Soient E et F deux $\mathbb{R}$-espaces vectoriels normés de dimension finie. Soient U un ouvert de E et $f : U \to F$ une application.
Dans ce chapitre, nous cherchons à généraliser la notion de dérivabilité et de développement limité à l'ordre un.
En général, on aura $E = \mathbb{R}^p$ et $F = \mathbb{R}$. Cependant nous pourrons nous ramener à cette situation.
Montrons comment se ramener à $E = \mathbb{R}^p$. Cette identification sera essentielle pour la notion de dérivées partielles. Soient $\mathcal{B} = (e_1, ..., e_p)$ une base de E et $\Phi : (x_1, ..., x_p) \to \displaystyle\sum_{i=1}^{p} x_i e_i$ qui est un isomorphisme de $\mathbb{R}^p$ dans E. On pose ainsi : $\tilde{f}(x_1, ..., x_p) = f(x_1 e_1 + ... + x_p e_p)$ pour $(x_1, .., x_p)$ dans $\Phi^{-1}(U) \subset \mathbb{R}^p$.
Par ailleurs Φ est une application continue (application linéaire entre espaces de dimension finie), donc $\Phi^{-1}(U)$ est un ouvert de $\mathbb{R}^p$. Ainsi $\tilde{f} : (x_1, ..., x_p) \to \tilde{f}(x_1, ..., x_p)$ est définie sur un ouvert de $\mathbb{R}^p$ à valeurs dans F. On pourra parfois noter $f(x_1, ..., x_p)$ au lieu de $\tilde{f}(x_1, ..., x_p)$.
Soit maintenant $\mathcal{U} = (u_1, ..., u_n)$ une base de F. On a donc : $f = \displaystyle\sum_{k=1}^{n} f_k u_k$ avec f_k dans $\mathcal{F}(U, \mathbb{R})$ pour tout k de $[\![1, n]\!]$. Ainsi on peut se ramener à l'étude de fonctions à valeurs réelles.
Par conséquent, $\tilde{f} = \displaystyle\sum_{k=1}^{n} \tilde{f}_k u_k$, permet de se ramener via les $\tilde{f}_k : (x_1, ..., x_p) \to \tilde{f}_k(x_1, ..., x_p)$ à l'étude d'une fonction définie sur un ouvert de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$.

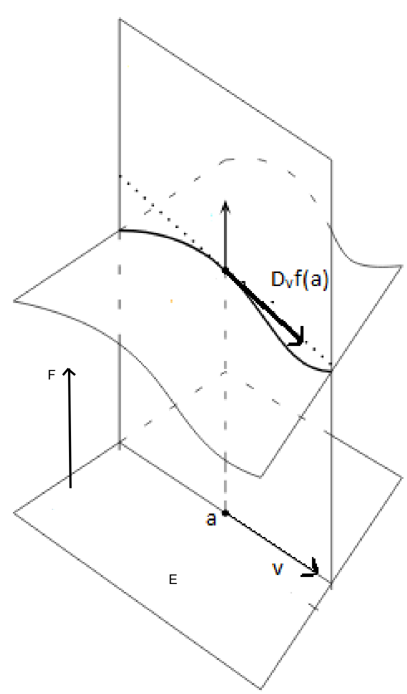
1 Dérivée selon un vecteur et dérivées partielles

1.1 Dérivées selon un vecteur

Définition 1.1.1 (Dérivée selon un vecteur) *Soient $a \in U$ et $v \in E$.*

- *On dit que f admet une dérivée selon le vecteur v au point a lorsque la fonction $\varphi_v : \left\{ \begin{array}{lcl} I & \to & F \\ t & \mapsto & f(a + tv) \end{array} \right.$ est dérivable en 0 (avec I un intervalle ouvert de $\mathbb{R}$ contenant 0).*
- *Le vecteur dérivé correspondant est appelé la dérivée de f suivant le vecteur v, et noté $D_v f(a)$:*

$$\lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} = \varphi_v'(0) = D_v f(a).$$



- *Si f admet en tout point a de U une dérivée suivant v, l'application $D_v f : a \mapsto D_v f(a)$ est l'application dérivée de f suivant v.*

Remarque 1.1.1 1. *On suppose v non nul. Comme U est un ouvert, alors $t \mapsto f(a+tv)$ est bien définie sur un intervalle ouvert contenant 0. En effet, il existe $r \in \mathbb{R}_+^*$, tel que : $B(a,r) \subset U$, donc si on pose $I = \left]-\frac{r}{\|v\|}, \frac{r}{\|v\|}\right[$, alors : $\forall t \in I,\ a+tv \in B(a,r) \subset U$ et donc $t \mapsto f(a+tv)$ est bien définie sur I.*

2. *Si $v=0$, alors $t \mapsto f(a+tv)$ est définie sur $\mathbb{R}$ et elle est constante (égale à $f(a)$). Dans ce cas, $D_v f(a) = 0$.*

Exemple 1.1.1 1. *Soient $f : \left\{ \begin{array}{lcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x^3+5y \end{array} \right.$, $a=(1,2)$ et $v=(-2,0)$. Déterminer $D_v f(a)$.*

On a : $\forall t \in \mathbb{R},\ f(a+tv) = f((1,2)+t(-2,0)) = f(1-2t,2) = (1-2t)^3 + 5\times 2 = 11 - 6t + 12t^2 - 8t^3$, puis : $\forall t \in \mathbb{R},\ \frac{d}{dt}(f(a+tv)) = -6 + 24t - 24t^2$, et enfin : $D_v f(a) = -6$.

2. *Dans $\mathcal{M}_n(\mathbb{R})$, déterminer $D_H \exp(0)$, avec $H \in \mathcal{M}_n(\mathbb{R})$.*

1.2 Dérivées partielles pour $E = \mathbb{R}^p$ et $F = \mathbb{R}$

Dans ce paragraphe, U est un ouvert de $\mathbb{R}^p$ et $f : (x_1, ..., x_p) \to f(x_1, ..., x_p)$ est une fonction définie sur U à valeurs dans $\mathbb{R}$.

Définition 1.2.1 (Application partielle) *Pour tout $a = (a_1, \ldots, a_p) \in U$ et $i \in [\![1,p]\!]$, l'application partielle en a selon la i-ème composante est définie par $f_{a,i} : t \mapsto f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p)$.*

Remarque 1.2.1 1. *L'application f étant définie sur un ouvert, alors pour tout i de $[\![1,n]\!]$, il existe un intervalle ouvert $]a_i - \varepsilon, a_i + \varepsilon[$ tel que $f_{a,i}$ soit définie sur cet intervalle. Ainsi $f_{a,i}$ est définie au voisinage de a_i.*

2. *Considérer la i-ème application partielle, revient à voir f comme fonction de sa i-ème variable.*
3. *Si f est continue en a, alors la i-ème application partielle est continue an a_i. En effet cela resulte de la composition de $t \mapsto (a_1, ..., t, ..., a_p)$ et de f qui sont continues.*
4. *ATTENTION, la continuité des applications partielles n'implique pas la continuité. Voici un contrexemple : Soit f la fonction définie sur $\mathbb{R}^2$ par $f(x,y) = 1 + \frac{2xy}{x^2+y^2}$ si $(x,y) \neq 0$ et $f(0,0) = 1$.*

   Soit $x \in \mathbb{R}$. Si $x \neq 0$, alors : $\forall y \in \mathbb{R},\ f(x,y) = 1 + \frac{2xy}{x^2+y^2}$ et donc $y \mapsto f(x,y)$ est continue. Par ailleurs on a : $\forall y \in \mathbb{R},\ f(0,y) = 1$ et donc $y \mapsto f(0,y)$ est continue. Par symétrie des variables, on a la continuité de $x \mapsto f(x,y)$ pour tout $y \in \mathbb{R}$.

   On a : $\forall t \in \mathbb{R}^*,\ f(t,t) = 1 + \frac{2t^2}{2t^2} = 2$. Ainsi $\lim_{t \to 0} f(t,t) = 2 \neq f(0,0)$, donc f n'est pas continue en $(0,0)$.

