16

Calcul différentiel et optimisation

L'objectif de ce chapitre est l'étude de propriétés généralisant la dérivation, dans le cadre de fonctions définies sur un e.v.n. de dimension finie E.
En MPSI, vous avez par exemple étudié le cas de fonctions de 2 variables, c'est-à-dire de fonctions $f : \Omega \to \mathbb{K}$ où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et Ω est un ouvert de $\mathbb{R}^2$. Tout élément de Ω peut être décrit dans la base canonique sous forme d'un couple $(x, y) \in \Omega$, ce qui permet de considérer une telle fonction comme une fonction de 2 variables réelles x et y. On a notamment vu dans ce cadre comment définir des dérivées partielles, et plus généralement des dérivées selon un vecteur.
On comprend facilement comment généraliser pour des fonctions de n variables, c'est à dire définies sur un ouvert Ω de $\mathbb{R}^n$ et à valeurs dans $\mathbb{K}$, ou plus généralement, à valeurs dans un $\mathbb{K}$-e.v.n. F.
De façon encore plus générale, on peut se demander si on peut s'affranchir de ce cadre très rigide de $\mathbb{R}^n$ pour la variable de départ, c'est-à-dire si on peut directement considérer des fonctions de Ω dans F, où Ω est un ouvert d'un $\mathbb{R}$-e.v.n. E, et F est un $\mathbb{K}$-e.v.n..
La notion de dérivée selon un vecteur se définit alors de même, et, si on est en dimension finie, est indépendante du choix des normes. Les dérivées partielles correspondent à des dérivées selon certains vecteurs particuliers formant une base de E. Les dérivées partielles peuvent donc se définir de façon plus générale en référence à une base de E.
Le point de vue des dérivées partielles est pratique d'un point de vue calculatoire, mais donne souvent un point de vue un peu trop partiel et morcelé sur la fonction f étudiée. Il peut être intéressant de regrouper les informations. Dans ce but, nous définirons la notion de différentielle en un point a, qui est un objet qui regroupe toutes les dérivées directionnelles en un point. Plus précisément, la différentielle de f au point a est (sous réserve d'existence et de linéarité) l'application qui à tout vecteur $\vec{u}$ de E associe le vecteur dérivé en a selon la direction de $\vec{u}$.
Cette notion de différentielle permet d'avoir des formules beaucoup plus compactes et naturelles que celles provenant de l'étude des dérivées partielles, et surtout, elle permet de s'affranchir de la donnée d'une base.

Dans tout le chapitre, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$, E est un $\mathbb{R}$-e.v.n. (nécessairement réel) de dimension finie, F un $\mathbb{K}$-e.v.n. de dimension finie, et Ω un ouvert de E. Dans certains contextes liés à l'optimisation (recherche d'extrema), il pourra cependant être intéressant de considérer des fonctions définies sur un fermé, ou même sur un compact, afin d'assurer l'existence d'extrema globaux.

I Compléments sur les fonctions vectorielles

Cette section a pour but de généraliser rapidement certaines notions vues dans le cadre de fonctions de la variable réelle et à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ au cas de fonctions de la variable réelle aussi, mais à valeurs dans un e.v.n. F de dimension finie. Nous ne nous attardons pas sur ces différentes propriétés, qui sont assez limpides une fois qu'on a compris que, du fait de l'équivalence des normes, il suffit de se ramener systématiquement aux coordonnées. Ces propriétés seront complétées dans les sections suivantes par des propriétés un peu plus spécifiques (dérivation de formes bilinéaires ou multilinéaires composées par des fonctions dérivables) qui seront vues comme cas particuliers de propriétés similaires pour les fonctions d'une variable vectorielle.

I.1 Extention des notations o et O

Dans cette sous-section, E est un $\mathbb{R}$-e.v.n. de dimension finie, F est un $\mathbb{K}$-e.v.n. de dimension finie, et $D \subset E$ est une partie quelconque de E.

Définition 16.1.1 – Extension des notations o et O

Soit f est une application de D dans F et $g : D \to \mathbb{R}$. Soit $x_0 \in \overline{D}$. On dira que

- $f(x) \underset{x_0}{=} o(g(x))$ s'il existe $\alpha : D \to F$ tel que
$$\exists V \in \mathcal{V}(x_0),\quad \forall x \in V \cap D,\quad f(x) = g(x)\alpha(x), \qquad \text{et} \qquad \lim_{x\to x_0} \alpha(x) = 0;$$
- $f(x) \underset{x_0}{=} O(g(x))$ s'il existe $\alpha : D \to F$ tel que
$$\exists V \in \mathcal{V}(x_0),\quad \forall x \in V \cap D,\quad f(x) = g(x)\alpha(x), \qquad \text{et} \qquad \alpha \text{ bornée sur } V \cap D;$$

Remarque 16.1.2

Le terme d'erreur sera toujours contrôlé par une fonction à valeur dans $\mathbb{R}$, et non une fonction vectorielle.

Proposition 16.1.3 – Caractérisation de o et O par les coordonnées

Soit $f : I \to F$ et $g : I \to \mathbb{R}$, et $\mathcal{B} = (b_1, \dots, b_n)$ une base de F, et pour tout $x \in I$
$$f(x) = \sum_{i=1}^{n} f_i(x) b_i.$$
Alors

- $f(x) \underset{x_0}{=} o(g(x))$ si et seulement si pour tout $i \in [\![1, n]\!]$, $f_i(x) \underset{x_0}{=} o(g_i(x))$;
- $f(x) \underset{x_0}{=} O(g(x))$ si et seulement si pour tout $i \in [\![1, n]\!]$, $f_i(x) \underset{x_0}{=} O(g_i(x))$;

◁ **Éléments de preuve.**
Écrire α avec ses coordonnées, et utiliser la norme $\|\cdot\|_{\mathcal{B},\infty}$. ▷

Exemples 16.1.4

1. $\begin{pmatrix} t \\ t^2 \\ \ln(t) \end{pmatrix} \underset{+\infty}{=} o(t^3)$
2. $\begin{pmatrix} e^t & \sin(t) \\ \frac{1}{1+t} & \mathrm{Arctan}(t)) \end{pmatrix} \underset{0}{=} \begin{pmatrix} 1+t & t \\ 1-t & t \end{pmatrix} + O(t^2).$
3. Pour $A \in \mathcal{M}_n(\mathbb{C})$, $e^A \underset{A\to 0}{=} \sum_{k=0}^{n} \frac{A^k}{k!} + o(|\!|\!| A |\!|\!|^n).$

4. On en déduit, pour $t \in \mathbb{R}$: $e^{At} \underset{t \to 0}{=} I_n + tA + o(t)$.

La situation la plus fréquente sera le cas d'un terme d'erreur en $\|x\|^\alpha$, x étant une variable de l'e.v.n. E de dimension finie.

Lemme 16.1.5 – Invariance vis-à-vis de la norme en dimension finie

Soit E de dimension finie et $\|\cdot\|_1$ et $\|\cdot\|_2$ deux normes quelconques sur E. Soit $D \subset E$ et $f : E \to F$, et $x_0 \in \overline{F}$. Soit $\alpha \in \mathbb{R}$. Alors

$$f(x) \underset{x \to x_0}{=} o(\|x\|_1^\alpha) \iff f(x) \underset{x \to x_0}{=} o(\|x\|_2^\alpha)$$

et

$$f(x) \underset{x \to x_0}{=} O(\|x\|_1^\alpha) \iff f(x) \underset{x \to x_0}{=} O(\|x\|_2^\alpha).$$

Notation 16.1.6 – Notation $o(h)$, $O(h)$

Pour $h \in E$, on trouve parfois par abus la notation $o(h)$ à la place de $o(\|h\|)$, et de même pour $O(h)$, dans le but d'alléger un peu les notations.

Lorsque $F = \mathbb{R}$ ou $\mathbb{C}$, on peut former des quotients, et on retrouve la caractérisation usuelle, généralisée au cas d'une variable vectorielle.

Proposition 16.1.7 – Caractérisation de o par le quotient

Si $g : D \to \mathbb{R}$ ne s'annule pas au voisinage de x_0, les propositions suivantes sont équivalentes:

(i) $f(x) \underset{x_0}{=} o(g(x))$;

(ii) $\dfrac{f(x)}{g(x)} \underset{x \to 0}{\longrightarrow} 0$.

Proposition 16.1.8 – Caractérisation de O par le quotient

Si $g : D \to \mathbb{R}$ ne s'annule pas au voisinage épointé de x_0, les propositions suivantes sont équivalentes:

(i) $f(x) \underset{x_0}{=} O(g(x))$;

(ii) $\dfrac{f(x)}{g(x)}$ est bornée au voisinage de x_0.

Les propriétés opératoires sur les o et O sont les mêmes que dans le cas réel (sommation, transitivité etc).

I.2 Dérivation de fonctions vectorielles d'une variable réelle

Dans cette sous-section, on suppose que $E = \mathbb{R}$ et $I \subset E$ est un intervalle ouvert. On considère donc $f : I \to F$ une fonction d'une variable réelle à valeurs dans E.

Définition 16.1.9 – Fonction dérivable en a

Soit $f : I \to F$ et $a \in I$. On dit que f est dérivable en a si le taux d'accroissement

$$\tau_{x_0} f(x) = \frac{f(x) - f(a)}{x - a}$$

admet une limite finie $\ell \in F$ lorsque x tend vers a. Le vecteur dérivé en a est alors défini par :

$$f'(a) = \lim_{x \to x_0} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

La deuxième limite est prise en considérant $h \in I - x_0 = \{x - x_0, x \in I\}$. Ainsi, si x_0 est la borne inférieure de I, il s'agit d'une limite lorsque h tend vers 0^+.

Définition 16.1.10 – Fonction dérivable sur I, fonction dérivée

Soit $f : I \to F$. On dit que f est dérivable sur I si f est dérivable en tout $x_0 \in I$. On définit alors la fonction dérivée $f' : I \to F$ par :

$$\forall x \in I, \quad f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}.$$

Exemples 16.1.11

1. Dérivée de $f : t \mapsto tX_0$, $X_0 \in \mathbb{K}$.
2. Plus généralement, dérivée de $f : t \mapsto g(t)X_0$, $g : I \to \mathbb{R}$ étant une fonction dérivable en a , et X_0 un vecteur de F.
3. Dérivée de $t \mapsto \mathrm{e}^{\mathrm{i}\, t}$.

Proposition 16.1.12 – Caractérisation de la dérivabilité par les coordonnées

Soit $f : I \to F$, $x_0 \in I$, et $\mathcal{B}$ une base de F. On note, pour tout $x \in I$,

$$f(x) = \sum_{i=1}^{n} f_i(x) b_i,$$

la décomposition de f suivant ses composantes sur la base $\mathcal{B}$. Les propositions suivantes sont équivalentes:

(i) f est dérivable en x_0 ;

(ii) pour tout $i \in [\![1, n]\!]$, f_i est dérivable en x_0.

Dans ce cas, on a

$$f'(x_0) = \sum_{i=1}^{n} f_i'(x_0) b_i.$$

◁ **Éléments de preuve.**
Former le taux d'accroissement coordonnée par coordonnée, et utiliser la caractéristique similaire pour les limites. ▷

La caractérisation par les coordonnées permet de montrer qu'en définissant la dérivabilité à gauche et à droite comme dans le cas d'une fonction réelle, et en utilisant la caractérisation de la dérivée par les dérivées à gauche et à droite sur chacune de coordonnée, cela permet d'obtenie une caractérisation similaire dans le cas vectoriel.

Corollaire 16.1.13 – Opérations sur les dérivées

Soit I et J des intervalles de $\mathbb{R}$, et F un e.v.n. de dimension finie.

1. La dérivation est linéaire : si $f, g : I \to F$ et $\lambda \in \mathbb{K}$, alors si f et g sont dérivables en $a \in I$, $\lambda f + g$ aussi, de dérivée
$$(\lambda f + g)'(a) = \lambda f'(a) + g'(a).$$
2. Si $\varphi : J \to I$ est une fonction réelle dérivable, et $f : I \to F$ une fonction vectorielle dérivable, alors
$$\forall t \in J, \quad (f \circ \varphi)'(t) = f' \circ \varphi(t) \cdot \varphi'(t).$$

◁ **Éléments de preuve.**
Voir coordonnées par coordonnées. ▷

Corollaire 16.1.14 – Dérivabilité à gauche, à droite

Soit $x_0 \in \mathring{I}$ et $f : I \to F$. La fonction f est dérivable en x_0 si et seulement si elle est dérivable à gauche et à droite en x_0 et $f'_g(x_0) = f'_d(x_0)$.

Corollaire 16.1.15 – Dérivable implique continu

Soit $f : I \to \mathbb{R}$ et $x_0 \in I$. Si f est dérivable en x_0, alors f est continue en x_0.

◁ **Éléments de preuve.**
La caractérisation précédente nous ramène au cas des fonctions d'une variable. On peut aussi utiliser directement la propriété suivante (DL à l'ordre 1, qu'on peut ensuite restreindre à l'ordre 0). ▷

Proposition 16.1.16 – Caractérisation de la dérivabilité par DL à l'ordre 1

Soit $f : I \to F$ et $x_0 \in I$. Les propositions suivantes sont équivalentes:

(i) f est dérivable en x_0 ;

(ii) il existe A et B dans F tels que $f(x_0 + h) =_{h\to 0} A + hB + o(h)$.

◁ **Éléments de preuve.**
Caractérisation de o par quotient, puisque t ne s'annulle pas au voisinage épointé de 0. ▷

On donne une première conséquence, qui est plutôt un exemple, suffisamment important pour qu'il soit un résultat du cours (il servira pour l'étude des équations différntielles).

