alors l'ordre de dérivation n'importe pas.

Dérivées Directionnelles :

Si on considère les dérivées partielles d'une fonction f , qu'on note $a = (x_1, \dots, x_n)$ le point d'étude et (e_i) la base canonique, on peut alors les caractériser par:

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \rightarrow 0} \frac{f(a + he_i) - f(a)}{h}$$

En effet ce sont par définition les dérivées des applications partielles, donc des limites de taux d'accroissement¹ qui correspondent à la limite ci-dessus.

Ce sont en fait des **dérivées directionnelles** dans des directions canoniques, mais on souhaite alors généraliser cette notion à tout vecteur unitaire v , on pose alors simplement:

$$D_v f(a) = \lim_{h \rightarrow 0} \frac{f(a + hv) - f(a)}{h}$$

C'est alors intuitivement la dérivée de la courbe obtenue par intersection avec le plan vertical suivant le vecteur choisi.

Inconvénients de cette définition:

On aimerait alors que cette définition remplisse les memes roles que ceux remplis par la dérivée usuelle, mais ce n'est pas du tout le cas, en particulier, on a les problèmes suivants:

- Un fonction qui admet des dérivées partielles en un point peut ne pas etre continue en ce point.
- Un fonction qui admet toute ses dérivées directionnelles en un point peut ne pas etre continue en ce point.
- Elles ne permettent pas directement de définir une application affine tangente à la courbe.
- Elles ne permettent pas d'étudier les extrema de la fonction.

En résumé, le concept de dérivée directionnelle manque de puissance.

¹Cette limite n'étant pas très pratique pour des calculs concrets, on verra au chapitre suivant une caractérisation plus utile dans le cas de fonctions différentiables.

VII — DIFFÉRENTIELLE

Nous allons à présent définir le vrai concept équivalent à la dérivée usuelle, qui nous permettra de généraliser l'idée d'application tangente en un point, d'étudier les extrema d'une fonction et d'étudier toutes les propriétés différentielles d'une fonction à plusieurs variables de manière générale.

Dans toute la suite, on s'appropriera la notation suivante des vecteurs de la **base duale**¹ de \mathbb{R}^n :

$$dx_i : (x_1, \dots, x_n) \mapsto x_i$$

Cas réel:

On rappelle qu'on dit qu'une fonction réelle f est dérivable en a de nombre dérivé $f'(a)$ si et seulement si²:

$$\forall h \in \mathbb{R} ; \frac{f(a+h) - f(a) - f'(a)h}{h} \xrightarrow{h \rightarrow 0} 0$$

Cette définition revient alors à dire que l'application linéaire qui approxime le mieux f en a est donnée par:

$$\begin{aligned} df_a : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto f'(a)dx(x) = f'(a)x \end{aligned}$$

Ce qui correspond à dire que la fonction $df_a(x) = f'(a)x$ est l'**application linéaire tangente** à f au point a . Attention, ici on dit bien linéaire, et donc ce n'est pas une tangente géométrique à la courbe, mais son pendant affine l'est.

On définit alors la **différentielle** de f comme la fonction qui à tout point où f est dérivable, lui associe son application linéaire tangente, ie:

$$\begin{aligned} df : \mathbb{R} &\longrightarrow \mathcal{L}(\mathbb{R}, \mathbb{R}) \\ a &\longmapsto f'(a)dx \end{aligned}$$

C'est cette notion de différentiabilité et d'application linéaire tangente qui va se généraliser dans tout sa force au cas général.

Définition:

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$, on dira que f est **différentiable** en un point a si et seulement si il existe une application linéaire $df_a : \mathbb{R}^n \rightarrow \mathbb{R}^p$ telle que:

$$\forall h \in \mathbb{R}^n ; \frac{f(a+h) - f(a) - df_a(h)}{\|h\|} \xrightarrow{h \rightarrow 0} 0$$

Ou en d'autres termes qu'il existe ε_a une fonction définie sur un voisinage de 0, telle que $\varepsilon_a(h) \xrightarrow{h \rightarrow 0} 0$ et que:

$$\forall h \in \mathbb{R}^n ; f(a+h) - f(a) - df_a(h) = \|h\| \varepsilon_a(h)$$

En outre, si f est différentiable sur un ouvert U , alors on définit sa différentielle de la même manière que dans le cas réel par:

$$\begin{aligned} df : \mathbb{R}^n &\longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) \\ a &\longmapsto df_a \end{aligned}$$

Mais cette définition ne nous permet pas de connaître la **forme** de la différentielle, et donc de la calculer, on doit donc chercher des conditions nécessaires sur cette fonction dans la section suivante.

¹Formellement, il s'agit d'un **champ de formes 1-linéaire** (on dira aussi **1-forme différentielle**) qui à chaque point a de \mathbb{R}^n associe la projection canonique sur la i -ème coordonnée de "l'espace tangent" en a , mais dans ce cas, $T\mathbb{R}^n = \mathbb{R}^n$, donc le champ de forme linéaire se réduit simplement en une simple forme linéaire (ou plutôt un champ constant).