Définition 1.2.2 (Dérivées partielles en un point) 1. *La fonction f admet une dérivée partielle en a par rapport à la i-ème variable si l'application partielle $f_{a,i}$ admet une dérivée en a_i. Cette valeur est notée $\partial_i f(a)$ ou $\frac{\partial f}{\partial x_i}(a)$. Autrement dit :*

$$\frac{\partial f}{\partial x_i}(a) = \frac{d}{dt}[f(a_1, ..., t, ..., a_p)]\bigg|_{t=a_i} = \lim_{t \to a_i} \frac{f(a_1, ..., t, ..., a_p) - f(a_1, ..., a_i, ..., a_p)}{t - a_i} =$$
$$\lim_{h \to 0} \frac{f(a_1, ..., a_i + h, ..., a_p) - f(a_1, ..., a_i, ..., a_p)}{h}.$$

2. *Si f admet une dérivée partielle en tout point de U, l'application $\partial_i f = \frac{\partial f}{\partial x_i} : a \mapsto \frac{\partial f}{\partial x_i}(a)$ définie sur U est la i-ème dérivée partielle de f sur U.*

Exemple 1.2.1

1. *Soit* $f : \mathbb{R}^p \to \mathbb{R}$ *définie par* $(x_1, ..., x_p) \mapsto \alpha_1 x_1^2 + ... + \alpha_i x_i^2 + ... + \alpha_p x_p^2$. *Déterminer les dérivées partielles de* f *en tout point de* $a = (a_1, ..., a_p)$ *de* $\mathbb{R}^p$.

   $\forall i \in [\![1, p]\!], \ \dfrac{\partial f}{\partial x_i}(a_1, ..., a_p) = 2\alpha_i a_i.$

2. *Pour* $(x, y) \in \mathbb{R}^2$, *on pose*
   $f(x, y) = x(1 - y)$ *si* $x \leq y$ *et* $f(x, y) = y(1 - x)$ *sinon. Étudier la continuité et l'existence des dérivées partielles.*

3. *En identifiant* $\mathcal{M}_n(\mathbb{R})$ *à* $\mathbb{R}^{n^2}$, *étudier l'existence de dérivées partielles de la fonction* det.

4. *ATTENTION Soit* $f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & xy^2 \end{array} \right.$ *Déterminer* $\dfrac{\partial f}{\partial x}(y, x)$ *pour tout* (x, y) *de* $\mathbb{R}^2$.

   On a $\dfrac{\partial f}{\partial x} : (x, y) \mapsto y^2$, *ainsi :* $\dfrac{\partial f}{\partial x}(y, x) = x^2$.

   Dans ce dernier exemple, on a calculé $\dfrac{\partial f}{\partial x}$ *et on a évalué cette fonction en* (y, x). *Attention à ne pas confondre avec la dérivée suivant* x *de* $(x, y) \mapsto f(y, x) = yx^2$ *!*

Proposition 1.2.1 (Dérivée partielle nulle selon une variable) *Supposons que f soit définie sur un ouvert convexe. On suppose que $\frac{\partial f}{\partial x_i}$ existe et elle est nulle sur U :*

$$\forall (x_1, ..., x_p) \in U,\ \frac{\partial f}{\partial x_i}(x_1, ..., x_p) = 0.$$

Alors f ne dépend pas de x_i, c'est-à-dire il qu'il existe une fonction g de $p-1$ variables telle que : $\forall (x_1, ..., x_p) \in U, f(x_1, ..., x_p) = g(x_1, ...x_{i-1}, x_{i+1}, ..., x_p)$.

Démonstration : Soit $\varphi : t \mapsto f(x_1, ...x_{i-1}, t, x_{i+1}, ..., x_p)$ et I son domaine de définition. Comme U est convexe, alors I est un intervalle et il est ouvert car U est ouvert. L'hypothèse entraîne que $\varphi' = 0$ sur I et donc φ est constante sur I.

1.3 Dérivées partielles dans le cas général

On peut étendre notre définition de dérivées partielles à une fonction
$f : (x_1, ..., x_p) \mapsto (f_1(x_1, ..., x_p), f_2(x_1, ..., x_p), ..., f_n(x_1, ..., x_p))$ définie sur U à valeurs dans $\mathbb{R}^n$. Pour tout a de U, on note $\frac{\partial f}{\partial x_i}(a) = \left(\frac{\partial f_1}{\partial x_i}(a), ..., \frac{\partial f_n}{\partial x_i}(a)\right)$, lorsque toutes les fonctions f_i admettent des dérivées partielles selon la i-ème variable en a.

Nous allons même étendre cette définition à $f : U \to F$, avec U un ouvert d'un $\mathbb{R}$-espace vectoriel normé de dimension finie E et F également un $\mathbb{R}$-espace vectoriel normé de dimension finie.
On se fixe $(e_1, ..., e_p)$ une base de E.

Définition 1.3.1 (Dérivées partielles) *Soient $a = \sum_{k=1}^{p} a_k e_k \in U$ et $i \in [\![1, p]\!]$. On dit que f admet une dérivée partielle en a selon la i-ème coordonnée, lorsque $D_{e_i} f(a)$ existe (la dérivée en a selon le vecteur e_i) La i-ième dérivée partielle est notée $\partial_i f(a)$ ou $\frac{\partial f}{\partial x_i}(a)$.*

$$\frac{\partial f}{\partial x_i}(a) = D_{e_i} f(a) = \lim_{t\to 0} \frac{f(a + he_i) - f(a)}{h} = \lim_{h\to 0} \frac{\tilde{f}(a_1, ..., a_i + h, ..., a_p) - \tilde{f}(a_1, ..., a_i, ..., a_p)}{h} =$$

$$\lim_{h\to 0} \frac{f(a_1e_1 + ... + (a_i + h)e_i + ... + a_pe_p) - f(a_1e_1 + ... + a_ie_i + ... + a_pe_p)}{h} = \frac{\partial \tilde{f}}{\partial x_i}(a_1, ..., a_p).$$

2 Différentielle

2.1 Différentiabilité et exemples

Considérons d'abord une fonction $f : \mathbb{R} \to \mathbb{R}$ dérivable en a, cela équivaut à l'existence d'un développement limité à l'ordre un : $f(a + h) = f(a) + lh + o(h)$, avec $l = f'(a)$. Ainsi on peut obtenir l'approximation $f(a+h) \approx f(a)+lh$, pour h petit. On constate que $h \mapsto lh$ est une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$.
Grâce au chapitre précédent, nous pouvons généraliser cela si f est à valeur dans un espace vectoriel normé de dimension finie F. Nous avons les mêmes expressions, sauf qu'ici l et $f'(a)$ sont dans F.
Nous allons maintenant étendre cette notion de développement limité à une fonction $f : U \to F$ avec E et F des espaces vectoriels normés de dimension finie et U un ouvert de E.

Définition 2.1.1 ($o(h)$ et $O(h)$) *Soit $g : U \to F$, avec U un ouvert de E contenant 0.*

1. *On dit que $g(h) = o(\|h\|)$ si $\lim_{h\to 0} \frac{g(h)}{\|h\|_E} = 0$, autrement dit, il existe $\varepsilon : U \to F$, avec $\lim_{h\to 0} \varepsilon(h) = 0$ et $g(h) = \|h\|_E \varepsilon(h)$.*

2. *On dit que que* $g(h) = O(\|h\|)$*, si* $h \mapsto \dfrac{g(h)}{\|h\|_E}$ *est bornée, autrement dit, il existe* $K \in \mathbb{R}_+$ *tel que :* $\forall h \in U,\ \|g(h)\|_F \leq K\|h\|_E$.

Définition 2.1.2 (Différentiablité) 1. *On dit que* f *est différentiable en un point* a *de* U *lorsqu'il existe une application linéaire* $L_a \in \mathcal{L}(E, F)$ *telle que pour tout* h *tel que* $a + h \in U$ *:*

$$f(a+h) \;=\; f(a) + L_a(h) + \|h\|\varepsilon(h) = f(a) + L_a(h) + o(\|h\|).$$

où $\varepsilon(h) \underset{h \to 0}{\to} 0$.

2. *On dit que* f *est différentiable sur* U *si elle est différentiable en tout point de* U.