Corollaire 16.1.17 – Dérivée de $t \mapsto \exp(At)$

Soit $A \in \mathcal{M}_n(CC)$. La fonction $f : t \mapsto \exp(At)$ est dérivable sur $\mathbb{R}$, et

$$\forall t \in \mathbb{R}, \quad f't) = A\exp(At).$$

Une autre conséquence de la caractérisation par les DL est l'interprétation du vecteur dérivé en terme d'approximation de la trajectoire.

Corollaire 16.1.18 – Tangente à la trajectoire

Soit $f : I \to F$ et $x_0 \in I$. On suppose que f est dérivable en x_0 et que $f'(x_0) \neq 0$. Alors la droite affine $f(x_0) + \mathbb{R}f'(x_0)$ est tangente à la trajectoire (dans le sens où c'est la droite qui approche au mieux la trajectoire au voisinage de $f(x_0)$).

Remarque 16.1.19 – Interprétation cinétique de la dérivée

Si f représente la trajectoire d'un point dans F, et si f est dérivable en x_0 :

- $\|f'(x_0)\|_F$ est la vitesse instantanée du point (dépendant de la norme choisie sur F)
- le vecteur $f'(x_0)$ est tangent à la trajectoire, et indique le sens de déplacement.

Exemple 16.1.20

Décrire la trajectoire, et la vitesse instantanée d'un point, dont la position est définie pour $t \in \mathbb{R}_+$ par

$$f(t) = (\cos(t), \sin(t), t).$$

Théorème 16.1.21 – Caractérisation des fonctions constantes

Soit $f : I \to F$, dérivable sur l'intervalle I, et à valeurs dans l'e.v.n. de dimension finie F. Les propositions suivantes sont équivalentes:

(i) f est constante sur I

(ii) $f' = 0$ sur I.

◁ **Éléments de preuve.**
Voir dans une base. ▷

I.3 Fonctions vectorielles de classe $\mathcal{C}^k$

On considère toujours dans cette sous-section le cas d'une fonction $f : I \to F$ définie sur un intervalle I ouvert, et à valeurs dans un e.v.n. F de dimension finie. Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de F, et,pour $x \in I$, $[f(x)]_{\mathcal{B}} = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$.

Si f est dérivable sur I, alors elle définit une fonction dérivée

$$\forall x \in I, \quad f'(x) = \sum_{k=1}^{n} f_i'(x) b_i, \qquad \text{i.e.} \qquad [f'(x)]_{\mathcal{B}} = \begin{pmatrix} f_1'(x) \\ \vdots \\ f_n'(x) \end{pmatrix}.$$

On a encore $f' : I \to F$, et on peut donc étudier la dérivabilité de f', et ainsi de suite, comme dans le cas réel.
Si f est dérivable p fois, on note $f^{(k)}$ sa dérivée d'ordre k.

Définition 16.1.22 – Fonction de classe $\mathcal{C}^k$

La fonction $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^k$ si elle est k fois dérivable et si $f^{(k)} = I \to F$ est continue.

Définition 16.1.23 – caractérisation de la classe $\mathcal{C}^k$ et expression de la dérivée

Avec les notations introduites ci-dessus,

1. f est k fois dérivable si et seulement si chaque f_i est k fois dérivable, et dans ce cas,

$$\forall x \in I, \quad f^{(k)}(x) = \sum_{i=1}^{n} f_i^{(k)}(x) \qquad \text{i.e.} \qquad [f^{(k)}(x)]_{\mathcal{B}} = \begin{pmatrix} f_1^{(k)}(x) \\ \vdots \\ f_n^{(k)}(x) \end{pmatrix};$$

2. f est de classe $\mathcal{C}^k$ si et seulement si chaque f_i l'est.

Proposition 16.1.24 – Règles opératoires sur les fonctions de classe $\mathcal{C}^k$

Soit I et J deux intervalles de $\mathbb{R}$ et F un $\mathbb{K}$-e.v.n. de dimension finie. Soit $f, g : I \to \mathbb{R}$, $\varphi : J \to I$ et $\lambda \in \mathbb{K}$.

1. Si f et g sont de classe $\mathcal{C}^k$, alors $f = \lambda g$ est de classe $\mathcal{C}^k$, et

$$(f + \lambda g)^{(k)} = f^{(k)} + \lambda g^{(k)}.$$

2. Si f et φ sont de classe $\mathcal{C}^k$, $f \circ \varphi$ est de classe $\mathcal{C}^k$. Il n'y a pas de formule simple en général

3. Si f est de classe $\mathcal{C}^k$ et $\varphi : t \mapsto at + b$, alors $(f \circ \varphi)^{(k)} = a^k (f^{(k)} \circ \varphi)$.

◁ **Éléments de preuve.**
Ces propriétés sont bien connues pour les fonctions à valeurs réelles. S'y ramener en travaillant sur les coordonnées. ▷

II Dérivées partielles et différentielles

On s'intéresse maintenant à la situation plus générale d'une fonction $f : \Omega \to F$, où Ω est un ouvert d'un $\mathbb{R}$-e.v.n. E de dimension finie, et F est un $\mathbb{K}$-e.v.n. de dimension finie.

II.1 Dérivées selon un vecteur

Définition 16.2.1 – Fonctions partielles

Soit $f : \Omega \to F$, $a \in \Omega$, et $\vec{u} \in E \backslash \{0\}$. La fonction partielle $f_{a,\vec{u}}$ au point a et de direction $\vec{u}$, est définie, en tout $t \in \mathbb{R}$ tel que $a + t\vec{u} \in \Omega$, par

$$f_{a,\vec{u}}(t) = f(a + t\vec{u}).$$

Remarque 16.2.2

1. La fonction partielle $f_{a,\vec{u}}$ est bien définie au voisinage de 0, car Ω est ouvert.
2. La flèche sur $\vec{u}$ est présente ici uniquement pour faire remarquer que c'est la structure affine naturelle de E qui est en jeu ici. Les éléments de E jouent donc un rôle double de points et de vecteurs. Les éléments en lesquels f est évalué correspondent plutôt à des points, alors que $\vec{u}$ donne une direction de translation, et doit plutôt être considéré comme un vecteur. La fonction $f_{a,\vec{u}}$ correspond plus ou moins, après reparamétrage, à la restriction de f à la droite affine passant par a et dirigée par le vecteur $\vec{u}$.

Définition 16.2.3 – Dérivée selon un vecteur

Soit $f : \Omega \to F$, $a \in \Omega$, et $\vec{u} \in E \backslash \{0\}$.

1. La fonction f est dérivable selon le vecteur $\vec{u}$ si la fonction partielle $f_{a,\vec{u}}$ est dérivable en 0.
2. La dérivée de f en a selon $\vec{u}$ est alors défini par
$$D_{\vec{u}} f(a) = f'_{a,\vec{u}}(0) = \lim_{t \to 0} \frac{f(a + t\vec{u}) - f(a)}{t} \in F.$$
3. Si f est dérivable selon $\vec{u}$ en tout point a de Ω, on peut définir la fonction $D_{\vec{u}} f : \Omega \to F$.

Remarque 16.2.4

1. Dans la suite, afin d'alléger les notations, on omettra la flèche sur u, en notant simplement $D_u f(a)$.
2. L'existence d'une dérivée selon u en a n'implique pas l'existence de dérivées selon d'autres vecteurs non colinéaires ; en revanche, si $v \neq 0$ est colinéaire à u, l'existence de $D_u f(a)$ équivaut à celle de $D_v f(a)$, et on peut relier les deux quantités.

Exemple 16.2.5

Étudier l'existence des dérivées en 0 selon $u \in \mathbb{R}^2$ de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \mathbb{1}_{\mathbb{R}\times\{0\}}(x,y)$$

Proposition 16.2.6 – Linéarité de D par rapport à u sur chaque droite

Soit $u \in E\backslash\{0\}$ et $f : \Omega \to F$ dérivable en $a \in \Omega$ selon u. Soit $\lambda \in \mathbb{R}^*$. Alors f est dérivable en a selon λu et

$$D_{\lambda u} f(a) = \lambda D_u f(a).$$

◁ **Éléments de preuve.**
Revenir à la définition par taux d'accroissement. ▷

Avertissement 16.2.7

L'existence de toutes les dérivées directionnelles en a n'est pas suffisante pour assurer la continuité en a

Exemple 16.2.8 – (figure 16.1)

Étudier les dérivées directionnelles en $0_{\mathbb{R}^2}$ et la continuité de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \begin{cases} \frac{x^2}{y} & \text{si } y \neq 0 \\ 0 & \text{si } y = 0. \end{cases}$$

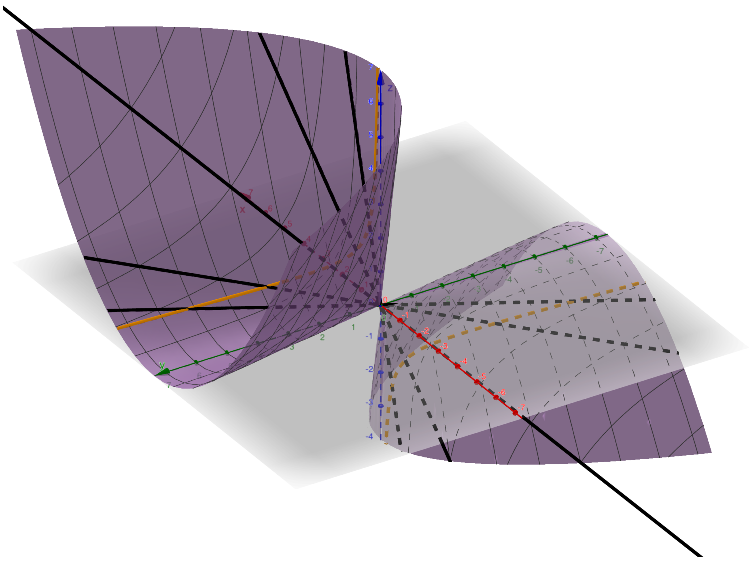

FIGURE 16.1 – Une fonction non continue en $(0,0)$ mais y admettant des dérivées selon toute direction

Définition 16.2.9 – Dérivées partielles dans une base

Soit $f : \Omega \to F$, $a \in \Omega$ et $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Pour $x \in E$, on note $X = (x_1, \dots, x_n)$ son vecteur coordonnées.

1. Soit $k \in [\![1, n]\!]$. La k-ième dérivée partielle de f dans la base $\mathcal{B}$ est la dérivée de f selon la direction b_k en a, et est notée $\partial_{\mathcal{B},k} f(a)$, ou, s'il n'y a pas d'ambiguïté sur la base utilisée,

simplement $\partial_k f(a)$. Ainsi

$$\partial_k f(a) = \partial_{\mathcal{B},k} f(a) = D_{b_k} f(a) = \lim_{t\to 0} \frac{f(a+tb_k) - f(a)}{t}.$$

2. Si f admet une k-ième dérivée partielle dans la base $\mathcal{B}$ en tout point de Ω, cela définit une application $\partial_k f : \Omega \to F$.

Remarque 16.2.10 – Identification à une fonction de plusieurs variables

Soit $f : \Omega \to F$ et $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Pour $x \in \Omega$, on note $[x]_{\mathcal{B}} = (x_1, \dots, x_n)$ ses coordonnées relativement à la base $\mathcal{B}$. On note

$$\tilde{\Omega} = \left\{ (x_1, \dots, x_n) \mid \sum_{k=1}^{n} x_k b_k \in \Omega \right\}.$$

Alors $\tilde{\Omega}$ est un ouvert de $\mathbb{R}^n$, et on peut identifier la fonction f à la fonction $\tilde{f} : \tilde{\Omega} \to F$ définie par

$$\tilde{f}(x_1, \dots, x_n) = f\left(\sum_{k=1}^{n} x_k b_k \right).$$

On obtient facilement, du fait de la possibilité de choisir la norme qu'on veut (en dimension finie) et donc notamment par exemple la norme $\|\cdot\|_{\mathcal{B},\infty}$ que :

- f est continue en x ssi $\tilde{f}$ est continue en $[x]_{\mathcal{B}}$;
- f admet en x une dérivée partielle selon b_k ssi $\tilde{f}$ admet en $[x]_{\mathcal{B}}$ une dérivée partielle selon e_k (vecteur de la base canonique), et alors

$$D_{b_k}(f)(a) = D_{e_k}(\tilde{f})(a).$$

Notation 16.2.11

Lorsque Ω est un ouvert de $\mathbb{R}^n$, on note aussi $\dfrac{\partial f}{\partial x_k}(a)$ la k-ième dérivée partielle par rapport à la base canonique. Ainsi, pour $a = (a_1, \dots, a_n)$, sous réserve d'existence,

$$\frac{\partial f}{\partial x_k}(a) = \lim_{t\to 0} \frac{f(a_1, \dots, a_k + t, \dots, a_n) - f(a_1, \dots, a_n)}{t}.$$

On l'appelle aussi dérivée partielle par rapport à la k-ième variable.