²Ce qui est équivalent à dire que $f(a+h) - f(a) - f'(a)h = o(|h|)$ au voisinage de 0

Forme nécessaire de la différentielle:

En exprimant les applications partielles de f et en utilisant la définition de la différentielle appliqué au quotient obtenu, on montre tout d'abord que f admet nécessairement toutes ses dérivées partielles en a et on a:

$$\frac{\partial f}{\partial x_i}(a) = df_a(e_i)$$

Alors, en exprimant un vecteur $x \in \mathbb{R}^n$ dans la base canonique et en utilisant la linéarité de la différentielle, on obtient l'expression:

$$df_a(x) = \sum_{k=1}^n x_k df_a(e_k) = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(a) x_k = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(a) dx_k$$

Finalement, on a bien donné une forme nécessaire à la différentielle de f et on a:

$$\begin{aligned} df : \mathbb{R}^n &\longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) \\ a &\longmapsto \sum_{k=1}^n \frac{\partial f}{\partial x_k}(a) dx_k \end{aligned}$$

C'est bien une application linéaire à valeurs dans \mathbb{R}^p car par composantes $\frac{\partial f}{\partial x_i}(a)$ est un vecteur et le produit $vd x$ est bien défini car dx est une forme linéaire.

Exemple: On considère la fonction $f : (x, y) \mapsto (xy, x + y)$, dont on souhaite calculer la différentielle en $a = (1, 1)$ alors on a:

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = (y, 1) \\ \frac{\partial f}{\partial y}(x, y) = (x, 1) \end{cases}$$

Donc la différentielle est donnée par:

$$df : (a, b) \mapsto (b, 1)dx + (a, 1)dy$$

Et donc la différentielle en a est donnée par:

$$df_a : (x, y) \mapsto (1, 1)x + (1, 1)y = (x + y, x + y)$$

Propriétés opératoires:

Tout d'abord, on peut montrer que l'**opérateur de différentiation** sur les applications différentiables sur un ouvert \mathcal{U} , qu'on notera $\mathcal{D}_{\mathcal{U}}$, est un opérateur différentiel:

$$\begin{aligned} d : \mathcal{D}_{\mathcal{U}} &\longrightarrow \mathcal{F}(\mathcal{U}, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)) \\ f &\longmapsto df \end{aligned}$$

C'est donc cet opérateur que nous appliquons à une fonction qui admet une différentielle. En particulier, l'intérêt de cette définition est alors de pouvoir affirmer les propriétés suivantes:

- La différentiation est **linéaire**.
- La différentiation suit **la règle du produit**.

La première propriété se généralise aux champs vectoriels, mais pas la seconde bien évidemment, car les opérations en jeu (effectuées à l'arrivée) ne sont pas définies pour des vecteurs.

Différentielle d'une composée:

On a aussi pour $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ différentiable en a et $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ différentiable en $f(a)$ la formule de différentielle d'une composée:

$$d(g \circ f)_a = dg_{f(a)} \circ df_a$$

Une autre interprétation de cette propriété sera possible plus tard via l'introduction de la Jacobienne.

Différentielle d'une réciproque:

Si on a $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ un **homéomorphisme** différentiable en a et si df_a est bijective alors f^{-1} est différentiable en $f(a)$ et on a:

$$d(f^{-1})_{f(a)} = (d(f)_a)^{-1}$$

Une autre interprétation de cette propriété sera possible plus tard via l'introduction de la Jacobienne.

Propriétés de régularité:

Nous nous devons maintenant de confirmer que cette définition est bien celle qui étends l'idée de dérivabilité classique. Et en effet c'est le cas et on a les propriétés suivantes:

Si une application est différentiable, elle est continue.

La réciproque est bien évidemment fausse.

Condition nécessaire de différentiabilité:

On aimerait maintenant réussir à trouver une condition nécessaire de différentiabilité, en effet, étudier la limite donnée par la définition est pénible et souvent complexe.

On montre alors la propriété fondamentale suivante:

Si une application est de classe \mathcal{C}^1 , alors elle est différentiable.

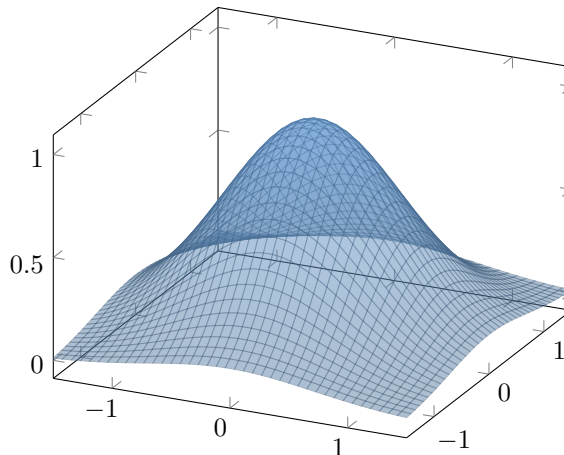
Ce qui nous permet simplement de calculer les dérivées partielles de f et d'étudier leur continuité pour en déduire la différentiabilité (dans les cas simples). Attention, ce n'est qu'une condition **suffisante**.