Remarque 2.1.1 *La notion de différentiabilité est indépendente de la norme choisie, car en dimension finie, les limites sont indépendantes de celle-ci.*

Proposition 2.1.1 (Unicité de la différentielle) *Si* f *est différentiable en* a*, alors il existe une unique application linéaire* $L_a : E \to F$ *telle que* $f(a+h) = f(a) + L_a(h) + o(\|h\|)$.

Démonstration :

Remarque 2.1.2 *(IMPORTANTE) Si on veut obtenir une propriété globale sur la différentielle, comme dans la preuve précédente (avec* $\forall u \in E,\ L_1(u) = L_2(u)$*), il est en général judicieux de fixer* $u \in E$ *et d'utiliser* tu*, avec* t *réel qui tend vers* 0 *dans la définition de la différentielle. En effet, cela permet d'utiliser la définition de la différentielle, qui est une propriété locale, pour montrer finalement une propriété globale.*

Définition 2.1.3 (Différentielle d'une application différentiable, développement limité) 1. *L'application linéaire* L_a *est appelée différentielle de* f *en* a *(ou application linéaire tangente à* f *en* a*) et notée* $df(a)$*, c'est un élément de* $\mathcal{L}(E, F)$.

2. $df(a)(h)$ *sera noté* $df(a).h$ *afin d'alléger les notations, mais cela ne désigne pas un produit scalaire.*

3. *Pour tout* h *tel que* $a + h \in U$*, on a donc :*

$$f(a+h) \;=\; f(a) + df(a)(h) + o(\|h\|).$$

C'est un développement limité d'ordre 1 *de* f *en* a.

Remarque 2.1.3 *Pour trouver une différentielle dans des espaces abstraits, il faut calculer* $f(a+h) = f(a) + ...$ *et identifier les termes dans lesquels « h n'apparaît qu'une fois » pour mettre en évidence la partie linéaire qui sera la différentielle.*
Nous verrons plus tard, dans le corollaire 2.3.1, un moyen simple de trouver une différentielle à l'aide des dérivées partielles dans les cas concrets. Cela sera commode pour les fonctions définies sur $\mathbb{R}^p$.

Proposition 2.1.2 (Différentiabilité et coordonnée) *Soient* $F_1, ..., F_p$ *des espaces vectoriels normés de dimension finie. Soit* $f = (f_1, ..., f_r) : U \to F_1 \times ... \times F_r$ *une application.* f *est différentiable en* $a \in U$ *si et seulement si pour tout* i *de* $[\![1, r]\!]$*, les applications* f_i *sont différentiable en* a.
Dans ce cas, on a : $df(a) : \left\{ \begin{array}{ccc} E & \to & F_1 \times ... \times F_r \\ h & \mapsto & (df_1(a).h, ..., df_r(a).h) \end{array} \right.$

Démonstration : On suppose que chaque f_i est différentiable en a. On a donc pour tout i de $[\![1,r]\!]$, on a : $f_i(a+h) = f_i(a) + L_{a,i}(h) + \|h\|\varepsilon_i(h)$, avec $L_{a,i}$ est dans $\mathcal{L}(E,F_i)$ et $\lim\limits_{h\to 0}\varepsilon_i(h) = 0$. Ainsi on a : $f(a+h) = f(a) + (L_{a,1}(h),...,L_{a,r}(h)) + \|h\|(\varepsilon_1(h),...,\varepsilon_r(h))$. L'application $h \mapsto (L_{a,1}(h),...,L_{a,r}(h))$ est linéaire et $\lim\limits_{h\to 0}(\varepsilon_1(h),...,\varepsilon_r(h)) = (0,...,0)$. Ainsi f est différentiable en a.

On suppose que f est différentiable en a. On a : $f(a+h) = f(a) + L_a(h) + \|h\|\varepsilon(h)$, avec $L_a = (L_{a,1},...,L_{a,r}) : E \to F_1 \times ... \times F_r$ linéaire et $\varepsilon = (\varepsilon_1,...,\varepsilon_r) : E \to F_1 \times ... \times F_r$, avec $\lim\limits_{0}\varepsilon = 0$. Donc pour tout i de $[\![1,r]\!]$, l'application $L_{a,i}$ est donc linéaire et $\lim\limits_{0}\varepsilon_i = 0$ et :
$f_i(a+h) = f_i(a) + L_{a,i}(h) + \|h\|\varepsilon_i(h)$, ce qui prouve la différentiablité de f_i en a.

Exemple 2.1.1

1. *Soit E un espace euclidien et f un endomorphisme autoadjoint de E et $u \in E$. Soit g définie sur E par $g : x \mapsto \frac{1}{2}(f(x)|x) - (u|x)$. Montrer que g est différentiable et préciser sa différentielle.*

Remarque 2.1.4 *Plus généralement. Soient E un espace vectoriel de dimension n et $B : E \times E \to \mathbb{R}$ une forme bilinéaire. Alors B est différentiable sur $E \times E$ et $dB(u_0,v_0) : (u,v) \mapsto B(u_0,v) + B(u,v_0)$.* *En effet, comme E est de dimension finie, alors B est continue et il existe $K \in \mathbb{R}_+$ tel que :*

$\forall (x,y) \in E \times E,\ |B(x,y)| \leq K\|x\|_\infty.\|y\|_\infty \leq K\|(x,y)\|^2$.

$E \times E$ est un ouvert de $E \times E$. Soit $a = (u_0,v_0)$ et $h = (u,v)$. On a :

$B(a+h) = B(u_0+u, v_0+v) = B(u_0,v_0) + B(u_0,v) + B(u,v_0) + B(u,v) =$

$B(a) + L(u,v) + B(u,v)$, *avec* $L : (u,v) \mapsto B(u_0,v) + B(u,v_0)$.

Montrons que L est linéaire sur $E \times E$. Soient $(x,y),(x',y') \in E \times E$ et $\alpha,\beta \in \mathbb{R}$. On a : $L\left(\alpha(x,y) + \beta(x',y')\right) = L\left(\alpha x + \beta x', \alpha y + \beta y'\right) = B(u_0, \alpha y + \beta y') + B(\alpha x + \beta x', v_0) = \alpha\left(B(u_0,y) + B(x,v_0)\right) + \beta\left(B(u_0,y') + B(x',v_0)\right) = \alpha L(x,y) + \beta L(x',y')$.

Par ailleurs : $|B(u,v)| \leq K\|(u,v)\|^2 = o(\|(u,v)\|) = (\|h\|)$ *au voisinage de* $(0,0)$.

Ainsi B est différentiable en (u_0,v_0) et : $dB(a) = dB((u_0,v_0)) = L$.

2. *On munit $\mathcal{M}_n(\mathbb{R})$ d'une norme $\|\ .\ \|$ d'algèbre (par exemple une norme subordonnée ou $\|\ .\ \|_2$) : $\forall A,B \in \mathcal{M}_n(\mathbb{R}),\ \|AB\| \leq \|A\|.\|B\|$. Nous avons vu dans le chapitre 7 que si $\|H\| < 1$, alors $I_n - H$ est inversible et $(I_n - H)^{-1} = \displaystyle\sum_{k=0}^{+\infty} H^k$.*
   Soit $f : M \mapsto M^{-1}$ définie sur l'ouvert $GL_n(\mathbb{R})$. Montrer l'existence et déterminer $df(I_n)$, puis en faire de même pour $df(M)$, avec $M \in GL_n(\mathbb{R})$.

3. *Montrer que* $f : \begin{cases} \mathbb{R}_n[X] & \to & \mathbb{R} \\ P & \mapsto & \displaystyle\int_0^1 \sin(tP(t))dt \end{cases}$ *est différentiable et donner sa différentielle.*

Proposition 2.1.3 (Différentiable implique continue) 1. *Soit $a \in U$. On suppose f différentiable en a. Alors f est continue en a.*

2. *On suppose f différentiable sur U, alors f est continue sur U.*

Démonstration : Nous ne montrons que la proposition *1.*, la deuxième en découlant immédiatement. Si f est différentiable en a, alors $f(a+h) = f(a) + df(a)(h) + \|h\|\varepsilon(h)$, avec $\lim_{h\to 0} \varepsilon(h) = 0$. Ainsi on a : $\lim_{h\to 0} \|h\|\varepsilon(h) = 0$. De plus $df(a)$ est une application linéaire entre E et F et comme ces espaces sont de dimension finie, alors $df(a)$ est continue. Ainsi $\lim_{h\to 0} df(a)(h) = df(a)(0) = 0$.
Par conséquent : $\lim_{h\to 0} f(a+h) = f(a)$ par somme et donc f est continue en a.