Proposition 16.2.12 – identification des dérivées partielles

Avec les notations de la remarque 16.2.10,

$$\frac{\partial \tilde{f}}{\partial x_k}(a) = D_{e_k}(f)(a) = \partial_{\mathcal{B},k} f(a).$$

Notation 16.2.13

Après avoir fixé la base $\mathcal{B}$ et identifié f à $\tilde{f}$, on pourra donc aussi identifier la k-ième dérivée partielles par rapport à la base $\mathcal{B}$ à la dérivée partielle par rapport à la k-ième variable de $\tilde{f}$, et on s'autorisera à écrire, s'il n'y a pas d'ambiguïté sur la base :

$$\partial_k f(a) = \frac{\partial f}{\partial x_k}(a).$$

Par ailleurs, il arrive souvent, dans le cadre d'une application d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}^p$ (situation à laquelle on se ramène par choix de bases), de travailler coordonnée par coordonnée. Dans ce cadre, on est ramené au calcul de dérivées partielles d'une fonction à valeurs dans $\mathbb{R}$:

Proposition 16.2.14 – Expression de la dérivée partielle d'une fonction à valeurs dans $\mathbb{R}^p$

Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}^p$, qu'on écrit $f = (f_1, \dots, f_p)$. Ainsi, pour tout $x \in \Omega$, $f(x) = (f_1(x), \dots, f_p(x)) \in \mathbb{R}^p$, et pour tout $k \in [\![1, p]\!]$, $f_k : \Omega \to \mathbb{R}$.
Alors f admet une dérivée partielle par rapport à la k-ième variable en a ssi c'est le cas de chacun des f_i, et dans ce cas,

$$\frac{\partial f}{\partial x_k}(a) = \left(\frac{\partial f_1}{\partial x_k}(a), \dots, \frac{\partial f_p}{\partial x_k}(a)\right).$$

Par ailleurs, on rappelle comment, concrètement, calculer les dérivée partielles d'une fonction dont on connaît l'expression :

Méthode 16.2.15 – Calcul d'une dérivée partielle

Soit f une application de $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}$ (ce à quoi on peut se ramener d'après toutes les propriétés précédentes d'identification ou de travail par coordonnées). Pour calculer la dérivée partielle par rapport à x_k, on utilise les règles usuelles de calcul de la dérivée en utilisant l'expression de f, dans laquelle on dérive par rapport à la variable x_k, en considérant que toutes les autres variables sont des constantes.

Exemple 16.2.16

Calculer la dérivée partielle par rapport à x de $f : (x, y, z) \mapsto (xy \ln(x + e^z), \sin(xe^z) \cos(y))$.

Avertissement 16.2.17

L'existence des dérivées partielles par rapport à une base n'implique pas l'existence de dérivées selon toutes les directions en un point a.

Exemple 16.2.18 – figure 16.2

Étudier les dérivées directionnelles en $(0, 0)$ de $f : (x, y) \mapsto \sqrt[4]{|xy|}$

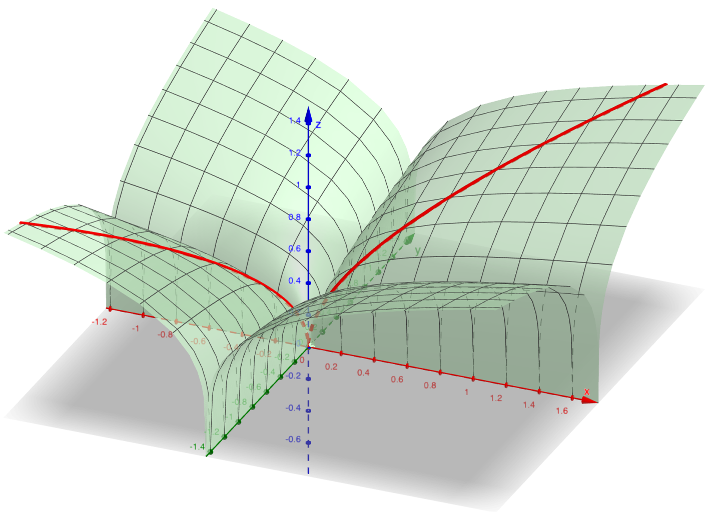

FIGURE 16.2 – Une fonction admettant des dérivées partielles en $(0, 0)$, mais pas de dérivée selon $(1, 1)$

Remarque 16.2.19

Cependant, sous des hypothèses supplémentaires de régularité, les dérivées partielles suffisent à retrouver les dérivées directionnelles. Par exemple, dans le cas des fonctions de 2 variables, vous avez montré l'année dernière que si f est de classe $\mathcal{C}^1$ au voisinage de a (dans le sens où ses deux dérivées partielles existent et sont continues au voisinage de a), alors, en notant $\nabla f(X) = \begin{pmatrix} \frac{\partial f}{\partial x}(X) \\ \frac{\partial f}{\partial x}(X) \end{pmatrix}$:

- pour tout $u = (u_x, u_y)$, $D_u f(a) = \langle \nabla f(a), u \rangle = u_x \frac{\partial f}{\partial x}(a) + u_y \frac{\partial f}{\partial y}(a)$,
- au voisinage de $a = (a_x, a_y)$, en notant $X = (x, y)$,

$$f(X) = f(a)+(x-a_x)\frac{\partial f}{\partial x}(a)+(y-a_y)\frac{\partial f}{\partial y}(a)+o(\|X-a\|) = f(a)+\langle X-a, \nabla f(a)\rangle+o(\|X-a\|)$$

(formule de Taylor-Young)

II.2 Différentielles

En particulier, d'après la remarque clôturant le paragraphe précédent, les fonctions de 2 variables suffisamment régulières se comportent à première approximation comme la fonction affine

$$X \mapsto f(a) + \langle X - a, \nabla f(a) \rangle,$$

donc la partie linéaire est l'application linéaire est

$$Y \mapsto \langle Y, \nabla f(a) \rangle = D_Y f(a).$$

Ainsi, sous certaines conditions de régularité, la dérivée directionnelle en a est une application linéaire en la direction, et permet de définir un développement limité à l'ordre 1.
C'est l'existence de cette approximation linéaire qui définit la notion de différentiabilité d'une application (tout comme la dérivabilité d'une fonction d'une variable équivaut à l'existence d'un DL_1).

Définition 16.2.20 – Différentiabilité et différentielle de f

Soit $f : \Omega \to F$, et $a \in \Omega$.

1. On dit que f est différentiable en a s'il existe une application linéaire $u \in \mathcal{L}(E, F)$ telle que au voisinage de a,
$$f(x) = f(a) + u(x-a) + o(\|x-a\|).$$
2. Dans ce cas, l'application linéaire u est notée $\mathrm{d}f(a)$, et est appelée différentielle de f au point a, ou application linéaire tangente à f en a.

Ainsi, la différentielle $\mathrm{d}f(a)$ est caractérisée par :

(i) sa linéarité ($\mathrm{d}f(a) \in \mathcal{L}(E, F)$)

(ii) le DL $f(a+h) \underset{0}{=} f(a) + [\mathrm{d}f(a)](h) + o(\|h\|)$

Pour alléger les notations, on écrira souvent $o(h)$ au lieu de $o(\|h\|)$ (et de même aux ordres supérieurs), et on omettra les crochets et parenthèses en écrivant

$$[\mathrm{d}f(a)](h) = \mathrm{d}f(a) \cdot h.$$

Remarque 16.2.21

Attention à bien comprendre les significations des différentes parenthèses. Si f est différentiable sur tout Ω par exemple,

$$\mathrm{d}f \ : \ \Omega \to \mathcal{L}(E, F)$$

est une fonction qui à tout point de Ω associe une application liénaire, qui est la partie linéaire de l'approximation affine de f en a. Pour tout a,

$$\mathrm{d}f(a) \ : E \to F$$

est donc une application linéaire, qu'on peut ensuite évaluer en $h \in E$. Nous verrons plus loin que, comme dans le cas d'une fonction de deux variables, $\mathrm{d}f(a) \cdot h$ est la dérivée directionnelle en a selon u.

Cette définition est à rapprocher de la caractérisation de la dérivabilité des fonctions d'une variable réelle par le DL_1. D'ailleurs, cette caractérisation permet d'en déduire le rapport entre dérivée et diférentielle dans le cas d'une fonction d'un intervalle I de $\mathbb{R}$ dans $\mathbb{R}$.

Proposition 16.2.22 – Différentielle d'une application d'une variable réelle

Soit Ω un ouvert de $\mathbb{R}$, $f : \Omega \to \mathbb{R}$ et $a \in \mathbb{R}$. Alors f est différentiable en a si et seulement si f est dérivable en a, et dans ce cas :

$$df(a) : \left| \begin{array}{ccc} \mathbb{R} & \longrightarrow & F \\ h & \longmapsto & f'(a) \cdot h \end{array} \right.$$

Cette proposition permet de voir la plupart des résultats suivants comme généralisation de la situation réelle bien connue.

Méthode 16.2.23 – Montrer que f est différentiable, méthode 1

Pour montrer que f est différentiable en a et calculer sa différentielle, on peut effectuer un DL_1 en a.

Exemples 16.2.24

1. Différentiabilité et différentielle en $(0,2,1)$ de la fonction définie sur $\mathbb{R}^3$ par $f(x,y,z) = \frac{\sin(x)-1}{1+x-y-z}$.
2. Différentiabilité en M et différentielle de $M \mapsto M^k$, définie sur $\mathcal{M}_n(\mathbb{R})$ à valeurs dans $\mathcal{M}_n(\mathbb{R})$
3. Différentiabilité et différentielle en 0 de $\exp : \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$.

Proposition 16.2.25 – Différentiable implique continu

Soit $f : \Omega \to F$ une application différentiable en a. Alors f est continue en a.

◁ **Éléments de preuve.**
Par DL ! Remarquer que $\mathrm{d}f(a)$ est continue en 0 puisque linéaire en dimension finie. ▷

Proposition 16.2.26 – Différentielle et dérivées directionnelles

Soit $f : \Omega \to F$ une application différentiable en a. Alors f admet en a des dérivées selon toute direction $u \in E$, et

$$D_u f(a) = \mathrm{d}f(a) \cdot u.$$

En particulier, si $\mathcal{B} = (b_1, \dots, b_n)$ est une base de E, et si f est identifiée à une fonction des variables $x_1, \dots, x_n$ coordonnées dans la base $\mathcal{B}$,

$$\forall k \in [\![1,n]\!], \quad \frac{\partial f}{\partial x_k}(a) = \mathrm{d}(a) \cdot b_k.$$

◁ **Éléments de preuve.**
Encore par DL. ▷

Corollaire 16.2.27 – Matrice de la différentielle

Soit $f : \Omega \to F$ une application différentiable en a, et $\mathcal{B}$ et $\mathcal{C}$ deux bases de E et F respectivement. Soit $(f_1, \dots, f_p)$ les fonctions coordonnées de f, définies sur E à valeurs dans $\mathbb{K}$. Alors

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\mathrm{d}f(a)) = \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant n}} \in \mathcal{M}_{p,n}(\mathbb{K}).$$

Corollaire 16.2.28 – Détermination de la différentielle par les dérivées partielles

Si f est différentiable en a, ses dérivées partielles selon une base $\mathcal{B}$ déterminent la différentielle. Plus précisément, pour tout $h \in E$ tel que $[h]_{\mathcal{B}} = (h_1, \dots, h_n)$,

$$\mathrm{d}f(a) \cdot h = \sum_{k=1}^{n} h_k \partial_k f = \sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k},$$

les dérivées partielles et les coordonnées étant prises relativement à la base $\mathcal{B}$.

Avertissement 16.2.29 – Dérivées directionnelles versus différentielle

L'existence des dérivées directionnelles ne suffit pas à obtenir la différentiabilité. L'hypothèse de différentiabilité de l'énoncé précédent est donc important.

En effet, l'exemple 16.1 définit une fonction admettant des dérivées directionnelles mais non continue en 0, donc non différentiable.

Définition 16.2.30 – Jacobienne d'une application de $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}^p$

Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$, la matrice de $\mathrm{d}f(a)$ relativement aux bases canoniques est appelée matrice jacobienne de f en a, et notée $J_f(a)$. Ainsi :

$$J_f(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_p}{\partial x_1} & \cdots & \dfrac{\partial f_p}{\partial x_n} \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{K}),$$

où $f = (f_1, \dots, f_p)$.

Méthode 16.2.31 – Montrer que f est différentiable, méthode 2

- Dans le cas où $E = \mathbb{R}^n$, $F = \mathbb{R}^p$, étudier l'existence des dérivées partielles en a, et considérer l'application linéaire v canoniquement associée à la matrice jacobienne. Si

$$f(a+h) - f(a) - v(h) = o(h),$$

  f est différentiable en a de différentielle $\mathrm{d}f(a) = v$.
- Adaptation immédiate dans les autres situations après avoir fixé des bases.

Exemple 16.2.32

Étudier la différentiabilité en $(0,0)$ de $f : (x,y) \mapsto \frac{x^3}{x^2+y^4}$, prolongée par 0 en 0.
NB : en exercice, assurez-vous que trouver un DL_1 de façon directe n'est pas aisé du tout !

Proposition 16.2.33 – Différentielle d'une application constante

Soit $f : \Omega \to \mathbb{R}$ une application constante. Alors f est différentiable sur Ω, de différentielle nulle.

Autrement dit, pour tout $a \in \Omega$, $\mathrm{d}f(a)$ est l'application linéaire nulle de E dans F.

Proposition 16.2.34 – Différentielle d'une AL

Soit $f \in \mathcal{L}(E,F)$. Alors f est différentiable en tout point $a \in E$, et pour tout $a \in E$,

$$\mathrm{d}f(a) = f.$$

Ainsi, $\mathrm{d}f$ est constante de valeur f.

Les notations $\mathrm{d}x$, $\mathrm{d}y$, etc, très utilisée par les physiciens notamment, peuvent être définies rigoureusement comme des cas particuliers de la proposition précédente.

Définition 16.2.35 – Notation $\mathrm{d}x_i$

Étant donné un espace E de dimension n, muni d'une base $\mathcal{B} = (b_1, \dots, b_n)$, on note $\mathrm{d}x_i$ la différentielle de la projection

$$p_i : \sum_{j=1}^{n} x_j p_j \mapsto x_i.$$

Ainsi, en tout $a \in E$,

$$\mathrm{d}x_i(a) = p_i \qquad \text{i.e.} \qquad \forall h \in E,\ \mathrm{d}x_i(a) \cdot h = h_i,$$

où h_i est la coordonnée selon b_i de h sur la base $\mathcal{B}$.