Plan Tangent:

Si on revient à des problématiques plus pratiques, et notamment dans le cas d'une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, alors on sait que la différentielle est l'application linéaire tangente au point a , et on aimerait avoir une expression de l'application **affine** tangente en (a, b) , et de manière analogue au cas réel, on peut définir:

$$\tau_{(a,b)} := \left\{ (x, y, z) \in \mathbb{R}^3 ; z = df_a(x - a, y - b) + f(a, b) \right\}$$

C'est le **plan affine tangent** à la courbe de f au point (a, b) .



Graphes de la fonction $f : (x, y) \mapsto \exp(-x^2 - y^2)$

On définit de même l'**espace affine tangent** à la courbe de f au point (a_1, \dots, a_n) par:

$$\tau_{(a_1, \dots, a_n)} := \left\{ (x_1, \dots, x_n, z) \in \mathbb{R}^{n+1} ; z = df_a(x_1 - a_1, \dots, x_n - a_n) + f(a_1, \dots, a_n) \right\}$$

Gradient:

On va maintenant définir un autre opérateur différentiel sur l'ensemble des fonctions de classe $\mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$, attention, ici il s'agit bien du cas particulier des champs de scalaires. On définit alors **l'opérateur gradient**:

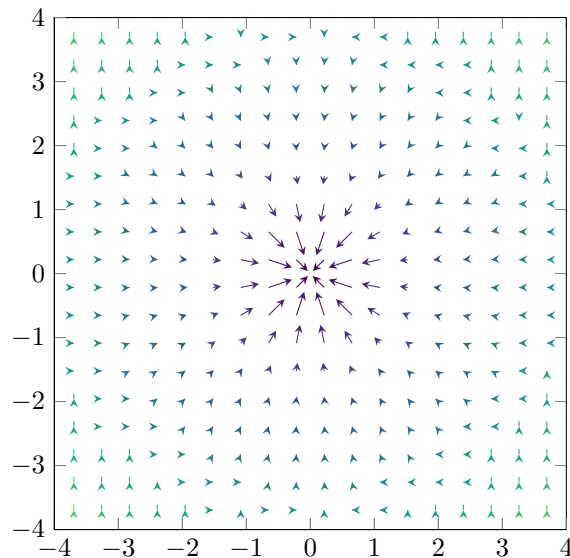
$$\begin{aligned}\nabla : \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}) &\longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \\ f &\longmapsto \nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)\end{aligned}$$

Le gradient de f est donc un **champ de vecteurs**, dont on verra une interprétation ci-dessous, mais tout d'abord, on peut remarquer l'identité suivante en un point a , qui le lie avec la différentielle:

$$\forall h \in \mathbb{R}^n ; \langle \nabla f(a) | h \rangle = \left\langle \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) | h \right\rangle = df_a(h)$$

On cherche maintenant à interpréter ce que représente le gradient par rapport au champs scalaire f , et on peut alors montrer la propriété géométrique suivante:

Le gradient est le champ de vecteurs qui donne la direction de la plus grand pente ascendante à la surface de la courbe de f .



Gradient de la fonction $f : (x, y) \mapsto \exp(-x^2 - y^2)$

Une autre propriété remarquable du gradient est qu'il est toujours **orthogonal** aux lignes de niveaux.

Caractérisation des dérivées directionnelles:

Pour un vecteur unitaire v donné, on peut alors montrer la caractérisation très utile de la dérivée directionnelle de f dans la direction de v , en effet on a:

$$D_v f(a) = df_a(v) = \langle \nabla f(a) | v \rangle$$

En effet, il suffit alors d'écrire la formule de Taylor à l'ordre 1 sur la quantité ci-dessous et utiliser les différentes propriétés des objets considérés:

$$f(a + hv) - f(a) = df_a(hv) + o(\|hv\|)$$

Exemple: Pour $f(x, y) = xy + x$ et $v = \frac{1}{\sqrt{2}}(1, 1)$ alors on a:

$$df_{(x,y)} = (y+1)dx + xdy$$

Et donc finalement:

$$D_v f(x, y) = \frac{1}{\sqrt{2}}(y + x + 1)$$

On peut remarque que si on prends $v = e_i$, on retombe évidemment sur les dérivées partielles.

Jacobienne:

On considère maintenant plus généralement l'espace des fonctions de classe $\mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^p)$, et on définit alors l'opérateur suivant:

$$J : \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^p) \longrightarrow \mathcal{M}_{p,n}(\mathcal{F}(\mathbb{R}^n, \mathbb{R}^p))$$
$$f \longmapsto \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_n} \end{pmatrix}$$

On appellera alors **jacobienne** de f la matrice Jf , et on remarque que c'est simplement la matrice¹ constituée des dérivées partielles de f .