Proposition 2.1.4 (Différentielle d'une application constante) *On suppose f constante sur U, alors f est différentiable sur U et : $\forall a \in U,\ df(a) = 0$.*

Démonstration : Soit $a \in U$ et h tel que $a+h$ soit encore dans U. On a alors : $f(a+h) - f(a) = 0 = 0 + o(\|h\|)$. Or l'application $h \to 0$ est une application linéaire, donc par unicité de la différentielle, on a : $df(a)(h) = 0$, soit $df(a) = 0$.

Remarque 2.1.5 *ATTENTION, la réciproque est fausse. Nous verrons plus tard qu'elle est vraie si U est connexe par arcs.*

Proposition 2.1.5 (Différentielle d'une application linéaire) *Soit $g \in \mathcal{L}(E,F)$, alors g est différentiable sur U et : $\forall a \in U,\ dg(a) =$*

Démonstration : Soit $a \in U$ et h tel que $a+h$ soit encore dans U. Par linéarité de g, on a alors : $g(a+h) = g(a) + g(h) = g(a) + g(h) + o(\|h\|)$. Or l'application $h \to g(h)$ est une application linéaire, donc par unicité de la différentielle, on a : $dg(a)(h) = g(h)$, soit $dg(a) = g$.

Proposition 2.1.6 (Différentielle des fonctions d'une seule variable) *Soit $f : I \to F$ une fonction définie sur un intervalle ouvert I de $\mathbb{R}$, et $a \in I$.*
Alors : f est différentiable en a si et seulement si elle est dérivable en a.
Dans ce cas $df(a)$: et en particulier $df(a)(1) = f'(a)$.

Démonstration :

1. Supposons f dérivable en a. Ainsi f admet un développement limité à l'ordre un : $f(a+t) = f(a) + tf'(a) + o(t)$. Or l'application $t \mapsto tf'(a)$ est linéaire de $\mathbb{R}$ dans F, donc f est bien différentiable en a de différentielle $df(a) : t \mapsto tf'(a)$.
2. Supposons f différentiable en a, de différentielle $df(a) \in \mathcal{L}(\mathbb{R}, F)$. Ainsi par linéarité, on a : $\forall t \in \mathbb{R}, df(a)(t) = df(a)(t.1) = tdf(a)(1)$. On pose : $\alpha = df(a)(1)$ et donc on a : $df(a) : t \mapsto \alpha t$. Alors par définition de la différentiablité : $f(a+t) = f(a) + \alpha t + o(t)$, donc : $\dfrac{f(a+t) - f(a)}{t} = \alpha + o(1)$, puis f est dérivable en a, avec : $f'(a) = \alpha = df(a)(1)$.

2.2 Différentiabilité et dérivées partielles

Proposition 2.2.1 (Différentiable implique existence de dérivées selon tout vecteur) *Si f est différentiable en a, alors pour tout vecteur v de E la fonction f admet une dérivée en a selon v et :*

$$D_v f(a) \ = df(a)(v).$$

Démonstration :

Remarque 2.2.1 *La réciproque est fausse : une fonction peut admettre une dérivée selon n'importe quel vecteur sans pour autant être différentiable. Contre-exemple : on considère*

$$f : (x,y) \mapsto \begin{cases} \dfrac{x^3+y^3}{x^2+y^2} & si \quad (x,y) \neq (0,0) \\ 0 & si \quad (x,y) = (0,0) \end{cases} \quad \textit{définie sur } \mathbb{R}^2.$$

1. *Soit $v = (h,k) \in \mathbb{R}^2$. Déterminer $D_{(h,k)}f(0,0)$, la dérivée en $a = (0,0)$ de f selon le vecteur $v = (h,k)$.*
2. *f est-elle différentiable en $(0,0)$?*

1. *On suppose $(h,k) \neq (0,0)$. On a : $\forall t \in \mathbb{R}^*,\ f(a+tv) = f(th,tk) = \dfrac{t^3(h^3+k^3)}{t^2(h^2+k^2)} = t\dfrac{h^3+k^3}{h^2+k^2}$, ceci est encore vrai si $t = 0$, car $f(0,0) = 0$. Ainsi : $D_{(h,k)}f(0,0) = \dfrac{d}{dt}f(a+tv)\Big|_{t=0} = \dfrac{h^3+k^3}{h^2+k^2}$.*
   Par ailleurs $D_{(0,0)}f(0,0) = 0$, grâce à la remarque 1.1.1.

2. *Si f est différentiable en $(0,0)$, alors grâce à la proposition précédente :*
$\forall (h,k) \in \mathbb{R}^2,\ df\left((0,0)\right).(h,k) = D_{(h,k)}f(0,0)$. *Ainsi l'application* $u : (h,k) \mapsto D_{(h,k)}f(0,0)$ *est linéaire, par linéarité de la différentielle. Or* $u(1,0) = u(0,1) = 1$, *mais* $u(1,1) = 1 \neq u(1,0) + u(0,1)$, *ce qui contredit la linéarité. Ainsi f n'est pas différentiable en $(0,0)$.*

Exemple 2.2.1 1. $f : \left\{ \begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \|x\| \end{array} \right.$ *est-elle différentiable en 0 ?*

2. *Soient $f, g \in \mathcal{F}(\mathbb{R}^n, \mathbb{R})$ différentiables et $M = \max(f,g)$.*
On note $F = \{x \in \mathbb{R}^n,\ f(x) = g(x)\}$, $U = \{x \in \mathbb{R}^n,\ f(x) > g(x)\}$ et $V = \{x \in \mathbb{R}^n,\ f(x) < g(x)\}$.
Il est clair que sur l'ouvert $U = (f-g)^{-1}(\mathbb{R}_+^*)$, la fonction M est différentiable, car : $\forall x \in U,\ M(x) = f(x)$. On a la même chose sur V.
Soit $a \in F$. Montrer que M est continue en a, puis donner une condition nécessaire et suffisante pour que M soit différentiable en a.

Réciproquement, si $df(a) = dg(a) = L$, alors il existe $\varepsilon_1 : E \to \mathbb{R}$ et $\varepsilon_2 : E \to \mathbb{R}$ telles que $\lim\limits_0 \varepsilon_1 = \lim\limits_0 \varepsilon_2 = 0$ et $f(a+h) = \underbrace{f(a)}_{=M(a)} + L(h) + \|h\|\varepsilon_1(h)$ et $g(a+h) = \underbrace{g(a)}_{M(a)} + L(h) + \|h\|\varepsilon_2(h)$. Ainsi on a : $M(a+h) = M(a) + L(h) + \|h\| \underbrace{\max(\varepsilon_1(h), \varepsilon_2(h))}_{\substack{\to \\ h\to 0}\, 0}$. Ce qui prouve la différentiablité de M en a.

Corollaire 2.2.1 (Différentiable implique existence de dérivées partielles) *Soit $\mathcal{B} = (e_1, ..., e_p)$ une base de E.*
On suppose que f est différentiable en $a \in U$. Alors pour tout j de $[\![1,p]\!]$, $\dfrac{\partial f}{\partial x_j}(a) = \partial_j f(a)$ existe et :

$$\frac{\partial f}{\partial x_j}(a) = \partial_j f(a) = df(a)(e_j).$$

Démonstration : On a grâce à la proposition précédente et la définition de dérivée partielle : $df(a)(e_j) = D_{e_j}f(a) = \partial_j f(a)$.

Remarque 2.2.2 *(IMPORTANTE) S'il existe $i \in [\![1,p]\!]$ tel que $\dfrac{\partial f}{\partial x_j}(a)$ n'existe, pas, alors f n'est pas différentiable en a.*
Si on reprend l'exemple 1.2.1, alors $f(x,y) = x(1-y)$ si $x \leq y$ et $f(x,y) = y(1-x)$ sinon, n'est pas différentiable sur $\Delta = \{(x,x),\ x \in \mathbb{R}\}$.