Remarque 16.2.36

L'expression de $\mathrm{d}f(a)$ en fonction des dérivées partielles se réécrit donc :

$$\forall h \in E,\quad \mathrm{d}f(a) \cdot h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i(a) \cdot h, \qquad \text{soit:} \qquad \mathrm{d}f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i(a),$$

ou encore, sur un domaine de différentiabilité de f :

$$\mathrm{d}f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i,$$

relation prisée des physiciens.

La notation $\mathrm{d}x_i$, bien qu'adhérente au programme et utile pour justifier les notations des physiciens, n'est pas explicitement au programme.

II.3 Gradient

On généralise dans cette section la notion de gradient introduite l'année dernière dans le cadre de fonctions de fonctions de deux variables réelles à valeurs dans $\mathbb{R}$. Dans ce contexte, le gradient a été défini en un point a comme le vecteur de $\mathbb{R}^2$ formé des deux dérivées partielles en a, autrement dit des deux dérivées selon les vecteurs d'une base.

On pourrait définir le gradient de la même manière, par rapport à une base, mais c'est plus commode de le définir sans référence à une base particulière. Pour cela, on se place dans le contexte où E est un espace euclidien et $F = \mathbb{R}$, et on utilise le lemme suivant, qu'on a déjà utilisé pour justifier l'existence de l'adjoint d'un endomorphisme.

Lemme 16.2.37 – Théorème de représentation de Riesz

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et $\varphi \in E^*$ une forme linéaire. Alors il existe un unique vecteur $a \in E$ tel que
$$\forall x \in E, \quad \varphi(x) = \langle a, x \rangle .$$

Proposition/Définition 16.2.38 – Gradient d'une application numérique

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et Ω un ouvert de E. Soit $f : \Omega \to \mathbb{R}$, et $a \in \Omega$ tel que f doit différentiable en a. Alors $\mathrm{d}f(a) \in E^*$. On définit alors le gradient de f en a comme étant l'unique vecteur $\nabla f(a) \in E$ tel que
$$\forall h \in E, \quad \mathrm{d}f(a) \cdot h = \langle \nabla f(a), h \rangle .$$

Théorème 16.2.39 – Expression du gradient en b.o.n.

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et Ω un ouvert de E. Soit $f : \Omega \to \mathbb{R}$, et $a \in \Omega$ tel que f doit différentiable en a. Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base orthonormale de E. Alors
$$\nabla f(a) = \sum_{k=1}^{n} \partial_k f(a) b_k,$$
les dérivées partielles étant prises dans la base $\mathcal{B}$. Ainsi :
$$[\nabla f(a)]_{\mathcal{B}} = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix}.$$

Corollaire 16.2.40 – Gradient d'une fonction de n variables réelles

On munit $\mathbb{R}^n$ (identifié à $\mathcal{M}_{n,1}(\mathbb{R})$) de sa structure euclidienne canonique. Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ une application différentiable en $a \in \Omega$. Alors
$$\nabla f(a) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

Remarque 16.2.41

Si $\Omega \subset \mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ est différentiable en a, alors la jacobienne en a est
$$J_f(a) = \mathrm{Mat}_{bc}(\mathrm{d}f(a)) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) \in \mathcal{M}_{1,n}(\mathbb{R}).$$
Ainsi, $J_f(a) = \nabla f(a)^\top$.

Proposition 16.2.42 – Première interprétation géométrique du gradient

Soit $\Omega \subset \mathbb{R}^n$, et $f : \Omega \to \mathbb{R}$ différentiable en $a \in \Omega$. Le gradient $\nabla f(a)$ est positivement colinéaire au vecteur unitaire u tel que $D_u f(a)$ soit maximal :
$$\max_{\|u\|=1} D_u f(a) \iff u = \frac{\nabla f(a)}{\|\nabla f(a)\|}$$

◁ **Éléments de preuve.**
C'est l'inégalité de Cauchy-Schwarz et son cas d'égalité, le choix du signe étant dicté par le fait que $D_{-u}f(a) = -D_u f(a)$, par linéarité de $\mathrm{d}f(a)$. ▷

Ainsi, en tout point de différentiabilité, le gradient indique la direction de plus forte pente.

II.4 Opérations sur les applications différentiables

On se place à nouveau dans la situation générale : E est un e.v.n. de dimension finie sur $\mathbb{R}$, F est un e.v.n. de dimension finie de $\mathbb{K}$ et Ω est un ouvert de E.

Proposition 16.2.43 – Linéarité

Soit f et $g : \Omega \to F$ différentiables en a, et $\lambda, \mu \in \mathbb{K}$. Alors $f + \lambda g$ est différentiable en a et

$$\mathrm{d}(f + \lambda g)(a) = \mathrm{d}f(a) + \lambda \mathrm{d}g(a).$$

Définition 16.2.44

Revenir à la définition par DL.

Proposition 16.2.45 – Différentielle d'une forme multilinéaire

Soit Ω un ouvert de E, $F_1, \dots, F_n$ des $\mathbb{K}$-e.v.n. de dimension finie, et $f_i : \Omega \to F_i$ différentiables en $a \in \Omega$. Soit $M : F_1 \times \cdots \times F_n \to F$ une application multilinéaire. Alors

$$M(f_1, \dots, f_n) : x \mapsto M(f_1(x), \dots, f_n(x))$$

est différentiable en a, et

$$\forall h \in E, \quad \mathrm{d}M(f_1, \dots, f_n)(a) \cdot h = \sum_{k=1}^{n} M(f_1(a), \dots, \mathrm{d}f_k(a) \cdot h, \dots, f_n(a)).$$

◁ **Éléments de preuve.**
Par DL et majoration de la forme multilinéaire (par continuité) ▷

Exemple 16.2.46

1. Soit $f, g : \Omega \to \mathbb{K}$, $\Omega \subset E$. Exprimer la différentielle de fg.
2. Si $f : \Omega \to F$ avec F euclidien, et si f est différentiable en a, montrer que $\|f\|^2$ est différentiable en tout point et exprimer sa différentielle.
3. En particulier, si E est un espace euclidien, exprimer la différentielle en tout point de $X \mapsto \|X\|^2$.

Corollaire 16.2.47 – Dérivée de $M(f_1, \dots, f_p)$

Soit I un intervalle ouvert de $\mathbb{R}$, $F_1, \dots, F_n, F$ des $\mathbb{K}$-e.v.n. de dimension finie, et $f_i : I \to F_i$. Soit M une forme multilinéaire de $F_1 \times \cdots \times F_n$ dans F. On pose :

$$g : t \mapsto M(f_1(t), \dots, f_n(t)).$$

Soit $a \in I$. Si pour tout $i \in [\![1, n]\!]$, f_i est dérivable en a, alors g aussi, et

$$g'(a) = \sum_{k=1}^{n} M(f_1(a), \dots, f'(a), \dots, f_n(a)).$$

Remarque 16.2.48

C'est une généralisation de la dérivée d'un produit (le produit est bilinéaire !)

Corollaire 16.2.49

Voici 3 cas particuliers classiques, a

1. Si L est une application linéaire de F dans G, et $f : I \to F$ dérivable en a, alors $(L \circ f)$ aussi, et
$$(L \circ f)'(a) = L(f'(a)).$$
2. Soit F un espace euclidien, et $f, g : I \to E$ deux applications dérivables sur un intervalle I. Alors $\langle f, g \rangle$ aussi, et
$$\frac{\mathrm{d}}{\mathrm{d}t} \langle f(t), g(t) \rangle = \langle f'(t), g(t) \rangle + \langle f(t), g'(t) \rangle .$$
3. Soit E un espace vectoriel de dimension n, et $\mathcal{B}$ une base de E. Alors, si $f_1, \ldots, f_n$ sont à valeurs dans E et dérivables,
$$\frac{\mathrm{d}}{\mathrm{d}t} \det_{\mathcal{B}}(f_1, \ldots, f_n) = \sum_{k=1}^{n} \det_{\mathcal{B}}(f_1, \ldots, f_k', \ldots, f_n).$$

Exemples 16.2.50

1. Exprimer la dérivée de $t \mapsto A(t)B(t)$, où
$$A : t \mapsto \mathcal{M}_{n,p}(\mathbb{C}) \qquad \text{et} \qquad B : t \mapsto \mathcal{M}_{p,q}(\mathbb{C})$$
sont deux applications dérivables sur I.
Par exemple, exprimer la dérivée de $\begin{pmatrix} 2 & 4 & 1+t \\ 3 & 2+2t & 2 \\ 1 & t & t \end{pmatrix} \begin{pmatrix} 2t & 4 & -t \\ 0 & 1 & 1-t \\ 2 & 1 & t \end{pmatrix}$.
2. Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et $X : I \to E \backslash \{0\}$ dérivable sur I. Montrer que $t \mapsto \|X(t)\|$ est dérivable sur I, et exprimer sa dérivée.
3. Dans le même contexte, montrer que $t \mapsto \frac{X(t)}{\|X(t)\|}$ est dérivable, et exprimer sa dérivée.
4. Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $X : I \mapsto \mathcal{M}_{n,1}(\mathbb{K})$ et $Y : I \mapsto \mathcal{M}_{p,1}(\mathbb{K})$. On suppose X et Y dérivables. Exprimer la dérivée de $X(t)^\top AY(t)$.
5. Elle se généralise à des fonctions à valeurs dans $\mathcal{M}_n(\mathbb{C})$.
Par exemple, exprimer la dérivée de $X(t)^n$, lorsque $X : I \to \mathcal{M}_n(\mathbb{K})$.
6. Exprimer la dérivée de l'application qui à $t \in \mathbb{R}$ associe $\begin{vmatrix} t-2 & 3 & 6 \\ 0 & t+1 & 2 \\ 2 & 5 & t-3 \end{vmatrix}$

Proposition 16.2.51 – Différentielle d'une composée

Soit E, et F deux e.v.n. de dimension finie sur $\mathbb{R}$, et G un e.v.n. de dimension finie sur $\mathbb{K}$. Soit Ω un ouvert de E et Ω' un ouvert de f, et $f : \Omega \to F$ et $g : \Omega' \to G$ deux applications. Soit $a \in \Omega$. On suppose que

(i) f est différentiable en a

(ii) $f(a) \in \Omega'$ et g est différentiable en $f(a)$.

Alors $g \circ f$ est différentiable en a, et

$$\mathrm{d}(g \circ f)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a).$$

◁ **Éléments de preuve.**
Composer les DL. ▷

Corollaire 16.2.52 – Composition par une fonction réelle

Soit Ω un ouvert d'un e.v.n. E, et I un intervalle ouvert de $\mathbb{R}$. Soit $f : \Omega \to \mathbb{R}$ et $g : I \to \mathbb{R}$. Si f est différentiable en $a \in \Omega$, si $f(a) \in I$ et si g dérivable en $f(a)$, alors $g \circ f$ est différentiable en a et

$$\mathrm{d}(g \circ f)(a) = g'(f(a))\mathrm{d}f(a).$$

Exemples 16.2.53

1. Différentiabilité de $x \mapsto \|x\|$ en tout $x \neq 0$ dans un espace euclidien.
2. Différentiabilité de $\|f\|$ lorsque f l'est.

Corollaire 16.2.54 – Dérivée le long d'un arc

Soit I un intervalle ouvert de $\mathbb{R}$, Ω un ouvert de E. Soit $\gamma : I \to \Omega$ et $f : \Omega \to F$. Si γ est dérivable en $a \in I$, et f différentiable en $\gamma(a)$, alors $f \circ \gamma$ est dérivable en a et

$$(f \circ \gamma)'(a) = \mathrm{d}f(a) \cdot \gamma'(a) = \left\langle \nabla f(a), \gamma'(a) \right\rangle.$$

Exemples 16.2.55

Dérivée de $t \mapsto f(a + th)$. Cohérence en 0.

Corollaire 16.2.56 – Règle de la chaîne

Avec les notations et hypothèses de la proposition 16.2.51, si $\mathcal{B} = (b_1, \dots, b_n)$ est une base de E, et $\mathcal{C} = (c_1, \dots, c_p)$ une base de F,

$$\forall k \in [\![1, n]\!], \partial_{\mathcal{B},k}(g \circ f)(a) = \sum_{i=1}^{p} \partial_{\mathcal{C},i}\, g(f(a))\partial_{\mathcal{B},k} f(a).$$

Avertissement 16.2.57

Attention à ne pas confondre $\partial_i g(f(a))$ et $\partial_i(g \circ f)(a)$.. Quelle est la différence entre les deux ?

Corollaire 16.2.58 – Reexpression de la règle de la chaîne dans $\mathbb{R}^n$

Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$, et si les bases considérées sont les bases canoniques, on obtient, avec les mêmes hypothèses :

$$\forall k \in [\![1, n]\!], \frac{\partial(g \circ f)}{\partial x_k}(a) = \sum_{i=1}^{p} \frac{\partial g}{\partial y_i}(f(a))\frac{\partial f}{\partial x_k}(a),$$

où $(x_1, \dots, x_n)$ représentent les coordonnées génériques dans $E = \mathbb{R}^n$ et $(y_1, \dots, y_p)$ les coordonnées génériques dans $F = \mathbb{R}^p$.
De façon équivalente, en exprimant coordonnée par coordonnée, $J_{g \circ f}(a) = J_g(f(a)) \times J_f(a)$.