En particulier, on a donc l'identité suivante:

$$Jf = \begin{pmatrix} \nabla f_1 \\ \vdots \\ \nabla f_n \end{pmatrix}$$

C'est aussi la matrice des gradients des applications partielles de f .

Si on l'évalue maintenant en un point $a \in \mathbb{R}^n$, on obtient alors une matrice réelle et on a la caractérisation fondamentale suivante:

La jacobienne prise en un point est la matrice de la différentielle en ce point.

En particulier on retrouve donc la règle de la composition de la différentielle vu sous l'angle matriciel:

$$J(f \circ g)_a = Jf_{g(a)} \times Jg_{g(a)}$$

Ainsi que la règle de l'inversion d'une fonction inversible:

$$(Jf_a)^{-1} = J(f^{-1})_{f(a)}$$

¹A coefficients dans l'anneau des fonctions.

VII — THÉORÈMES GÉNÉRAUX

Dans cette section, nous allons utiliser toutes les notions construites au chapitre précédent pour essayer de généraliser les grands théorèmes sur les fonctions réelles dérivables, en particulier les deux théorèmes suivants:

- Le théorème de la bijection
- Le théorèmes des accroissements finis

Le problème principal pour obtenir un analogue au premier théorème est le suivant:

Il existe des fonctions non-injectives telles qu'en tout point leur différentielle soit inversible !

Ce qui implique alors que nous ne pouvons pas prouver l'injectivité **globale** d'une fonctions à plusieurs variables avec le critère simple de la non-annulation de la différentielle. Ceci nous menera donc à la recherche de théorème permettant de montrer la bijectivité de ces fonctions autrement.

Théorème d'inversion locale:

On considère $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ de classe \mathcal{C}^1 et un point $a \in \mathbb{R}^n$, alors si sa différentielle est **inversible** en a , on a:

La fonction induit un \mathcal{C}^1 -difféomorphisme d'un voisinage V_a dans $f(V_a)$.

On peut interpréter ce théorème de manière plus naturelle sous la forme:

Si la différentielle ne s'annule pas en un point, alors la fonction est localement inversible en ce point.

Théorème d'inversion globale:

On considère $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ une fonction **injective** et de classe \mathcal{C}^1 , alors si en tout point d'un ouvert \mathcal{U} sa différentielle est **inversible**, on a:

La fonction induit un \mathcal{C}^1 -difféomorphisme de \mathcal{U} sur $f(\mathcal{U})$.

On peut interpréter ce théorème de manière plus naturelle sous la forme:

Sous réserve d'injectivité, si la différentielle ne s'annule pas sur un ouvert, alors la fonction est inversible sur cet ouvert.

Théorème des fonctions implicites:

On considère une fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1(\mathbb{R}^n)$ et on s'intéresse aux courbes Γ d'équation:

$$f(x_1, \dots, x_n) = 0$$

Si il est possible d'exprimer une variable en fonction des autres, ie d'isoler une variable x_i , alors on dira que l'équation E définit x_i comme **une fonction implicite** des $n - 1$ autres variables. On s'intéresse ici à des condition suffisantes sur f pour que de telles fonctions implicites existent.

On peut alors montrer grâce au **théorème d'inversion locale** que si $(x_1, \dots, x_n) \in \Gamma$, et que la i -ème dérivée partielle ne s'annule pas, alors **localement** l'équation définit bien x_i comme fonction implicite des $n - 1$ autres variables pour un certaine fonction ϕ de classe $\mathcal{C}^1(\mathbb{R}^{n-1})$, ie on a:

$$x_i = \phi(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$$

Prenons un exemple simple pour illustrer le propos

On considère la courbe Γ suivante:

$$f(x, y) = x^2 + y^2 - 1 = 0$$

On reconnaît immédiatement le cercle unité, le théorème des fonctions implicites affirme alors qu'à l'exception des points $(\pm 1, 0)$, cette équation définit implicitement y comme une fonction de x et en effet, on sait que par exemple dans le demi-plan supérieur, on a:

$$y = \sqrt{1 - x^2}$$

Ce théorème justifie une technique courante dans la littérature anglo-saxonne nommé **différentiation implicite**, en effet si on considère une équation d'une fonction à deux variables (par exemple), alors on peut "voir"¹ une des variables comme fonction de l'autre et simplement différentier, et par exemple on peut de cette manière trouver les dérivées des fonctions réciproques:

$$\begin{aligned} y = \cos^{-1}(x) &\iff \cos(y) = x \\ &\iff \sin(y)y' = -1 && \text{(On "voit" } y \text{ comme fonction de } x \text{ et on différentie.)} \\ &\iff y' = \frac{-1}{\sin(y)} \\ &\iff y' = \frac{-1}{\sin(\cos^{-1}(x))} \\ &\iff y' = \frac{1}{-\sqrt{1 - x^2}} \end{aligned}$$

Théorème des accroissements finis:

On considère ici un champs scalaire² défini et différentiable sur un ouvert \mathcal{U} , alors pour tout point $a, b \in \mathcal{U}$ tel que le segment $[a, b]$ les reliant soit dans \mathcal{U} , on a³:

$$\exists c \in \mathcal{U} ; f(b) - f(a) = df_c(b - a)$$

En particulier si \mathcal{U} est convexe, cette propriété est vrai pour tout couple (a, b) .