Corollaire 2.2.2 (Expression de la différentielle à l'aide des dérivées partielles) *1. Soit $\mathcal{B} = (e_1, ..., e_p)$ une base de E.*

On suppose que f est différentiable en $a \in U$. Alors pour tout $h = \sum_{j=1}^{p} h_j e_j \in E$, on a :

$$df(a).h = \sum_{j=1}^{p} h_j \partial_j f(a) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a).$$

2. *On se place dans le cas particulier où $E = \mathbb{R}^p$. On suppose f différentiable en $a \in U$, alors :*

$$\forall (h_1, ..., h_p) \in \mathbb{R}^p,\ df(a).(h_1, ..., h_p) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a).$$

Démonstration :

Pour le deuxième point, appliquer le premier point à la base canonique de $\mathbb{R}^p$.

Remarque 2.2.3 *1. On se place dans le cas $E = \mathbb{R}^p$ et $F = \mathbb{R}$. Si f est différentiable en $a = (a_1, ..., a_p) \in U$, alors grâce au développement limité à l'ordre un, il existe une fonction ε telle que $\lim_a \varepsilon = 0$ et pour tout $x = (x_1, ..., x_p) \in \mathbb{R}^p$,*

$$f(x) = f(a) + \sum_{i=1}^{p} (x_i - a_i) \frac{\partial f}{\partial x_i}(a) + o_{x \to a}(\|x - a\|) = f(a) + \sum_{i=1}^{p} (x_i - a_i) \frac{\partial f}{\partial x_i}(a) + \|x - a\| \varepsilon(x).$$

On peut écrire ceci de la façon suivante : il existe une fonction ε_1 telle que $\lim_0 \varepsilon_1 = 0$ et pour $h = (h_1, ..., h_p) \in \mathbb{R}^p$, on a :

$$f(a + h) = f(a) + \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a) + o_{h \to 0}(\|h\|) = f(a) + \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a) + \|h\| \varepsilon_1(h).$$

Dans ce cas nous avons donc $f(a + h) \approx f(a) + \sum_{i=1}^{p} h_i \partial_i f(a)$, quand h est petit.

Lien avec la notation en physique : si jamais vous avez $y = f(x_1, ..., x_p)$, le développement limité à l'ordre un peut s'interpréter comme l'approximation au premier ordre de l'accroissement dy de la grandeur y lorsque l'on opère de petits accroissements $dx_1, ..., dx_p$, sur les grandeurs $x_1, ..., x_p$.

On a donc $dy = f(x_1 + dx_1, ..., x_p + dx_p) - f(x_1, ..., x_p) \approx \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(x_1, ..., x_p) dx_i$.

En mathématiques dx_i désigne plutôt l'application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}$ définie par $dx_i(h) = h_i$, avec $h = (h_1, ..., h_p)$. Ainsi on a l'égalité d'applications linéaires : $df(a) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a) dx_i$.

2. *ATTENTION, avoir des dérivées partielles n'implique ni la différentiabilité, ni même la continuité. Contre-exemple : soit $f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & \left\{ \begin{array}{ll} 1 + \dfrac{2xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 1 & \text{si } (x, y) = (0, 0) \end{array} \right. \end{array} \right.$. Montrer que f admet des dérivées partielles sur $\mathbb{R}^2$, mais cependant qu'elle n'est ni continue, ni différentiable sur $\mathbb{R}^2$.*

Définition 2.2.1 (Matrice jacobienne) *Soit $a \in U$ et on suppose f différentiable en a. On appelle matrice jacobienne de f en un point a de U, relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$ (de respectivement E et F), la matrice de la différentielle de f en a dans ces bases : $J_f(a) = Mat_{\mathcal{B},\mathcal{B}'}(df(a))$.*

Proposition 2.2.2 (Matrice jacobienne et dérivées partielles) *Nous reprenons le cadre de la définition précédente.*
On pose $\mathcal{B} = (e_1, ..., e_p)$ une base de E et $\mathcal{B}' = (e'_1, ..., e'_n)$ une base de F. Comme f est à valeurs dans F, on peut donc écrire $f = \sum_{i=1}^{n} f_i e'_i$, avec $f_1, ..., f_n$ des fonctions allant de U dans $\mathbb{R}$.
Alors chaque fonction f_i, avec $i \in [\![1, n]\!]$, admet des dérivées partielles en a suivant chaque coordonnée et

$$J_f(a) = [\partial_j f_i(a)]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} \in \mathcal{M}_{n,p}(\mathbb{R}).$$

Démonstration : Montrons d'abord que pour tout $i \in [\![1, n]\!]$ et $j \in [\![1, p]\!]$, la dérivée partielle $\partial_j f_i(a)$ existe.
Comme f est différentiable en a, alors elle admet des dérivées partielles suivant toutes les coordonnées en a. Or : $\partial_j f(a) = \lim_{t\to 0} \frac{1}{t}(f(a+te_j)-f(e_j))$. Une fonction vectorielle admet une limite en un point si et seulement si chaque fonction coordonnée admet une limite en ce point, ainsi $\lim_{t\to 0} \frac{1}{t}(f_i(a+te_j) - f_i(e_j))$ existe et donc $\partial_j f_i(a)$ est bien définie.
Grâce au corollaire précédent, on a :
$\forall j \in [\![1, p]\!],\ df(a).e_j = \frac{\partial f}{\partial x_1}(a) \times 0 + ... + \frac{\partial f}{\partial x_j}(a) \times 1 + ... + \frac{\partial f}{\partial x_n}(a) \times 0 = \frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(a) e'_i$, car $e_j = (0, ..., \underbrace{1}_{j\text{-ème position}}, ..., 0)$ et ainsi la matrice jacobienne de f en a est :

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_p}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_p}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \cdots & \frac{\partial f_n}{\partial x_p}(a) \end{pmatrix}.$$

Définition 2.2.2 (Matrice jacobienne quand $E = \mathbb{R}^p$ et $F = \mathbb{R}^n$) *Si $E = \mathbb{R}^p$ et $F = \mathbb{R}^n$, alors f est définie par : $\forall (x_1, .., x_p) \in U,\ f(x_1, ..., x_p) = (f_1(x_1, ..., x_p), ..., f_n(x_1, ..., x_p))$. Si f est différentiable en a, la matrice jacobienne de f en a est*

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_p}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_p}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \cdots & \frac{\partial f_n}{\partial x_p}(a) \end{pmatrix} = \left[\frac{\partial f_i}{\partial x_j}(a)\right]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} \in \mathcal{M}_{n,p}(\mathbb{R}).$$

Autrement dit les bases $\mathcal{B}$ et $\mathcal{B}'$ de la définition 2.2.1 sont les bases canoniques de E et F.

Exemple 2.2.2 *Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ différentiable telle que : $\forall x, y \in \mathbb{R}^n,\ \|f(x) - f(y)\| \geq \|x - y\|$.*

1. *Montrer que : $\forall x \in \mathbb{R}^n, \forall h \in \mathbb{R}^n,\ \|df(x)(h)\| \geq \|h\|$.*
2. *Soit $x \in \mathbb{R}^n$, en déduire que la matrice $J_f(x)$ est inversible.*

2.3 Fonctions de classe $\mathcal{C}^1$

Définition 2.3.1 (Fonction de classe $\mathcal{C}^1$) *On dit que f est de classe $\mathcal{C}^1$ sur U, lorsque f est différentiable sur U et l'application $df : \left\{ \begin{array}{lcl} U & \to & \mathcal{L}(E,F) \\ a & \mapsto & df(a) \end{array} \right.$ est continue.*

On note $\mathcal{C}^1(U,F)$ l'ensemble des applications de classe $\mathcal{C}^1$ allant de U dans F.

Exemple 2.3.1 *Soit $u \in \mathcal{L}(E,F)$. Alors u est de classe $\mathcal{C}^1$ sur E.*

Théorème 2.3.1 (Caractérisation des applications $\mathcal{C}^1$ par les dérivées partielles) *Les assertions suivantes sont équivalentes :*

1. *La fonction f est de classe $\mathcal{C}^1$ sur U.*
2. *Toutes les dérivées partielles de f sont définies et continues.*
   Autrement dit si on se fixe une base $\mathcal{B} = (e_1, ..., e_p)$ de E, alors les applications $\dfrac{\partial f}{\partial x_j}$ sont définies et continues sur U pour tout j de $[\![1, p]\!]$.