Exemple 16.2.59

Soit f une application de $\mathbb{R}^3$ dans $\mathbb{R}$ admettant des dérivées partielles en tout (x, y, z). Exprimer les dérivées partielles de

$$g(u, v, w) = f(u^2 + v, 2v, 3w - e^v).$$

Proposition 16.2.60 – Différentielle de $(f_1, \dots, f_p)$

Soit Ω un ouvert de E, et $F_1, \dots, F_q$ des $\mathbb{K}$-e.v.n. de dimension finie, et soit $f_i : \Omega \to F_i$, $i \in [\![1, q]\!]$. Les propositions suivantes sont équivalentes:

(i) Pour tout $i \in [\![1, q]\!]$, f_i est différentiable en $a \in \Omega$

(ii) $f = (f_1, \dots, f_q) : E \to F_1 \times \dots \times F_q$ est différentiable en a.

Si c'est le cas,

$$\mathrm{d}f(a) = (\mathrm{d}f_1(a), \dots, \mathrm{d}(f_q)(a)) \in \mathcal{L}(E, F_1 \times \dots \times F_q),$$

c'est-à-dire :

$$\forall h \in E, \quad \mathrm{d}f(a) = (\mathrm{d}f_1(a) \cdot h, \dots, \mathrm{d}f_q(a) \cdot h).$$

Remarque 16.2.61

Retrouver de la sorte l'expression de la jacobienne.

III Applications de classe $\mathcal{C}^n$

III.1 Applications de classe $\mathcal{C}^1$

Définition 16.3.1 – Application de classe $\mathcal{C}^1$

Soit $f : \Omega \to F$. On dit que f est de classe $\mathcal{C}^1$ sur Ω si :

(i) f est différentiable en tout point de Ω

(ii) $\mathrm{d} : \Omega \to \mathcal{L}(E, F)$ est continue

Exemples 16.3.2

1. Soit $f : E \to F$ une application linéaire, alors f est de classe $\mathcal{C}^1$.
2. Soit f une application constante, alors f est de classe $\mathcal{C}^1$.

Proposition 16.3.3 – Opérations sur les applications de classe $\mathcal{C}^1$

Soit E, E' des $\mathbb{R}$-e.v.n. de dimension finie, et F, $F_1, \dots, F_q$ des $\mathbb{K}$-e.v.n. de dimension finie. Soit Ω, Ω' des ouverts de E et E' respectivement.

1. Soit $f, g : \Omega \to F$ de classe $\mathcal{C}^1$ et $\lambda \in \mathbb{K}$. Alors $f + \lambda g$ est de classe $\mathcal{C}^1$.
2. Soit pour $i \in [\![1, q]\!]$, $f_i : \Omega \to F_i$, et $M : F_1 \times \dots \times F_q \to F$ une forme multilinéaire. Si les f_i sont de classe $\mathcal{C}^1$, alors $M(f_1, \dots, f_q)$ est de classe $\mathcal{C}^1$.
3. Soit $f : \Omega' \to \Omega$ et $g : \Omega \to F$. Si f et g sont de classe $\mathcal{C}^1$ sur Ω' et Ω respectivement, alors $g \circ f$ est de classe $\mathcal{C}^1$ sur Ω.
4. Si $f : \Omega \to \mathbb{K}$ ne s'annule pas et est de classe $\mathcal{C}^1$, alors $\frac{1}{f}$ aussi.

5. Soit pour $i \in [\![1,q]\!]$, $f_i : \Omega \to F_i$, et $f : \Omega \to F_1 \times \cdots \times F_q$ définie par

$$\forall x \in \Omega, \quad f(x) = (f_1(x), \dots, f_q(x)).$$

Alors f est de classe $\mathcal{C}^1$ si et seulement si tout f_i est de classe $\mathcal{C}^1$.

◁ **Éléments de preuve.**

Attention à bien comprendre sur quelle variable porte la continuité (le point $x \in \Omega$ en lequel on considère la différentielle). Cela conditionne la validité de l'argument donné. Par exemple, pour la composée, la différentielle $\mathrm{d}(g \circ f)$ s'exprime comme composée de deux différentielles, mais la composée porte sur la variable h, et non sur a qui est à voir plutôt comme un paramètre. Donc on ne peut pas obtenir la continuité de cette différentielle par propriété de continuité d'une composée.

En revanche, puisque la composition se fait entre applications linéaires d'espaces de dimension finie, l'application $(f,g) \mapsto g \circ f$ est une application bilinéaire entre espaces de dimension finie, donc elle-même continue. C'est ce point qui est en jeu ici.

Pour 2, on peut remarquer que si $M : E^n \to F$ est une application multilinéaire, alors $\widetilde{M} : E^{k-1} \times \mathcal{L}(E,F) \times E^{n-k} \to \mathcal{L}(E,F)$ définie par

$$\widetilde{M}(x_1, \dots, x_{k-1}, \alpha, x_{k+1}, \dots, x_n)(x) = M(x_1, \dots x_{k-1}, \alpha(x), x_{k+1}, \dots, x_n)$$

est également multilinéaire, donc continue, tous les espaces en jeu étant de dimension finie. ▷

Vous avez défini l'année dernière le caratère $\mathcal{C}^1$ des fonctions de 2 variables à l'aide des dérivées partielles. On retrouve ce point de vue via la caractérisation 16.3.5. On démontre d'abord le résultat suivant, qui n'apparaît qu'implicitement au programme (cela peut être vu comme un sous-produit de la caractérisation 16.3.5 et de la définition de la différentielle)

Théorème 16.3.4 – Formule de Taylor-Young à l'ordre 1 pour les dérivées partielles

Soit $f : \Omega \to \mathbb{K}$ une application à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. Soit $\mathcal{B}$ une base de E, et $a \in \Omega$. On suppose que pour tout f, $\partial_k f$ est définie et continue au voisinage de a. Alors

$$f(a+h) \underset{h \to 0}{=} f(a) + \sum_{k=0}^{n} h_k \partial_k f(a) + o(h).$$

où $(h_1, \dots, h_n) = [h]_{\mathcal{B}}$

◁ **Éléments de preuve.**

Utiliser la formule de Taylor pour les fonctions d'une variable, successivement pour chacune des n variables. ▷

Théorème 16.3.5 – Caractérisation de la classe $\mathcal{C}^1$ par les dérivées partielles

Soit $f : \Omega \to F$ et $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Les propositions suivantes sont équivalentes:

(i) f est de classe $\mathcal{C}^1$ sur Ω ;

(ii) f admet des dérivées partielles en tout point selon la base $\mathcal{B}$, et pour tout $k \in [\![1,n]\!]$, $\partial_k f : \Omega \to F$ est continue.

◁ **Éléments de preuve.**

Démonstration non exigible.

- Se ramener à $f_i : \Omega \to \mathbb{K}$, en considérant les coordonnées dans une base $\mathcal{C}$ de F.
- Utiliser la décomposition $\mathrm{d}f_i(a) = \sum_{k=1}^{n} \partial_k f_i$.

Cela revient en gros à utiliser le fait que $\mathrm{d}f$ est continue si et seulement si $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\mathrm{d}f)$ est continue, ce qui est la matrice des dérivées partielles des coordonnées dans $\mathcal{C}$ (l'analogue de la jacobienne)
Dans le sens $(ii) \implies (i)$, ne pas oublier de montrer d'abord la différentiabilité de f, en utilisant la formule de Taylor-Young. ▻

Remarque 16.3.6

La formule de Taylor 16.3.4 se réécrit alors simplement

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + o(h),$$

qui n'est rien de plus que la combinaison de la définition de la différentielle et du gradient ∇f.

Corollaire 16.3.7 – Continuité via les dérivées partielles

Si $E = \mathbb{R}^n$ et si f admet des dérivées partielles par rapport à tout x_i en a, et que ces dérivées partielles sont continues au voisinage de a, alors f est continue au voisinage de a.

Avertissement 16.3.8

Comme on l'a déjà vu, l'existence seule des dérivées partielles (et même plus généralement des dérivées directionnelles) ne suffit pas à établir la continuité.

Théorème 16.3.9 – Intégrale de la différentielle le long d'un arc

Soit $f : \Omega \to F$ une application de classe $\mathcal{C}^1$, et $\gamma : [0,1] \to \Omega$ un chemin de classe $\mathcal{C}^1$ tel que $\gamma(0) = a$ et $\gamma(1) = b$. Alors

$$f(b) - f(a) = \int_0^1 \mathrm{d}f(\gamma(t))\gamma'(t)\ \mathrm{d}t.$$

◅ **Éléments de preuve.**
L'intégrande est la dérivée de $f \circ \gamma : [0,1] \to F$. ▻

Remarque 16.3.10

Ce théorème peut être vu comme une généralisation du théorème fondamentale de l'analyse, puisqu'en prenant $E = \mathbb{R}$ et $\gamma : t \mapsto a(1-t) + bt$, puis en faisant un changement de variable affine, on retrouve

$$\int_a^b f(t)\ \mathrm{d}t = f(b) - f(a)$$

lorsque $f : [a,b] \to F$ est de classe $\mathcal{C}^1$.

Exemple 16.3.11

1. Si f est de classe $\mathcal{C}^1$ sur l'ouvert Ω et si $[a,b] \subset \Omega$, alors

$$f(b) - f(a) = \int_0^1 \mathrm{d}f(a(1-t) + bt) \cdot (b-a)\ \mathrm{d}t$$

2. De façon équivalente, si f est de classe $\mathcal{C}^1$ sur l'ouvert Ω et si $[a, a+u] \subset \Omega$, alors

$$f(a+u) - f(a) = \int_0^1 \mathrm{d}f(a+tu) \cdot u\ \mathrm{d}t.$$

3. Si f est de classe $\mathcal{C}^1$ sur l'ouvert Ω et si $\gamma : [0,1] \to \Omega$ est un lacet de classe $\mathcal{C}^1$ (*i.e.*

$\gamma(a) = \gamma(b)$), alors

$$\int_0^1 \mathrm{d}f(\gamma(t)) \cdot \gamma'(t)\ \mathrm{d}t = 0.$$

Théorème 16.3.12 – Caractérisation des fonctions constantes (1)

Soit Ω un ouvert convexe de E, et $f : \Omega \to F$. Les propositions suivantes sont équivalentes:

(i) f est constante sur Ω

(ii) f est différentiable et $\mathrm{d}f$ est nulle sur Ω.

◁ **Éléments de preuve.**

- Le sens direct est déjà vu.
- Pour le sens réciproque, utiliser le théorème 16.3.9 sur une paramétrisation du segment $[a, b]$ pour montrer que pour tout $a, b \in \Omega$, $f(a) = f(b)$.

▷

Ce théorème admet une généralisation qui est aussi au programme, mais sans sa démonstration (que nous esquissons quand même).

Théorème 16.3.13 – Caractérisation des fonctions constantes (2)

Soit Ω un ouvert connexe par arcs de E, et $f : \Omega \to F$. Les propositions suivantes sont équivalentes:

(i) f est constante sur Ω

(ii) f est différentiable et $\mathrm{d}f$ est nulle sur Ω.

◁ **Éléments de preuve.**

Démonstration hors-programme.

C'est toujours la réciproque qui reste à voir.

On pourrait essayer d'adapter la preuve précédente, en montrant que si Ω est ouvert et connexe par arcs, on peut toujours joindre deux points par un arc de classe $\mathcal{C}^1$. Mais cette construction n'est pas si évidente que cela à faire.

On procède donc autrement, en « grignotant » petit à petit.

Soit $a \in \Omega$, et $\Omega' \subset \Omega$ l'ensemble des points $x \in \Omega$ tels que $f(a) = f(x)$. Si $\Omega' \neq \Omega$, considérer $b \in \Omega \backslash \Omega'$ et joindre a à b. Montrer l'existence d'un point c sur cet arc, étant sur la frontière de Ω'. Trouver une contradiction en utilisant la caractérisation (1) et la convexité d'une petite boule centrée en c.

▷

III.2 Plan tangent et noyau de la différentielle

Définition 16.3.14 – Vecteur tangent à une partie, figure 16.3

Soit $X \subset E$ une partie de E, et $x \in X$.

1. Un vecteur $v \in E$ est tangent à X en x s'il existe $\varepsilon > 0$ et un arc $\gamma :]-\varepsilon, \varepsilon[\to X$ dérivable en 0 tel que $\gamma'(0) = v$.
2. On note $T_x X$ l'ensemble des vecteurs tangents à X en x.

Exemples 16.3.15

1. Quels sont les vecteurs tangents à $\varnothing$?
2. 0 est tangent à tout X non vide.

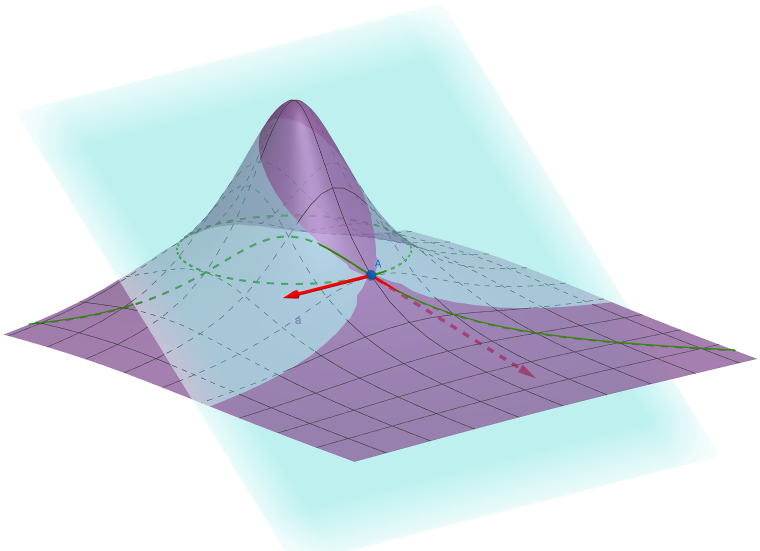

Figure 16.3 – Un sous-ensemble de $\mathbb{R}^3$ et deux vecteurs tangents en A.