Corollaire 1 : Inégalité des accroissements finis:

En particulier, si la différentielle df_x est **bornée** par un réel M sur \mathcal{U} , alors on a aussi:

$$|f(b) - f(a)| \leq M(b - a)$$

Cette majoration reste vraie, contrairement au théorème, dans le cas de champs de vecteurs et on a alors pour des normes bien choisies:

$$\|f(b) - f(a)\| \leq M \|b - a\|$$

En particulier, cela montre donc que les applications différentiables sont toutes **localement** lipschitziennes, et même globalement lipschitziennes si la différentielle est bornée.

Corollaire 2 : Applications constantes :

On peut alors généraliser le critère des fonctions réelles qui nous affirme que si la dérivée d'une fonction est nulle, elle est constante. En effet, si $f : \mathcal{U} \rightarrow \mathbb{R}$ différentiable sur \mathcal{U} ouvert et **connexe**⁴, alors:

Si la différentielle de f est **nulle** sur \mathcal{U} , alors f est **constante** sur \mathcal{U} .

¹Sous les hypothèses de régularité du théorème bien évidemment.

²Il n'existe pas d'équivalent au théorème des accroissements finis dans le cas de fonctions à valeurs vectorielles, considérer par exemple la courbe $t \mapsto (\cos(t), \sin(t))$.

³La démonstration de cette propriété revient à se ramener au cas réel en posant $g(t) = f((1 - t)a + tb)$

⁴La preuve est directe si \mathcal{U} est convexe, pour passer à la connexité, il faut alors remarque que toutes les boules sont convexes.

Corollaire 3 : Applications constantes par rapport à une variable :

On s'intéresse alors à ce qu'on peut déduire de l'annulation d'une seule des dérivées partielles, et en se ramenant à l'étude des application partielles on peut alors montrer que si $f : \mathcal{U} \rightarrow \mathbb{R}$ différentiable sur \mathcal{U} ouvert et **convexe**² et que $\frac{\partial f}{\partial x_i}$ est **nulle** sur \mathcal{U} , alors il existe une fonction g telle que:

$$\forall (x_1, \dots, x_n) \in \mathcal{U} \quad f(x_1, \dots, x_n) = g(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$$

Intuitivement, on l'interprète simplement par le fait que f ne dépend pas de la variable x_i sur cet ouvert.

C'est ce corollaire qui nous permettra de résoudre nos premières équations aux dérivées partielles. Voici un exemple élémentaire, on considère une fonction f de classe $\mathcal{C}^1(\mathbb{R}^2)$ qui vérifie:

$$\frac{\partial f}{\partial x}(x, y) = 0$$

Alors d'après ce corollaire, on a directement, pour $g \in \mathcal{C}^1(\mathbb{R})$ une certaine fonction réelle:

$$f(x, y) = g(y)$$

²**ATTENTION:** Ce résultat est faux sur un domaine juste connexe. Prendre la fonction qui définie sur \mathbb{R}^2 privée d'un axe bien choisi qui vaut 0 sur $\{(x, y), xy < 0\}$ et x^2 ailleurs.

VII — OPTIMISATION I

Un des intérêts de la notion de dérivée d'une fonction réelle est de pouvoir étudier les extrema de cette fonction, en particulier, un point capital du cours d'analyse réelle est la définition de **point critique**, qui sont les points d'annulation de la dérivée, et on a alors la condition nécessaire suivante:

Si un point est un extremum, alors c'est un point critique.

On veut généraliser cette notion au cas des **champs scalaires**, pour à terme être capable de trouver les extrema de telles fonctions.

Points critiques:

De manière parfaitement analogue, si f est une fonction définie sur un ouvert \mathcal{U} , alors on définit les **points critiques**, qui sont les points d'annulation de la différentielle, et on a alors la condition nécessaire suivante:

Si un point est un extremum, alors c'est un point critique.

Informellement, on comprends naturellement que les seuls extrema possibles seront ceux où l'application linéaire tangente est nulle, donc un plan horizontal. Néanmoins la condition n'est que nécessaire, en effet on considère la fonction:

$$f : (x, y) \mapsto xy$$

Alors sa différentielle est nulle en $(0, 0)$, mais ce point n'est pas un extremum, il suffit d'étudier quelques restrictions pour le montrer. On a donc besoin d'outils plus puissants pour trouver ces extrema.