Démonstration : Admis.
Le plus compliqué est le sens retour. Donnons une démonstration dans un cas particulier avec $E = \mathbb{R}^2$, muni de $\|.\|_1$, et $F = \mathbb{R}$, afin d'avoir une idée de la preuve.
Soit $a = (a_1, a_2) \in U$ et $h = (h_1, h_2) \in \mathbb{R}^2$. On écrit $f(a+h) - f(a)$ sous la forme
$f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1 + h_1, a_2) - f(a_1, a_2)$.
Les égalités des accroissements finis donnent θ_1 et θ_2 dans $[0, 1]$ tels que :

$$f(a_1+h_1, a_2+h_2) - f(a_1+h_1, a_2) = h_2 \frac{\partial f}{\partial x_2}(a_1+h_1, a_2+\theta_2 h_2) \text{ et } f(a_1+h_1, a_2) - f(a_1, a_2) = h_1 \frac{\partial f}{\partial x_1}(a_1+\theta_1 h_1, a_2),$$

car les fonctions $t \mapsto f(a_1 + h_1, a_2 + th_2)$ et $t \mapsto f(a_1 + th_1, a_2)$ sont continues sur $[0, 1]$ et dérivables sur $]0, 1[$.
Soit $\varepsilon > 0$. La continuité des dérivées partielles en a donne une constante $\eta > 0$ telle que pour tout $h \in B(0, \eta)$ on a

$$\left| \frac{\partial f}{\partial x_1}(a+h) - \frac{\partial f}{\partial x_1}(a) \right| \leq \varepsilon \text{ et } \left| \frac{\partial f}{\partial x_2}(a+h) - \frac{\partial f}{\partial x_2}(a) \right| \leq \varepsilon.$$

Pour $h \in B(0, \eta)$, on a alors par l'inégalité triangulaire

$$\left| f(a+h) - f(a) - \frac{\partial f}{\partial x_1}(a) h_1 - \frac{\partial f}{\partial x_2}(a) h_2 \right|$$

$$\leq \left|\frac{\partial f}{\partial x_1}(a_1+\theta_1 h_1, a_2)h_1 - \frac{\partial f}{\partial x_1}(a)h_1\right| + \left|\frac{\partial f}{\partial x_2}(a_1+h_1, a_2+\theta_2 h_2)h_2 - \frac{\partial f}{\partial x_2}(a)h_2\right| \leq \varepsilon(|h_1|+|h_2|) = \varepsilon\|h\|_1.$$

On pose $\varepsilon_1(h) = \dfrac{f(a+h)-f(a)-\frac{\partial f}{\partial x_1}(a)h_1 - \frac{\partial f}{\partial x_2}(a)h_2}{\|h\|_1}$. Grâce à ce que nous venons de montrer, on a : $\lim\limits_{0} \varepsilon_1 = 0$. On a donc montré que

$$f(a+h)-f(a) = h_1\frac{\partial f}{\partial x_1}(a) + h_2\frac{\partial f}{\partial x_2}(a) + \|h\|_1\varepsilon_1(h).$$

Pour une généralisation à p variables, partie de : $f(a_1+h_1,...,a_p+h_p) - f(a_1,...,a_p) =$
$\sum_{i=1}^{p}(f(a_1+h_1,...,a_{i-1}+h_{i-1},a_i+h_i,a_{i+1},...,a_p) - f(a_1+h_1,...,a_{i-1}+h_{i-1},a_i,a_{i+1},...,a_p))$
et adapter la preuve précédente.

Remarque 2.3.1 *(IMPORTANTE) Dans des cas concrets, pour montrer qu'une application est de classe $\mathcal{C}^1$, il est souvent plus pratique de calculer des dérivées partielles plutôt que la différentielle. De plus pour montrer qu'une application est différentiable, il sera souvent plus simple de montrer qu'elle est de classe $\mathcal{C}^1$ à l'aide du calcul explicite des dérivées partielles.*

Corollaire 2.3.1 (Expression de la différentielle pour une application $\mathcal{C}^1$) *On suppose $f : U \to F$ de classe $\mathcal{C}^1$.*

1. *Soit $\mathcal{B} = (e_1,...,e_p)$ une base de E.*
   Alors pour tout $h = \sum_{j=1}^{p} h_j e_j \in E$, on a : $df(a).h = \sum_{j=1}^{p} h_j \partial_j f(a) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a)$.
2. *On se place dans le cas particulier où $E = \mathbb{R}^p$. On suppose f différentiable en $a \in U$. On a alors : $\forall (h_1,...,h_p) \in \mathbb{R}^p,\ df(a).(h_1,...,h_p) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a)$.*

Démonstration : Comme f est de classe $\mathcal{C}^1$ sur U, alors elle y est différentiable et on utilise ensuite le corollaire 2.2.2.

Remarque 2.3.2 *(IMPORTANT) Pour donner une expression explicite d'une différentielle dans des cas concrets, il peut être pratique de calculer les dérivées partielles pour montrer qu'elle est de classe $\mathcal{C}^1$, puis d'utiliser ensuite le corollaire précédent.*

Exemple 2.3.2

1. *$f : \left\{\begin{array}{lcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x_1,x_2) & \mapsto & e^{-x_1}(x_1^2 - \cos(x_1 x_2)) \end{array}\right.$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, car après calcul, ses dérivées partielles sont continues sur $\mathbb{R}^2$.*
2. *Toute fonction polynomiale sur $\mathbb{R}^p$ est de classe $\mathcal{C}^1$ et donc différentiable.*
3. *Donner une valeur approchée de $1,02^{1,99}$.*

   On pose $f : (x,y) \mapsto x^y = e^{y\ln(x)}$ définie sur l'ouvert $\mathbb{R}_+^* \times \mathbb{R}$.
   Soit $(x,y) \in \mathbb{R}_+^* \times \mathbb{R}$. On a : $\frac{\partial f}{\partial x}(x,y) = yx^{y-1}$ et : $\frac{\partial f}{\partial y}(x,y) = \ln(x)e^{y\ln(x)} = \ln(x)x^y$.
   Ainsi les dérivées partielles de f sont continue sur $\mathbb{R}_+^* \times \mathbb{R}$, donc f y est $\mathcal{C}^1$ et donc différentiable. On a donc :
   $\forall (h,k) \in \mathbb{R}^2,\ df((x,y)).(h,k) = h\frac{\partial f}{\partial x}(x,y) + k\frac{\partial f}{\partial y}(x,y) = hyx^{y-1} + k\ln(x)x^y$, puis pour (h,k) voisin de $(0,0)$, on a :
   $f((1,2)+(h,k)) = f(1,2) + df((1,2)).(h,k) + o(\|(h,k)\|) = f(1,2) + 2h + o(\|(h,k)\|) =$
   $1 + 2h + o(\|(h,k)\|)$.
   Enfin : $1,02^{1,99} = f(1.02, 1.99) = f((1,2)+(0.02,-0.01)) \approx 1 + 2\times 0.02 = 1.04$.
4. *Soit $f : \left\{\begin{array}{lcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \left\{\begin{array}{ll} xy\dfrac{x^2-y^2}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{array}\right. \end{array}\right.$ Montrer que f est continue, puis de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.*

Corollaire 2.3.2 ($\mathcal{C}^1$ et jacobien) *On suppose que f est de classe $\mathcal{C}^1$ sur U.*

1. *Soient $\mathcal{B} = (e_1, ..., e_p)$ une base de E et $\mathcal{B}' = (e'_1, ..., e'_n)$ une base de F. On a :*

$$J_f(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dfrac{\partial f_1}{\partial x_2}(a) & \cdots & \dfrac{\partial f_1}{\partial x_p}(a) \\ \dfrac{\partial f_2}{\partial x_1}(a) & \dfrac{\partial f_2}{\partial x_2}(a) & \cdots & \dfrac{\partial f_2}{\partial x_p}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(a) & \dfrac{\partial f_n}{\partial x_2}(a) & \cdots & \dfrac{\partial f_n}{\partial x_p}(a) \end{pmatrix}.$$