3. Décrire $T_x X$ lorsque $x \in \mathring{X}$.
4. Décrire les vecteurs tangents à $\mathbb{R} \times \mathbb{R}_+$ en $(0,0)$.

D'autres exemples importants sont mis en résultats du cours :

Proposition 16.3.16 – Vecteurs tangents à un sous-espace affine

Soit X un sous-espace affine de E de direction V. Alors, pour tout $B \in X$, $T_B(X) = V$.

◁ **Éléments de preuve.**

L'inclusion directe est évidente en paramétrant des droites incluses dans X passant par B.

Réciproquement la dérivée de γ est limite d'un taux d'accroissement qui est dans V. Comme on est en dimension finie, V est fermé. ▷

Proposition 16.3.17 – Vecteurs tangents à une sphère, figure 16.4

Soit E un espace euclidien, $c \in E$ et $r > 0$. Soit $X = S(c,r)$ et $x \in S(c,r)$. Alors $T_x X = \overrightarrow{cx}^{\perp}$.
En d'autres termes, l'ensemble des vecteurs tangents à une sphère en un point x est l'orthogonal du rayon défini par x.

◁ **Éléments de preuve.**

Considérer les grands cercles passant par x.

Pour l'autre inclusion, vérifier que si γ est à valeurs dans S, sa dérivée est orthogonale au rayon (dériver un produit scalaire). ▷

Exemple 16.3.18

Décrire sous forme d'une équation le plan vectoriel $T_A(S)$ où $S = S(0,1) \subset \mathbb{R}^3$ et $A = (a,b,c) \in S$.

Exemple 16.3.19 – Vecteur tangent au graphe d'une fonction

Soit $f : \Omega \to \mathbb{R}$ une application numérique. Son graphe $\mathcal{G}$ est un sous-ensemble de l'e.v.n. $\mathrm{E} \times \mathbb{R}$. Soit $x_0 \in \Omega$, on peut donc définir l'ensemble des vecteurs tangents au graphe de f au point $(x_0, f(x_0)) \in \mathcal{G}$. La figure 16.3 est en fait (à condition de rajouter les axes au bon endroit) le graphe de la fonction

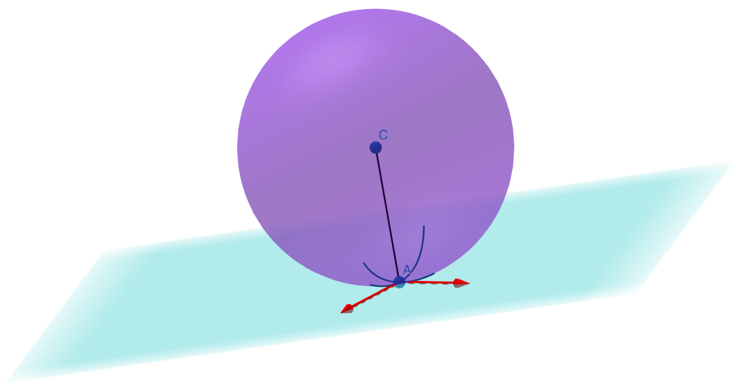

FIGURE 16.4 – Plan tangent à une sphère

$(x, y) \mapsto \dfrac{2}{1 + 2x^2 + 2y^2}$, le point de tangence étudié étant en $(1, 0)$.

Théorème 16.3.20 – Description d'un espace tangent à un graphe

Soit Ω un ouvert de $\mathbb{R}^2$, $f : \Omega \to \mathbb{R}$, et $X_0 \in \Omega$. On suppose f différentiable en X_0. L'ensemble $T_{X_0}f$ des vecteurs tangents à la courbe de f en $(X_0, f(X_0))$ est le plan d'équation

$$h\frac{\partial f}{\partial x}(X_0) + k\frac{\partial f}{\partial y}(X_0) = \ell,$$

c'est-à-dire l'ensemble des vecteurs $H = (h, k, \ell) \in \mathbb{R}^3$ vérifiant l'équation ci-dessus.

◁ **Éléments de preuve.**
Comme dans les autres exemples, étudier les propriétés des arcs du graphe, et réciproquement, trouver des arcs adéquats. ▷

Dans le cas général ($\Omega \subset E$, E de dimension finie), on obtient la description suivante :

Théorème 16.3.21 – Espace tangent à un ensemble défini par une équation

Soit $f : \Omega \to \mathbb{R}$ une fonction numérique de classe $\mathcal{C}^1$, $X = f^{-1}(\{0\})$. Ainsi X est l'ensemble défini par l'équation $f(x) = 0$.
Soit $a \in X$. Si $\mathrm{d}f(a) \neq 0$, l'ensemble des vecteurs tangents à X en a est égal à

$$T_aX = \mathrm{Ker}(\mathrm{d}f(a)).$$

◁ **Éléments de preuve.**
Hors-programme. ▷

Exemple 16.3.22

1. Retrouver à l'aide de ce théorème la description des vecteurs tangents à un espace affine de $\mathbb{R}^n$, décrit par une équation
$$a_1x_1 + \cdots + a_nx_n = b.$$
2. De même, retrouver la description des vecteurs tangents de la sphère unité de l'espace euclidien $\mathbb{R}^n$.
3. Retrouver la description de l'espace des vecteurs tangents au graphe d'une fonction $f : \Omega \to \mathbb{R}$ où $\Omega \subset \mathbb{R}^2$, lorsque f est de classe $\mathcal{C}^1$ au voisinage de X_0. Généraliser pour $\Omega \subset \mathbb{R}^n$.

Proposition 16.3.23 – Deuxième interprétation géométrique du gradient

Avec les hypothèse du théorème 16.3.21, si on suppose de plus que E est un espace euclidien, $T_aX = \nabla f(a)^{\perp}$.

◁ **Éléments de preuve.**

Dans le contexte euclidien, cette formulation est équivalente à cette du théorème, par définition du gradient.

On peut le démontrer indépendamment du théorème, en considérant un arc à valeurs dans X, et en dérivant $f \circ \gamma$. ▷

Exemple 16.3.24

Retrouver l'équation des hyperplans des vecteurs tangents à la sphère unité de $\mathbb{R}^n$.

Remarque 16.3.25

1. Si $E = \mathbb{R}^2$, X est la ligne de niveau de valeur 0. Plus généralement, en considérant $f - b$ au lieu de f, on peut décrire de la sorte toutes les lignes de niveau. La proposition affirme que le gradient en a est orthogonal à la droite tangente à la ligne de niveau en a, ce qu'on peut réexprimer de façon plus abrégée en disant que le gradient est orthogonal aux lignes de niveau.
2. Ainsi, si f est de classe $\mathcal{C}^1$ sur un ouvert de $\mathbb{R}^2$, en combinant les deux interprétations, la direction de plus forte pente au point a est toujours orthogonale à la ligne de niveau en a.

Exemple 16.3.26 – Figure 16.5

Soit $f : \mathbb{R}^2 \backslash \{(0,0)\} \to \mathbb{R}$ définie par

$$f(x,y) = \frac{x^2+y^3}{x^4+y^2}.$$

En exprimant le gradient, déterminer la tangente à la courbe de niveau en $(1,1)$. La figure 16.5 illustre cet exemple. Le gradient en $(1,1)$ est le vecteur $\overrightarrow{AB}$, reporté sur le plan $z = 1$, dans lequel est tracée en rouge la courbe de niveau au point $(1,1)$.

III.3 Applications de classe $\mathcal{C}^k$

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E, $(i,j) \in [\![1,n]\!]$, et $f : \Omega \to F$ admettant une dérivée partielle $\partial_i f$ sur tout Ω. Si la fonction $\partial_i f : \Omega \to F$ admet elle-même une j-ième dérivée partielle en $a \in \Omega$, on peut donc définir $\partial_j \partial_i f(a)$, la j-ième dérivée partielle de $\partial_i f$. De façon plus générale, on définit par récurrence sur k les dérivées d'ordre k :

Définition 16.3.27 – Dérivées partielles d'ordre k

On définit, si c'est possible, les dérivées d'ordre k de f par récurrence sur k :

- Si $k = 1$, il s'agit des dérivées partielles $\partial_i f$
- Soit $k > 1$ et $(i_1, \dots, i_k) \in [\![1,n]\!]^k$. La fonction f admet en $a \in \Omega$ une dérivée par rapport successivement aux $i_1, \dots, i_{k-1}$ et i_k-ièmes vecteurs de la base $\mathcal{B}$, notée $\partial_{i_k}\partial_{i_{k-1}} \dots \partial_{i_1} f(a)$, si et seulement si $\partial_{i_{k-1}} \dots \partial_{i_1} f(a)$ existe au voisinage (relatif dans Ω) de a et est dérivable selon b_{i_k} en a. Ainsi,

$$\partial_{i_k}\partial_{i_{k-1}} \dots \partial_{i_1} f(a) = \partial_{i_k}\left(\partial_{i_{k-1}} \dots \partial_{i_1} f\right)(a).$$

On trouve aussi les notations suivantes :

- $\partial_{i_1,\dots,i_k} f(a)$ (attention à l'ordre des indices !),

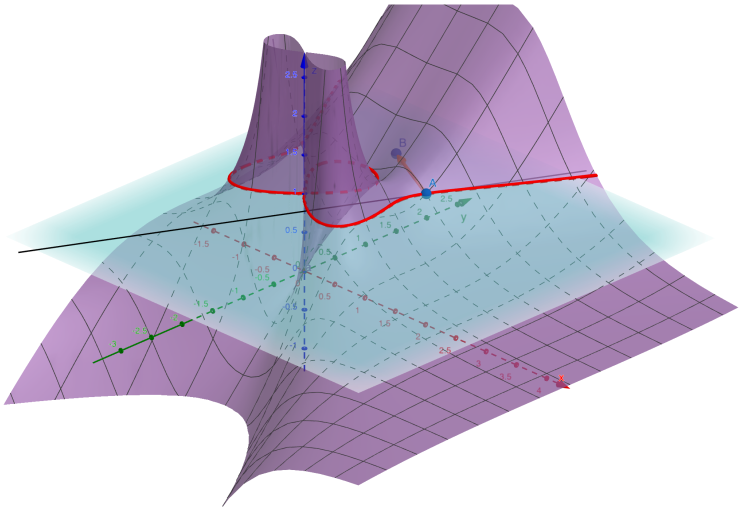

FIGURE 16.5 – Courbe de niveau $\dfrac{x^2+y^3}{x^4+y^2} = 1$, tangente et gradient en $A = (1,1,1)$.

- $\dfrac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(a)$, si $(x_1, \dots, x_n)$ représentent les coordonnées génériques dans la base $\mathcal{B}$, en particulier si $E = \mathbb{R}^n$ muni de la base canonique.

On peut définir des différentielles d'ordre supérieur, mais cette notion est hors programme, et les objets manipulés deviennent de plus en plus complexes. Pour cette raison, nous définissons la classe $\mathcal{C}^k$ à l'aide des dérivées partielles d'ordre k. Cette définition est une généralisation de la caractérisation de la classe $\mathcal{C}^1$ par les dérivées partielles premières.

Définition 16.3.28 – Applications de classe $\mathcal{C}^k$

Soit $f : \Omega \to F$ une application définie sur l'ouvert Ω de E.

1. On dit que f est de classe $\mathcal{C}^k$ sur Ω si toutes les dérivées d'ordre k de f existent (il y en a 2^k), et sont continues sur Ω.
2. On dit que f est de classe $\mathcal{C}^k$ au voisinage de $a \in \Omega$, s'il existe un ouvert $V \subset \Omega$ tel que $f_{|V}$ soit de classe $\mathcal{C}^k$ sur V.
3. On note $\mathcal{C}^k(\Omega, F)$ l'ensemble des fonctions de classe $\mathcal{C}^k$ sur Ω, à valeurs dans F. Par convention, $\mathcal{C}^0(\Omega, F)$ est l'ensemble des fonctions continues.
4. On dit que f est de classe $\mathcal{C}^\infty$ sur Ω si elle est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.
5. On note $\mathcal{C}^\infty(\Omega, F)$ l'ensemble des fonctions de classe $\mathcal{C}^\infty$ sur Ω, à valeurs dans F.

Remarque 16.3.29

Si f est de classe $\mathcal{C}^k$, alors :
- f est de classe $\mathcal{C}^j$ pour tout $j \leqslant k$;
- plus généralement, si $\ell \leqslant k$ et $(i_1, \dots, i_\ell) \in [\![0, n]\!]^\ell$, $\partial_{i_1,\dots,i_\ell} f$ est de classe $\mathcal{C}^{k-\ell}$.

Pour étudier la classe $\mathcal{C}^k$, il y a donc *a priori* 2^k dérivées à étudier. Mais en pratique, cela en fait moins, car sous des conditions assez peu restrictives, on va retrouver plusieurs fois la même, dans le sens où l'ordre de dérivation (*i.e.* le choix de l'ordre des variables par rapport auxquelles on dérive) n'importe pas. C'est ce que dit le théorème suivant.

Théorème 16.3.30 – Théorème de Schwarz

Soit $f : \Omega \to F$, $\mathcal{B} = (b_1, \dots, b_n)$ une base de E, et $a \in \Omega$. Si f est de classe $\mathcal{C}^2$ au voisinage de a, alors

$$\forall (i,j) \in [\![1,n]\!], \quad \partial_i \partial_j f(a) = \partial_j \partial_i f(a).$$

◁ **Éléments de preuve.**

Démonstration non exigible.