Hessienne:

On suppose maintenant que $f \in \mathcal{C}^2(\mathbb{R}^n)$, alors elle admet toutes ses dérivées partielles secondes et elle vérifie le **théorème de Schwarz**. On définit alors la **matrice Hessienne** de f par:

$$\mathcal{H}f_{(x,y)} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j \in \llbracket 1 ; n \rrbracket} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Cette matrice est **symétrique** d'après le théorème de Schwarz, elle définit donc une forme bilinéaire symétrique et une forme quadratique, qu'on appellera forme hessienne, qui jouera un rôle important pour la suite.

Classification des points critiques:

Dans le cas réel, pour classer les points critiques d'une fonction, on étudiait la convexité (locale) de cette fonction, et en particulier le signe de la dérivée seconde, ceci permettait de différencier les maxima, minima et point cols (ou point d'inflexions). D'une certaine manière, on généralise cette approche au cas général en étudiant le **signe de la Hessienne** au sens du signe d'une forme quadratique, et on a les cas suivants:

- Si la forme hessienne n'est **ni positive ni négative** en un point critique, alors le point est un **point col**.
- Si la forme hessienne est **définie positive** en un point critique, alors la fonction est localement convexe et le point est un **minimum**.
- Si la forme hessienne est **définie négative** en un point critique, alors la fonction est localement concave et le point est un **maximum**.
- Sinon la méthode ne permet pas de conclure.

On se ramène donc à l'étude des valeurs propres de la Hessienne, et de leurs signes ou dit autrement, de la signature de la forme quadratique Hessienne. Ceci s'interprète en disant que, de manière analogue au cas réel, le signe de la hessienne caractérise la **convexité locale** de la fonction.

VII — OPTIMISATION II

Pour aller plus loin, on peut s'intéresser à essayer d'optimiser une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sous la contrainte que $x \in \{x \in \mathbb{R}^n ; g(x) = 0\}$ où f, g sont lisses et où g satisfait les conditions du théorème des fonctions implicites (ie que dg soit de rang maximal, partout pour simplifier). En d'autres termes on veut résoudre le problème d'optimisation suivant:

$$(P) : \begin{cases} \sup f(x) \\ x \in C := \{x \in \mathbb{R}^n ; g(x) = 0\} \end{cases}$$

On cherche alors une méthode permettant de résoudre un tel problème, et plus particulièrement une méthode permettant de passer de ce problème à un **problème d'optimisation sans contraintes**.

Conditions nécessaires :

On considère un point $x \in \mathbb{R}^n$ qui est solution du problème, et v un vecteur **tangent** à $\{x \in \mathbb{R}^n ; g(x) = 0\}$, alors on peut montrer que nécessairement:

$$df_x(v) = 0 \iff \langle \nabla f(x) | v \rangle = 0$$

Ce qui s'interprète alors par le fait que ∇f_x est **orthogonal** à l'espace tangent à $\{x \in \mathbb{R}^n ; g(x) = 0\}$, or celui-ci (par le théorème des fonctions implicites) peut localement s'écrire comme:

$$\{x \in \mathbb{R}^n ; \langle \nabla g(x) | x \rangle = 0\} = \{\nabla g(x)\}^\perp$$

Finalement on trouve alors que f est orthogonale à un orthogonal et donc colinéaire au **gradient de la contrainte**, ie on a:

$$\exists \lambda \in \mathbb{R} ; \nabla f(x) = \lambda \nabla g(x)$$

Plus généralement, si on a m contraintes données par des fonction g_i , on a;

$$\exists (\lambda_i) \in \mathbb{R}^m ; \nabla f(x) = \sum_{i=1}^m \lambda_i \nabla g_i(x)$$

Lagrangien :

On encapsule alors toutes les contraintes du problème en posant la fonction suivante appelée **lagrangien** du problème:

$$\mathcal{L}(x, \lambda) = f(x) + \langle \lambda | g(x) \rangle$$

Alors c'est une fonction lisse de ses variables et si x est solution du problème, on montre que l'on a nécessairement:

$$\nabla \mathcal{L}(x, \lambda) = 0$$

Cette fonction nous donne alors une condition **nécessaire** que doit vérifier un extremum sous contrainte.

Méthode d'optimisation :

On peut donc construire une méthode qui permet de résoudre de tels problèmes. En effet, on peut appliquer l'algorithme suivant:

- On cherche à montrer l'**existence** d'un extremum sous contraintes (par exemple par compacité de C).
- On calcule les **points critiques du Lagrangien** qui sont alors tous les **candidats** possibles.

VII — FORMES DIFFÉRENTIELLES

VII — INTÉGRALES CURVILIGNES

Dans cette partie, nous allons généraliser l'intégrale de Riemann, c'est à dire l'intégrale d'une fonction le long d'un segment, aux **champs scalaires et vectoriels**, avec en tête l'idée suivante:

On veut pouvoir intégrer un champ le long d'une courbe lisse donnée.