2. *Si $E = \mathbb{R}^p$ et $F = \mathbb{R}^n$, alors f est définie par :*
$\forall (x_1, .., x_p) \in U,\ f(x_1, ..., x_p) = (f_1(x_1, ..., x_p), ..., f_n(x_1, ..., x_p)).$

$$J_f(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dfrac{\partial f_1}{\partial x_2}(a) & \cdots & \dfrac{\partial f_1}{\partial x_p}(a) \\ \dfrac{\partial f_2}{\partial x_1}(a) & \dfrac{\partial f_2}{\partial x_2}(a) & \cdots & \dfrac{\partial f_2}{\partial x_p}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(a) & \dfrac{\partial f_n}{\partial x_2}(a) & \cdots & \dfrac{\partial f_n}{\partial x_p}(a) \end{pmatrix} = \left[\frac{\partial f_i}{\partial x_j}(a)\right]_{\substack{1\le i\le n\\ 1\le j\le p}} \in \mathcal{M}_{n,p}(\mathbb{R}).$$

Remarque 2.3.3 *Pour déterminer un jacobien, il sera pratique de montrer qu'une application est de classe $\mathcal{C}^1$ en calculant les dérivées partielles et ensuite d'utiliser les formules du corollaire précédent.*

3 Opérations sur les fonctions différentiables et de classe $\mathcal{C}^1$

3.1 Opérations sur les fonctions de classe $\mathcal{C}^1$

3.1.1 Fonctions définies sur $\mathbb{R}^p$ à valeurs dans $\mathbb{R}$

a) Opérations de base

Ici on considère $E = \mathbb{R}^p$ et $F = \mathbb{R}$. Soit U un ouvert de E.

Proposition 3.1.1 (Opérations sur les fonctions de classe $\mathcal{C}^1$) *Soient f et g deux applications de classe $\mathcal{C}^1$ sur U.*

1. *$f+g$ est de classe $\mathcal{C}^1$ sur U et : $\forall i \in [\![1,p]\!],\ \partial_i(f+g) = \partial_i f + \partial_i g$.*
2. *Pour tout λ dans $\mathbb{R}$, λf est de classe $\mathcal{C}^1$ sur U et : $\forall i \in [\![1,p]\!],\ \partial_i(\lambda f) = \lambda \partial_i f$.*
3. *fg est de classe $\mathcal{C}^1$ sur U et : $\forall i \in [\![1,p]\!],\ \partial_i(fg) = (\partial_i f)g + f(\partial_i g)$.*
4. *Si g ne s'annule pas sur U, f/g est de classe $\mathcal{C}^1$ sur U et : $\forall i \in [\![1,p]\!],\ \partial_i \left(\dfrac{f}{g}\right) = \dfrac{(\partial_i f)g - f(\partial_i g)}{g^2}$.*
5. *Si $\varphi : I \to \mathbb{R}$, est de classe $\mathcal{C}^1$ sur un intervalle I de $\mathbb{R}$ et telle que : $f(U) \subset I$, alors $\varphi \circ f$ est de classe $\mathcal{C}^1$ sur U et : $\forall i \in [\![1,p]\!],\ \partial_i(\varphi \circ f) = \partial_i(f) \times \varphi' \circ f$.*

Démonstration : Vient des propriétés usuelles des fonctions de la variable réelle.

Remarque 3.1.1 *$\mathcal{C}^1(U, \mathbb{R})$ est une $\mathbb{R}$-algèbre.*

b) Règle de la chaîne et composition

Proposition 3.1.2 (Règle de la chaîne) *Soit $x_1, \ldots, x_p : I \to \mathbb{R}$ de classe $\mathcal{C}^1$ sur un intervalle I de $\mathbb{R}$ telle que : $\forall t \in I,\ (x_1(t), \ldots, x_p(t)) \in U$, et on suppose f de classe $\mathcal{C}^1$ sur U.*
Alors $g : t \longmapsto f(x_1(t), \ldots, x_p(t))$ est de classe $\mathcal{C}^1$ sur I et :

$$\forall t \in I,\ g'(t) =$$

Démonstration : Soit $t_0 \in I$ et $a = (x_1(t_0), \ldots, x_p(t_0))$. Soit $r(h) = g(t_0+h) - g(t_0) - \sum_{i=1}^{p} x_i'(t_0)\partial_i f(a)h$.
En effectuant un développement limité à l'ordre un au voisinage de t_0, on a :
$\forall i \in [\![1,p]\!],\ x_i(t_0+h) = x_i(t_0) + hx_i'(t_0) + h\varepsilon_i(h)$ avec $\lim\limits_{0} \varepsilon_i = 0$, donc :
$(x_1(t_0+h), ..., x_p(t_0+h)) = (x_1(t_0), ..., x_p(t_0)) + h(x_1'(t_0), ..., x_p'(t_0)) + h(\varepsilon_1(h), ..., \varepsilon_p(h))$ et donc en posant $\epsilon_1 = (\varepsilon_1, ..., \varepsilon_p)$, on obtient :

$$\begin{aligned} g(t_0+h) = f\left(\underbrace{a + \sum_{i=1}^{p} x_i'(t_0)he_i + h\epsilon_1(h)}_{=k(h)}\right) &= g(t_0) + df(a)(k(h)) + \|k(h)\|\epsilon_2(k(h)) \\ &= g(t_0) + h\sum_{i=1}^{p} x_i'(t_0)\partial_i f(a) + hdf(a)\left(\epsilon_1(h)\right) + \|k(h)\|\epsilon_2(k(h)) \end{aligned}$$

On a utilisé la linéarité de $df(a)$ et ϵ_1 et ϵ_2 sont des fonctions de limite nulle en 0 .
On pose : $r(h) = hdf(a)(\epsilon_1(h)) + \|k(h)\|\epsilon_2(k(h))$.

Or $\|k(h)\| = O_{h\to 0}(\|h\|)$ donc $\|k(h)\|\epsilon_2(k(h)) = o_{h\to 0}(\|h\|)$, et par continuité de l'application linéaire df_a (on est dimension finie), on a bien $\lim\limits_{h\to 0} df(a)(\epsilon_1(h)) = 0$.

Ainsi : $r(h) = o_{h\to 0}(\|h\|)$ soit : $g(t_0+h) = g(t_0) + h\sum\limits_{i=1}^{p} x_i'(t_0)\partial_i f(a) + o_{h\to 0}(\|h\|)$, ce qui donne une développement limité à l'ordre un en t_0 et prouve que g est dérivable en t_0 avec : $g'(t_0) = \sum\limits_{i=1}^{p} x_i'(t_0)\partial_i f(a)$.

Pour conclure, on note que la formule obtenue $g' : t \mapsto \sum\limits_{i=1}^{p} x_i'(t)\partial_i f(x_1(t),\dots,x_p(t))$ est continue sur I par opérations et compositions, donc g est de classe $\mathcal{C}^1$ sur I.