En fixant toutes les autres variables, on est ramené au cas de fonctions de 2 variables. Exprimer un DL_2 de $f(x+h,y+k)-f(x,y)$ en $H=(h,k)$, de deux manières différentes, en intercalant $f(x+h,y)$, ou $f(x,y+k)$, et en se ramenant à la formule de Taylor-Young à l'ordre 2 pour les fonctions d'une variable (en fixant l'autre). Identifier les deux DL obtenus (justifier proprement cette identification). ▷

Corollaire 16.3.31 – Théorème de Schwarz pour la classe $\mathcal{C}^k$

Soit $f : \Omega \to F$, $\mathcal{B} = (b_1, \dots, b_n)$ une base de E, et $a \in \Omega$. Si f est de classe $\mathcal{C}^2$ au voisinage de a, alors pour tout $(i_1, \dots, i_k) \in [\![1,n]\!]^k$, et toute permutation $\sigma \in \mathfrak{S}_k$,

$$\partial_{i_1,\dots,i_k} f(a) = \partial_{i_{\sigma(1)},\dots,i_{\sigma(k)}} f(a)$$

◁ **Éléments de preuve.**

Les dérivées partielles successives commutent 2 à 2. On peut donc ramener d'abord i_1 en position initiale, par échange successif, pour faire ensuite une récurrence. Ou directement utiliser le fait que $\mathfrak{S}_n$ est engendré par les transpositions $(i\ i+1)$. ▷

Ainsi, l'ordre des variables de dérivation importe peu, dès lors que la fonction est de classe $\mathcal{C}^k$, ce qui permet de regrouper éventuellement les dérivées selon les variables, afin d'écrire ces dérivées sous la forme :

$$\frac{\partial^k f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}, \quad \alpha_1 + \cdots + \alpha_n = k.$$

Cette notation signifie qu'on a dérivé α_1 fois par rapport à x_1, α_2 fois par rapport à x_2 etc. Certains α_i peuvent être nuls et peuvent être omis dans la notation. On notera par exemple $\dfrac{\partial^3 f}{\partial x \partial z^2}$ la dérivée d'ordre 3 d'une fonction de 3 variables x, y, z, obtenue en dérivant une fois par rapport à x et 2 fois par rapport à z, mais aucune fois par rapport à y.

Proposition 16.3.32 – Opérations sur les applications de classe $\mathcal{C}^k$

Soit E, E' des $\mathbb{R}$-e.v.n. de dimension finie, et F, $F_1, \dots, F_q$ des $\mathbb{K}$-e.v.n. de dimension finie. Soit Ω, Ω' des ouverts de E et E' respectivement.

1. Soit $f, g : \Omega \to F$ de classe $\mathcal{C}^k$ et $\lambda \in \mathbb{K}$. Alors $f + \lambda g$ est de classe $\mathcal{C}^k$.
2. Soit pour $i \in [\![1,q]\!]$, $f_i : \Omega \to F_i$, et $M : F_1 \times \cdots \times F_q \to F$ une forme multilinéaire. Si les f_i sont de classe $\mathcal{C}^k$, alors $M(f_1, \dots, f_q)$ est de classe $\mathcal{C}^k$.
3. Soit $f : \Omega' \to \Omega$ et $g : \Omega \to F$. Si f et g sont de classe $\mathcal{C}^k$ sur Ω' et Ω respectivement, alors $g \circ f$ est de classe $\mathcal{C}^k$ sur Ω.
4. Si $f : \Omega \to \mathbb{K}$ ne s'annule pas et est de classe $\mathcal{C}^k$, alors $\frac{1}{f}$ aussi.
5. Soit pour $i \in [\![1,q]\!]$, $f_i : \Omega \to F_i$, et $f : \Omega \to F_1 \times \cdots \times F_q$ définie par
$$\forall x \in \Omega, \quad f(x) = (f_1(x), \dots, f_q(x)).$$
f est de classe $\mathcal{C}^k$ si et seulement si tout f_i est de classe $\mathcal{C}^k$.

Ces propriétés restent vraies pour la classe $\mathcal{C}^\infty$

◁ **Éléments de preuve.**

Démonstrations non exigibles, un peu fastidieuse mais sans difficulté. Faire une récurrence, en dérivant une première fois. ▷

Exemples 16.3.33

1. Les fonctions polynomiales $f : \mathbb{R}^n \to \mathbb{K}$, de n variables réelles et à coefficients dans $\mathbb{K}$, sont de classe $\mathcal{C}^\infty$.
2. Les fractions rationnelles aussi, sur leur domaine de définition.
3. $\det : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ est de classe $\mathcal{C}^\infty$.
4. $M \mapsto M^{-1}$ définie sur $\mathrm{GL}_n(\mathbb{R})$ est de classe $\mathcal{C}^\infty$.

Dans le cas de fonctions d'une seule variable réelle, on rajoute une propriété, concernant la dérivée itérée d'une forme bilinaire. Lorsqu'on applique cette formule au cas du produit de deux réels, on retrouve la formule de Leinbniz classique.

Théorème 16.3.34 – Formule de Leibniz vectorielle

Soit E, F, G des $\mathbb{K}$-evn de dimension finie, et $B : EtimesF \to G$ des $\mathbb{K}$-evn. Soit I un intervalle, et $f : I \to E$ et $g : I \to F$ deux applications. Si f et g sont de classe $\mathcal{C}^k$, alors $B(f,g) : t \mapsto B(f(t), g(t))$ est aussi de classe $\mathcal{C}^k$, et

$$\forall t \in I, \quad B(f,g)^{(k)}(t) = \sum_{i=1}^{k} \binom{k}{i} B(f^{(i)}(t), g^{(k-i)}(t))\ \mathrm{d}t.$$

III.4 Hessienne et formule de Taylor-Young à l'ordre 2

Définition 16.3.35 – Matrice hessienne de f

Soit Ω un ouvert de $\mathbb{R}^n$, et $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$. La matrice hessienne de f en $a \in \Omega$ est

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{1 \leqslant i,j \leqslant n} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial {x_1}^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \dfrac{\partial^2 f}{\partial {x_n}^2}(a) \end{pmatrix}.$$

Proposition 16.3.36 – Symétrie de $H_f(a)$

Sous les mêmes hypothèses, $H_f(a) \in \mathcal{S}_n(\mathbb{R})$.

◁ **Éléments de preuve.**

C'est une réexpression du théorème de Schwarz. ▷

Lemme 16.3.37 – Expression de la dérivée seconde par rapport à un vecteur

Soit Ω un ouvert de $\mathbb{R}^n$, et $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$. Soit $a \in \Omega$ et $H \in \mathbb{R}^n$. Alors, la dérivée seconde selon le vecteur U au point a s'exprime à l'aide de la hessienne :

$$f''_{a,U}(0) = U^T H_f(a) U = \sum_{i=1}^{n} \sum_{j=1}^{n} u_i u_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a), \quad \text{où} \quad U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

La matrice hessienne va nous permettre d'exprimer de façon élégante la formule de Taylor-Young, c'est-à-dire l'expression d'un DL à l'ordre 2 à l'aide des dérivées d'ordre 1 et 2, sous les hypothèses idoines de régularité.

Nous avons déjà obtenu une première version de cette formule, lorsque f est une fonction de 2 variables de classe $\mathcal{C}^2$ au voisinage d'un point a, lors de la démonstration du théorème de Schwarz :

$$f(x+h,y+k) = f(x,y)+h\frac{\partial f}{\partial x}(x,y)+k\frac{\partial f}{\partial y}(x,y)+\frac{1}{2}h^2\frac{\partial^2 f}{\partial x^2}(x,y)+hk\frac{\partial^2 f}{\partial x \partial y}(x,y)+\frac{1}{2}k^2\frac{\partial^2 f}{\partial y^2}(x,y)+o(\|H\|^2),$$

où $H = \begin{pmatrix} h \\ k \end{pmatrix}$. À l'aide du lemme précédent, on obtient une réexpression synthétique de cette formule, en posant $X = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$f(X+H) = f(X) + \langle \nabla f(X), H \rangle + \frac{1}{2} H^\top H_f(x) H + o(\|H\|^2).$$

Plus généralement :

Théorème 16.3.38 – Formule de Taylor-Young à l'ordre 2

Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$ au voisinage de $a \in \Omega$. Alors f admet un DL à l'ordre 2 en a, s'exprimant :

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} h^\top H_f(a) h + o(\|h\|^2).$$

◁ **Éléments de preuve.**

Démonstration non exigible. On peut procéder comme dans le cas de 2 variables, en utilisant la formule de Taylor-Young pour les fonctions d'une variables, successivement sur chaque variable. On peut aussi exploiter la symétrie de $H_f(a)$ en considérant d'abord une b.o.n. de diagonalisation $(b_1, \dots, b_n)$, et en exploitanta la formule de Taylor à l'ordre 2 pour les fonctions partielles associées. Les exprimer par ε, et exprimer le cas général par linéairité et bilinéarité. ▷

III.5 Équations aux dérivées partielles, EDP

Une équation aux dérivées partielles (EDP) est une équation d'inconnue une fonction f de plusieurs variables, reliant ses différentes dérivées partielles. L'ordre de l'EDP est l'ordre de dérivation le plus important intervenant dans cette équation. Par exemple, l'équation suivante, pour une fonction inconnue f de 2 variables, l'équation

$$\frac{\partial f^3}{\partial x^3} = \frac{\partial f^2}{\partial y^2} + xy \frac{\partial^2 f}{\partial x \partial y}$$

est une EDP d'ordre 3. La résolution d'équations aux dérivées partielles est un problème en général dur. Nous nous contentons ici de donner quelques pistes, sur des exemples simples.

Le principe va souvent être de se ramener à des équations simples. Pour cela, nous commençons par une forme simple d'equation de degré 1.

Proposition 16.3.39 – Résolution de $\dfrac{\partial f}{\partial x_i} = 0$

Soit $i \in [\![1, n]\!]$. Les propositions suivantes sont équivalentes:

(i) $f : \mathbb{R}^n \to \mathbb{R}$ est de classe $\mathcal{C}^1$ et vérifie $\dfrac{\partial f}{\partial x_i} = 0$ sur Ω

(ii) il existe $g \in \mathcal{C}^1(\mathbb{R}^{n-1}, \mathbb{R})$ tel que pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$f(x_1, \ldots, x_n) = g(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

En d'autres termes, les solutions de l'EDP $\dfrac{\partial f}{\partial x_i} = 0$ sont les fonctions ne dépendant pas de la vaiable x_i et de classe $\mathcal{C}^1$ par rapport aux autres variables.

Exemple 16.3.40

Résolution de $a\dfrac{\partial f}{\partial x} + b\dfrac{\partial f}{\partial y} = 0$, par changement de variable.

Nous donnons deux exemples d'EDP d'ordre 2 :

Exemples 16.3.41

1. Résoudre $\dfrac{\partial^2 f}{\partial x \partial y} = 0$ où $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$.
2. Résoudre $\dfrac{\partial^2 f}{\partial x^2} - a\dfrac{\partial^2 f}{\partial y^2} = 0$, où $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$. On cherchera un changement de variable linéaire.

IV Optimisation

IV.1 Principes généraux

L'optimisation consiste en l'étude des extrema d'une fonction f (globaux ou locaux).

Définition 16.4.1 – Extrema

Soit D un sous-ensemble de E, et $f : D \to \mathbb{R}$, et $a \in D$. On dit que :

- f admet un minimum global (resp. maximum global) en a si pour tout $x \in D$, $f(x) \geqslant f(a)$ (resp. $f(x) \leqslant f(a)$) ;
- f admet un minimum global strict (resp. maximum global strict) en a si pour tout $x \in D \backslash \{a\}$, $f(x) > f(a)$ (resp. $f(x) < f(a)$) ;
- f admet un minimum local (resp. maximum local) en a s'il existe $V \in \mathcal{V}(a)$ tel que pour tout $x \in D \cap V$, $f(x) \geqslant f(a)$ (resp. $f(x) \leqslant f(a)$) ;
- f admet un minimum local strict (resp. maximum local strict) en a s'il existe $V \in \mathcal{V}(a)$ tel que pour tout $x \in D \cap V \backslash \{a\}$, $f(x) > f(a)$ (resp. $f(x) < f(a)$).

Le but de cette section est de donner des principes de recherche des extrema (locaux ou globaux).
Les sous-sections suivantes donneront des méthodes pour trouver en quels points de D il existe un extremum, local ou global. Pour commencer, nous nous intéressons à l'existence d'un extremum global (ou local par restriction à un voisinage).
Nous rappelons le théorème de compacité, outil incontournable dans ce contexte.

Théorème 16.4.2 – Compacité

Si D est compact et $f : D \to \mathbb{R}$ est continue, alors f admet un minimum et un maximum globaux.

Il est fréquent d'étudier des fonctions qui ne sont pas définies sur des compacts. Mais il est possible de s'y ramener, soit en restreignant (si D n'est pas borné), soit en prolongeant (si D est borné).

Méthode 16.4.3 – Prouver l'existence d'un extremum global

On peut bien entendu utiliser les méthodes de localisation des extrema des paragraphes suivants et comparer les valeurs à toutes les autres, mais il peut être intéressant de savoir justifier *a priori* l'existence d'un extremum global de $f : D \to \mathbb{R}$.