Nous verrons alors qu'il existe en fait trois constructions différentes qui mesurent différents phénomènes sur les champs considérés:

- L'intégrale d'un **champ scalaire** qui quantifie l'aire entre la courbe et le plan des paramètres.
- L'intégrale de **travail** d'un **champ vectoriel** qui quantifie la contribution du champ vectoriel au trajet de la courbe.
- L'intégrale de **flux** d'un **champ vectoriel** qui quantifie la tendance au champ de vecteurs à traverser la courbe.

Dans toute la suite, on considèrera un **arc orienté** Γ paramétré par une fonction $\gamma : t \mapsto \gamma(t)$ définie sur $[a; b]$ et de classe \mathcal{C}^1 . On rappelle à toutes fins utiles:

- L'abscisse curviligne de γ est donnée par: $ds = \|\gamma'(t)\| dt$
- Le vecteur normal à γ' est donnée par: $dn = (\gamma'_2(t), -\gamma'_1(t))dt$

Intégrale d'un champ scalaire:

On se donne un champ scalaire lisse $f : \Gamma \rightarrow \mathbb{R}$ et on cherche à quantifier l'aire située entre la courbe et le graphe \mathcal{G}_f , on peut alors écrire cette aire comme une somme de Riemann, qui tends alors vers une intégrale et on définit:

$$\int_{\gamma} f ds = \int_a^b f(\gamma(t)) \|\gamma'(t)\| dt$$

Intégrale de travail d'un champ vectoriel:

On se donne un champ vectoriel lisse $f : \Gamma \rightarrow \mathbb{R}^n$ et on cherche à quantifier la **contribution du champ vectoriel au trajet de la courbe**, ie au travail que le champ effectuerais sur une particule qui se déplacerait le long de la courbe, à nouveau grâce aux sommes de Riemann, on peut définir:

$$\int_{\gamma} F d\gamma = \int_a^b \langle F(\gamma(t)) | \gamma'(t) \rangle dt$$

Intégrale de flux d'un champ vectoriel:

On se donne un champ vectoriel lisse $f : \Gamma \rightarrow \mathbb{R}^n$ et on cherche à quantifier la **tendance au champ de vecteurs à traverser la courbe**, ie si le champ de vecteurs représenterais un fluide, à la quantité d'eau qui traverse la courbe en un instant donné. A nouveau grâce aux sommes de Riemann, on peut définir:

$$\int_{\gamma} F dn = \int_a^b \langle F(\gamma(t)) | n'(t) \rangle dt$$

Interprétation du produit scalaire:

On remarque que les intégrales curviligne d'un champ de vecteur font apparaître un produit scalaire, en effet cela découle directement des quantités géométriques que l'on cherche à calculer:

- Dans le cas d'une intégrale de travail, on considère la projection d'un vecteur du champ sur un vecteur tangent, c'est **la composante tangentielle** du champ, et on somme toutes ses composantes pour obtenir une contribution totale.
- Dans le cas d'une intégrale de flux, on considère la projection d'un vecteur du champ sur un vecteur normal, c'est **la composante normale** du champ, et on somme toutes ses composantes pour obtenir une contribution totale mais dans un autre sens.

Théorème du gradient:

On s'attarde un peu sur le cas vectoriel, et on se donne un champ vectoriel X tel que $X = \nabla F$ pour un certain champ scalaire F , on dira alors que X **dérive du potentiel** F , on se donne un chemin paramétré par γ sur $[a; b]$ et on cherche à calculer le travail de X sur ce chemin, on trouve alors que:

$$\begin{aligned}\int_{\gamma} X d\gamma &:= \int_a^b \langle \nabla F(\gamma(t)) | \gamma'(t) \rangle dt \\ &= \int_a^b dF_{\gamma(t)}(\gamma'(t)) dt && \text{(On utilise le lien gradient / différentielle.)} \\ &= \int_a^b d(F \circ \gamma)_t && \text{(Par la règle de la chaîne, ou en passant en composantes.)} \\ &= \int_a^b (F \circ \gamma)'(t) dt && \text{(Par le lien différentielle / dérivée.)} \\ &= F(\gamma(b)) - F(\gamma(a)) && \text{(Théorème fondamental de l'analyse.)}\end{aligned}$$

En particulier, on a montré le théorème suivant, généralisation du théorème fondamental de l'analyse, appelé souvent **théorème du gradient**:

Le travail le long d'un chemin d'un champ qui dérive d'un potentiel ne dépend que des bords du chemin choisi.

C'est en fait même un cas particulier d'un théorème très général qu'on verra plus tard appelé **théorème de Stokes** qui affirme que pour une forme différentielle **exacte**, ie qui s'écrit $d\omega$ pour ω une autre forme différentielle, alors on a que l'intégrale ne dépend que des bords:

$$\int_A d\omega = \int_{\partial A} \omega$$

Généralisations:

On imagine alors qu'il doit être possible de généraliser ces concepts d'intégrales sur une courbe (espace de dimension 1) à des intégrales sur des **surfaces, volumes ...**. On aurait alors les deux¹ concepts géométriques suivants:

- L'intégrale sur une surface d'un **champ scalaire** qui quantifierait le volume entre la courbe et le plan des paramètres.
- L'intégrale sur une surface d'un **champ vectoriel** qui quantifierait la tendance au champ de vecteurs à traverser la surface.

¹Le concept d'intégrale de travail ne se généralise pas conceptuellement aux dimensions plus grandes pour des raisons évidentes.

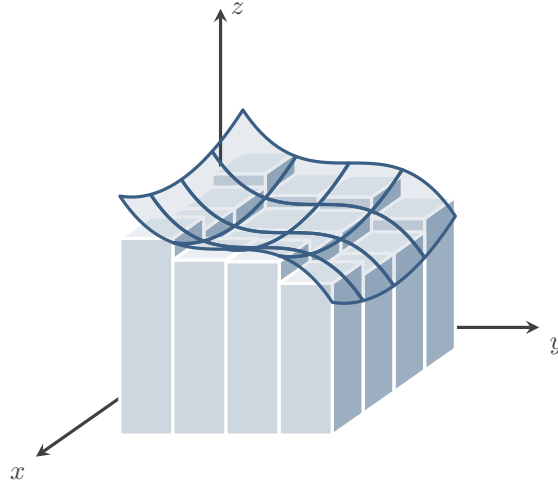
VII — INTÉGRALES MULTIPLES

Dans ce chapitre on cherche à généraliser la notion d'intégrale d'une fonction d'une seule variable sur un segment à la notion d'intégrale d'une fonction réelle **de plusieurs variables qu'on supposera bornée**, ici on considérera principalement $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ pour simplifier l'exposition. On a alors deux nouveaux concepts d'intégrales qui se présentent:

- La notion d'intégrale **curviligne** qui se propose d'intégrer à nouveau sur un chemin (donc toujours sur un espace de dimension 1).
- La notion d'intégrale **multiple** qui se propose d'intégrer sur une surface (donc sur un espace de dimension plus grande).

Ces deux notions différentes coexistent et permettent de résoudre des problèmes très différents, nous nous proposons ici tout d'abord de définir la notion d'intégrale multiple. Moralement, l'idée consiste à se donner un domaine raisonnable \mathcal{D} de \mathbb{R}^2 , et de chercher à calculer le **volume** au dessus \mathcal{D} et contenu sous la nappe définie par l'image de f .

Graphiquement, voilà ce que l'on cherche à réaliser, ici le domaine d'intégration étant un simple rectangle:



Outils préliminaires:

On souhaite intégrer f sur le rectangle $R = [a, b] \times [c, d]$, on considère alors une **subdivision** de R c'est à dire un couple de subdivisions de $[a, b]$ et $[c, d]$ respectivement, ie on se donne $(a = t_0, t_1, \dots, t_{n-1}, t_n = b)$ et $(c = t'_0, t'_1, \dots, t'_{n-1}, t'_n = d)$, alors ces subdivisions définissent des **sous-rectangles** par:

$$R_{ij} = [t_i, t_{i+1}] \times [t'_j, t'_{j+1}]$$

On appellera la collection de ces sous-rectangles une **subdivision** du rectangle R . On définit de manière évidente le **volume** (ici une aire) d'un rectangle R par:

$$\text{Vol}(R) = (b - a)(d - c)$$

Finalement on définit aussi les deux quantités suivantes pour tout rectangle R :

$$\begin{cases} m_R(f) &:= \{\inf(f(x)); x \in R\} \\ M_R(f) &:= \{\sup(f(x)); x \in R\} \end{cases}$$

Intégrale sur un rectangle:

Soit R le rectangle domaine d'intégration et P une subdivision de R , alors on définit les **sommes de Darboux supérieure et inférieures** par:

$$\begin{cases} S_P^+(f) = \sum_R M_R(f) \cdot \text{Vol}(R) \\ S_P^-(f) = \sum_R m_R(f) \cdot \text{Vol}(R) \end{cases}$$

Où la somme parcourt tout les rectangles de la subdivision. On définit alors les **intégrales de Darboux supérieure et inférieures** par:

$$\begin{cases} \int_R^+ f = \inf \left\{ S_P^+(f) ; P \text{ est une subdivision de } R \right\} \\ \int_R^- f = \sup \left\{ S_P^-(f) ; P \text{ est une subdivision de } R \right\} \end{cases}$$

*L'intégrale supérieure (resp. inférieure) est la somme **minimale** (resp. **maximale**) obtenue en considérant **toutes les subdivisions possibles** du rectangle.*

Enfin on dit que f est **intégrable** sur R si et seulement si l'**intégrale supérieure et inférieure** sont égales, et on la note formellement:

$$\int_R f dx_1 \dots dx_n \stackrel{\text{def.}}{=} \int_R^+ f$$

Intégrale sur un ensemble borné:

.... -> Admettons pour le moment que l'on sait intégrer sur des surfaces ...