1. Si D est compact et f continue, le théorème de compacité donne l'existence d'un maximum et et d'un minimum.
2. Si on parvient à trouver un compact $K \subset D$, et $a \in D$ tel que
$$\forall x \in D\backslash D, \quad f(x) \geqslant f(a),$$
alors f admet un minimum sur D, atteint dans K.
3. Si D est borné, $\overline{D}$ est borné. Si f admet une limite en tout point de $\overline{D}\backslash D$, et s'il existe $a \in D$ tel que
$$\forall b \in \overline{D}\backslash D, \quad \lim_{x \to b} f(x) \geqslant f(a),$$
alors f admet un minimum sur D.

Exemples 16.4.4 – figures 16.6 et 16.7

1. Montrer que $f : (x,y) \mapsto \dfrac{x^2y - 3y + y^2}{(1 + x^2 + y^2)^2}$ admet un minimum et un maximum sur $\mathbb{R}^2$.
2. Montrer que $g : (x,y) \mapsto \dfrac{\mathrm{e} - \mathrm{e}^{xy}}{1 - (x^2 + y^2)}$ admet un minimum sur $\mathring{B}(0,1)$.

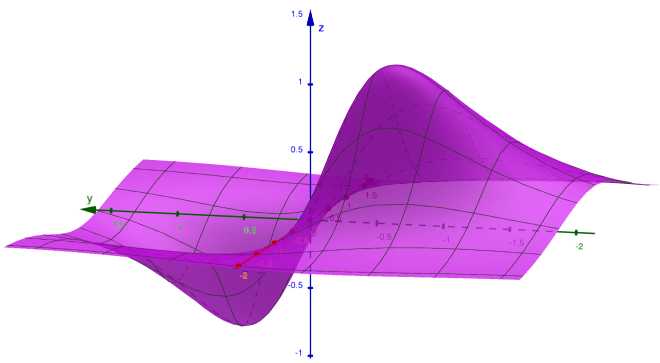

Figure 16.6 – $f : x \mapsto \dfrac{x^2y - 3y + y^2}{(1 + x^2 + y^2)} = 1$

IV.2 Condition nécessaire du premier ordre

Théorème 16.4.5 – CN du premier ordre pour un extremum local

Soit $f : D \to \mathbb{R}$. Si f admet un extremum local en $a \in \mathring{D}$ et si f est différentiable en a, alors $\mathrm{d}f(a) = 0$.

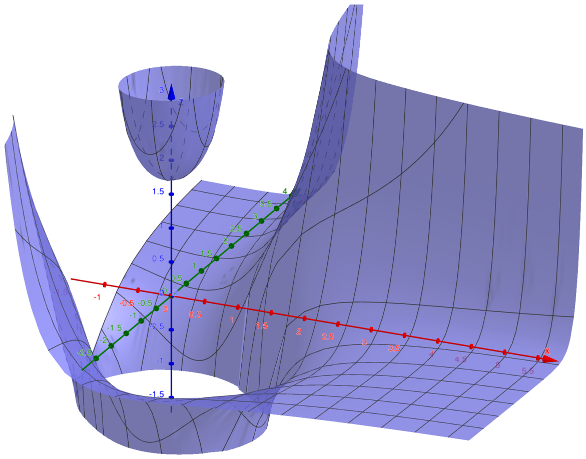

FIGURE 16.7 – $g : x \mapsto \dfrac{\mathrm{e} - \mathrm{e}^{xy}}{1 - (x^2 + y^2)}$

◁ **Éléments de preuve.**
Se ramener à la CN similaire pour les fonctions d'une variable réelle, par l'étude de toutes les dérivées directionnelles $\mathrm{d}f(a) \cdot h$. Ou alors, montrer directement la proposition 16.4.15 qui est une version plus générale de ce théorème. ▷

Ce théorème motive la définition suivante, comme dans le cas des fonctions d'une variable réelle.

Définition 16.4.6 – Point critique

Soit $f : D \to \mathbb{R}$. On dit que $a \in E$ est un point critique de f si $a \in \mathring{D}$ et $\mathrm{d}f(a) = 0$.

Remarque 16.4.7

Si f est une fonction d'une variable réelle, $a \in \mathring{D}$ est point critique si et seulement si f est dérivable en a et l'application linéaire $\mathrm{d}f(a) : h \mapsto f'(a)h$ est nulle, c'est-à-dire si et seulement si $f'(a) = 0$. On retrouve la notion de point critique définie en MPSI.

Corollaire 16.4.8 – CN portant sur le gradient dans le cas euclidien

Soit E un espace euclidien, $D \subset E$, et $f : D \to \mathbb{R}$. Si f admet un extremum en $a \in \mathring{D}$ en lequel f est différentiable, alors $\nabla f(a) = 0$.

Exemple 16.4.9

Rechercher les points critiques de $g : (x, y) \mapsto \dfrac{\mathrm{e} - \mathrm{e}^{xy}}{1 - x^2 - y^2}$ sur $\mathring{B}(0, 1)$, et en déduire l'existence d'un minimum global.

Remarque 16.4.10

Si le nombre de points critiques est peu élevé et si on a pu montrer l'existence d'extrema locaux, comme dans le paragraphe précédent, l'étude de la CN du premier ordre peut être suffisante pour conclure.

Avertissement 16.4.11

Comme dans le cas d'une fonction d'une variable réelle, $\mathrm{d}f(a) = 0$ n'est pas une CN pour avoir un extremum local.

Exemple 16.4.12 – figure 16.8

Quels sont les points critiques de $f : (x, y) \mapsto xy$? La fonction f admet-elle un extremum local ? Cet exemple montre au passage que le fait que les restrictions de f à toutes les droites passant par a admettent un extremum local en a n'est pas suffisant pour que f admette un extremum local en a.

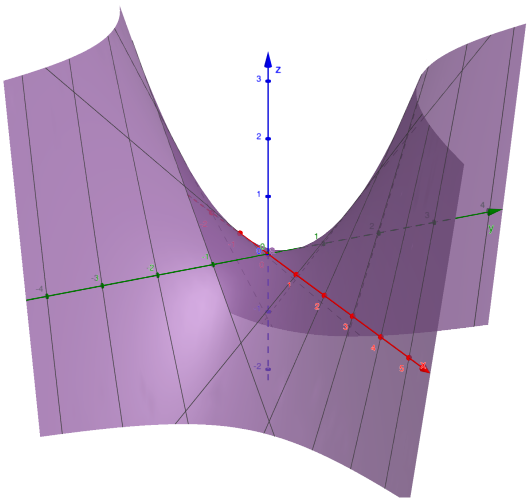

FIGURE 16.8 – $f : x \mapsto xy$

Méthode 16.4.13 – Rechercher les extrema locaux sur un ouvert

On cherche les points critiques de f, puis on étudie chaque point critique a :

- soit par étude directe du signe de $f(x) - f(a)$,
- soit en utilisant la CN d'ordre 2 qui fait l'objet d'un paragraphe ultérieur.

Méthode 16.4.14 – Rechercher les extrema globaux

Soit $f : D \to \mathbb{R}$, différentiable sur $\mathring{D}$.

- Rechercher les points critiques sur $\mathring{D}$. Les extrema globaux, s'ils existent, sont un point critique, ou un point de $D \backslash \mathring{D}$.
- Si on a pu justifier l'existence d'un extremum global, on peut le trouver en comparant les valeurs prises aux différents points critiques, ainsi que le maximum et/ou le minimum de f restreinte à $D \backslash \mathring{D}$.

Le paragraphe suivant donne justement des techniques en vue de l'étude des extrema d'une fonction retreinte. Cela permet notamment l'étude aux bords (si D lui-même peut être plongé dans un ouvert Ω de sorte que $D \backslash \mathring{D}$ soit une partie de Ω).

IV.3 Optimisation sous contrainte

Soit $f : \Omega \to \mathbb{R}$, et $X \subset D$. Optimiser f sous la contrainte X consiste à étudier les extrema de f en se retreignant à X (la contrainte). Il s'agit donc d'étudier les extrema d'une restriction. Nous commençons par une CN similaire à celle qu'on a trouvé dans le cas global (qui d'ailleurs peut être vu comme conséquence de celle-ci).

Proposition 16.4.15 – CN du premier ordre pour une restriction

Soit Ω un ouvert de E, X une partie de Ω, et $f : \Omega \to \mathbb{R}$. Si $f_{|X}$ admet un extremum local en $a \in X$, alors

$$\forall h \in T_a X, \quad \mathrm{d}f(a) \cdot h = 0.$$

En d'autres termes, $\mathrm{d}(a)$ s'annule en tout vecteur tangent à X, c'est-à-dire $T_a X \subset \operatorname{Ker} \mathrm{d}f(a)$.

◁ **Éléments de preuve.**
Dériver $f(\alpha(t))$, où $\alpha = [-\varepsilon, \varepsilon] \to X$ vérifie $\alpha(0) = a$. ▷

Remarques 16.4.16

1. En prenant $X = \mathring{D}$, puisque X est un voisinage de a, $T_a X = E$, et on retrouve donc le théorème 16.4.5.
2. Si E est euclidien, la proposition 16.4.15 affirme que si $f_{|X}$ admet un extremum local en a, $\nabla f(a) \perp T_a X$.

Si l'ensemble X est défini par une équation $g(x) = 0$, on peut combiner le résultat précédent et le théorème 16.3.21.

Théorème 16.4.17 – Optimisation sous contrainte

Soit $f, g : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^1$, et $X = \{x \in \Omega,\ g(x) = 0\}$. Si $f_{|X}$ admet un extremum local en $a \in X$ tel que $\mathrm{d}g(a) \neq 0$, alors $\mathrm{d}f(a)$ et $\mathrm{d}g(a)$ sont colinéaires.

Exemple 16.4.18

Après avoir montré leur existence, déterminer les extrema globaux de $f : (x, y) \mapsto x^2 + xy - 2y^2$ sur $\overline{B}(0, 1)$.

IV.4 Étude à l'ordre 2

Si $f : I \to \mathbb{R}$ est une fonction définie sur un intervalle ouvert de $\mathbb{R}$, et si $a \in I$ est un point critique de f, on peut déterminer s'il s'agit effectivement d'un extremum par un argument de convexité locale, ce qui, en cas de régularité suffisante, se ramène à l'étude du signe de $f''(a)$. Nous recherchons une condition similaire lorsque f est une fonction définie sur un ouvert de $\mathbb{R}^n$, à l'aide de la matrice hessienne $H_f(a)$.

Proposition 16.4.19 – CN du second ordre

Soit $f : \Omega \to \mathbb{R}$ (où Ω est un ouvert de $\mathbb{R}^n$) de classe $\mathcal{C}^2$. Si f admet un minimum local (resp. maximum local) en $a \in \Omega$, alors a est un point critique de f et $H_f(a) \in \mathcal{S}^+(\mathbb{R})$ (resp. $-H_f(a) \in \mathcal{S}^+(\mathbb{R})$).

◁ **Éléments de preuve.**
D'après la formule de Taylor-Young, le signe de $f(a + h) - f(a)$ est celui de $h^\top H_f(a) h$. ▷

Théorème 16.4.20 – CS du second ordre

Soit $f : \Omega \to \mathbb{R}$ (où Ω est un ouvert de $\mathbb{R}^n$) de classe $\mathcal{C}^2$, et soit $a \in \Omega$. Si a est un point critique de f et si $H_f(a) \in \mathcal{S}^{++}(\mathbb{R})$ (resp. $-H_f(a) \in \mathcal{S}^{++}(\mathbb{R})$), alors f admet un minimum local strict (resp. maximum local strict) en a.

◁ **Éléments de preuve.**
Toujours le contrôle du signe par Taylor-Young, en travaillant avec une norme adaptée à $H_f(a)$. ▷

Avertissement 16.4.21

- La CN du second ordre n'est pas une CS.
- La CS du second ordre n'est pas une CN

Exemples 16.4.22

Étudier les CN et CS du second ordre pour les fonctions suivantes, en 0.

1. $f : (x, y) \mapsto (x^2 + y^2)^2$.
2. $g : (x, y) \mapsto x^3y^3$.

Méthode 16.4.23 – Étude d'un point critique

Pour savoir si un point critique correspond à un extremum local, on étudiera le signe de la hessienne,

- soit directement par mises sous forme canonique successives de la forme quadratique $h \mapsto h^\top H_f(a)h$,
- soit par étude des valeurs propres de $H_f(a)$.

On peut alors conclure :

- Si $H_f(a)$ n'est pas positive (*i.e.* s'il existe une valeur propre strictement négative, ou s'il existe h tel que $h \mapsto h^\top H_f(a)h < 0$, alors f n'admet pas d'extremum local en a
- Si $H_f(a)$ est strcitement positive (si $\mathrm{Spec}(H_f(a)) \subset \mathbb{R}_+^*$), alors f admet un minimum local.

On ne peut pas conclure de cette façon si ces conditions ou les conditions symétriques (pour un maximum) ne sont pas satisfaites, à savoir dans le cas où $0 \in \mathrm{Spec}(H_f(a))$.

Corollaire 16.4.24 – Explicitation pour $n = 2$

Soit Ω un ouvert de $\mathbb{R}^2$, $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$, et a un point critique de f. Alors :

- Si $\det(H_f(a)) > 0$ et $\mathrm{tr}(H_f(a)) > 0$, alors f admet un minimum local strict en a
- Si $\det(H_f(a)) > 0$ et $\mathrm{tr}(H_f(a)) < 0$, alors f admet un maximum local strict en a
- si $\det(H_f(a)) < 0$, alors f n'admet pas d'extremum local en a.

Exemples 16.4.25

Étudier la nature du point critique $(0, 0)$ dans les 3 cas suivants :

1. $f_1(x, y) = y^4 + x^2y + x^2 + 2xy + 2y^2$
2. $f_2(x, y) = x^3y + x^2 - 4xy + y^2$
3. $f_3(x, y) = xy^2 + x^3 - x^2 + 2xy - 2y^2$.