Exemple 3.1.1 *1. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. Soit $g : t \mapsto f(t^2, t^3)$. Montrer que g est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et calculer g'.*

2. Donner une condition nécessaire et suffisante sur $\alpha \in \mathbb{R}$ pour que

$$f : \left\{ \begin{array}{lcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \left\{ \begin{array}{ll} \dfrac{|xy|^\alpha}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{array} \right. \end{array} \right. \quad \text{soit de classe } \mathcal{C}^1 \text{ sur } \mathbb{R}^2.$$

• si f est de classe $\mathcal{C}^1$, alors $g : t \mapsto f(t,t)$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Mais on a : $\forall t \in \mathbb{R}^*$, $\dfrac{g(t)-g(0)}{t} = \dfrac{|t|^{2(\alpha-1)}}{2t} = \left\{ \begin{array}{lll} t^{2\alpha-3} & \text{si} & t>0 \\ -t^{2\alpha-3} & \text{si} & t<0 \end{array} \right.$.
On a les même limites à gauche et à droite de 0 si et seulement si $2\alpha - 3 > 0$, soit $\alpha > 3/2$.
• Réciproquement, on suppose que : $\alpha > 3/2$. Pour les même raisons que précédement, $\varphi : x \mapsto |x|^\alpha$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et : $\forall x \in \mathbb{R}$, $\varphi'(x) = \varepsilon(x)\alpha|x|^{\alpha-1}$, avec $\varepsilon(x) = 1$ si $x \geq 0$ et $\varepsilon(x) = -1$, si $x < 0$.
Comme : $\forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, $f(x,y) = \dfrac{\varphi(x)\varphi(y)}{x^2+y^2}$, alors f est de classe $\mathcal{C}^1$ sur l'ouvert $\mathbb{R}^2 \setminus \{(0,0)\}$.
De plus, on a : $\forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, $\dfrac{\partial f}{\partial x}(x,y) = \varphi(y)\dfrac{\varphi'(x)(x^2+y^2) - 2x\varphi(x)}{(x^2+y^2)^2}$.
On a : $\forall x \in \mathbb{R}^*$, $\dfrac{f(x,0)-f(0,0)}{x} = 0$, donc quand x tend vers 0, on a : $\dfrac{\partial f}{\partial x}(0,0) = 0$.
Soit $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$. En posant $x = r\cos(\theta)$ et $y = r\sin(\theta)$, avec r dans $\mathbb{R}_+^*$, on a :
$\left|\dfrac{\partial f}{\partial x}(x,y)\right| \leq r^\alpha \left(\dfrac{\alpha r^{\alpha-1} \times r^2 + 2r \times r^\alpha}{r^4}\right) = (\alpha+2)r^{2\alpha-3}$.
Ainsi $\lim\limits_{(xy)\to(0,0)} \dfrac{\partial f}{\partial x}(x,y) = 0 = \dfrac{\partial f}{\partial x}(0,0)$, donc $\dfrac{\partial f}{\partial x}$ est aussi continue en $(0,0)$. Par symétrie des variables, on a le même résultat pour $\dfrac{\partial f}{\partial y}$.

*3. **(Identité d'Euler)***
Une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ est dite α-homogène ($\alpha \in \mathbb{R}$) si :
$\forall t \in \mathbb{R}_+^*,\ \forall (x,y) \in \mathbb{R}^2,\quad f(tx,ty) = t^\alpha f(x,y) \quad (*)$.

(a) Soit f de classe $\mathscr{C}^1$ et α-homogène.

i. Montrer que $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont $(\alpha-1)$-homogènes.

ii. Montrer l'identité d'Euler : $x\dfrac{\partial f}{\partial x}(x,y) + y\dfrac{\partial f}{\partial y}(x,y) = \alpha f(x,y)$.

(b) Réciproquement, soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$ vérifiant l'identité d'Euler et soit $(x,y) \in \mathbb{R}^2$. On définit la fonction $\phi :]0,+\infty[\to \mathbb{R}$ par : $\forall t > 0 : \phi(t) = f(tx,ty)$. Trouver une relation entre ϕ et ϕ' et en déduire que f est α-homogène.

Corollaire 3.1.1 (Règle de la chaîne pour la composition) *Soit U un ouvert de $\mathbb{R}^p$ et V un ouvert de $\mathbb{R}^m$. Soient $i \in [\![1,p]\!]$ et $x_i : V \to \mathbb{R}$ de classe $\mathcal{C}^1$. On considère $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ et on pose $g : (u_1, ..., u_m) \to f\left(x_1(u_1, ..., u_m), ..., x_p(u_1, ..., u_m)\right)$. Alors la fonction g est de classe $\mathcal{C}^1$ sur V et :*

$$\forall j \in [\![1,m]\!],\ \frac{\partial g}{\partial u_j}(u_1, ..., u_m) =$$

Démonstration : Faisons la preuve pour $\dfrac{\partial g}{\partial u_1}$. Nous devons dériver $u_1 \mapsto f\left(x_1(u_1, ..., u_m), ..., x_p(u_1, ..., u_m)\right)$. Pour $i \in [\![1,p]\!]$, on pose $z_i : u_1 \mapsto x_i(u_1, ..., u_m)$. Nous devons donc dériver $u \mapsto f(z_1(u), ..., z_p(u))$. Grâce à la règle de la chaîne, nous avons ainsi : $\dfrac{\partial g}{\partial u_1} : u \mapsto z_1'(u)\partial_1 f(z_1(u), ..., z_p(u)) + ... + z_p'(u)\partial_p f(z_1(u), ..., z_p(u))$. Comme nous avons $z_i'(u) = \dfrac{\partial x_i}{\partial u_1}(u, ..., u_m)$, on obtient le résultat.

c) Changement de variable pour les équations aux dérivées partielles

Adaptons le corollaire 3.1.1 dans le cas $p = m = 2$

Proposition 3.1.3 (Composition dans le cas $\mathbb{R}^2$) *On suppose que U et V sont des ouverts de $\mathbb{R}^2$. Soient $f : U \to \mathbb{R}$ et $(x,y) : \begin{cases} V & \to & U \\ (u,v) & \mapsto & (x(u,v), y(u,v)) \end{cases}$ des fonctions de classe $\mathcal{C}^1$.*
Alors, la fonction $g : (u,v) \mapsto f(x(u,v), y(u,v))$ est de classe $\mathcal{C}^1$ sur V et

$$\frac{\partial g}{\partial u}(u,v) =$$

$$\frac{\partial g}{\partial v}(u,v) =$$

Nous pouvons le reformuler ainsi :

$$\begin{aligned} \partial_1 g(u,v) &= \partial_1 x(u,v)\partial_1 f(x(u,v), y(u,v)) + \partial_1 y(u,v)\partial_2 f(x(u,v), y(u,v)) \\ \partial_2 g(u,v) &= \partial_2 x(u,v)\partial_1 f(x(u,v), y(u,v)) + \partial_2 y(u,v)\partial_2 f(x(u,v), y(u,v)). \end{aligned}$$

ou matriciellement :

$$\begin{pmatrix} \dfrac{\partial g}{\partial u} \\ \dfrac{\partial g}{\partial v} \end{pmatrix} =$$

Remarque 3.1.2 *(IMPORTANTE) Cette dernière proposition sert à effectuer des changements de variables pour résoudre des équations aux dérivées partielles. Il s'agit de poser $x = x(u,v)$ et $y = y(u,v)$ et d'étudier l'équation aux dérivées partielles vérifée par $g : (u,v) \mapsto f(x(u,v), y(u,v))$. Cependant il faudra bien veiller à ce que les changements de variables soient bijectifs pour avoir des équivalences entre les équations aux dérivées partielles et aussi pour revenir aux variables intiales.*

Voici les deux changements de variable les plus classiques :

- Le changement de variable linéaire : $x(u,v) = au + bv$ et $y(u,v) = cu + dv$.
  Dans ce cas on pose $g : (u,v) \mapsto f(au + bv, cu + dv)$. Nous avons :

$$\frac{\partial g}{\partial u}(u,v) =$$

$$\frac{\partial g}{\partial v}(u,v) =$$

Exemple 3.1.2 *Trouver toutes les fonctions f de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ telles que :*
$2\dfrac{\partial f}{\partial x} - 3\dfrac{\partial f}{\partial y} = f$ *(E), on posera le changement de variable $x = 2u + v$ et $y = -3u - v$.*

Remarque 3.1.3 *On utilisera le changement de variable linéaire pour des équations du type $\lambda\dfrac{\partial f}{\partial x} + \mu\dfrac{\partial f}{\partial y} = h(x,y)$, avec h une fonction et λ et μ des constantes. Le but, comme dans l'exemple précédent, est de trouver un changement de variable tel que $\dfrac{\partial g}{\partial u}(u,v) = k(u,v)$. Comme on a : $\dfrac{\partial g}{\partial u}(u,v) = a\dfrac{\partial f}{\partial x} + c\dfrac{\partial f}{\partial y}$, il suffit de choisir $a = \lambda$ et $c = \mu$. Nous voulons que le changement de variable soit bijectif, donc l'application linéaire $(u,v) \mapsto (au + bv, cu + dv)$ doit être bijective. Sa matrice dans la base canonique est $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et elle doit être inversible. Dans ce cas pour revenir aux variables initiales x et y quand on a trouvé g, il faut exprimer u et v en fonction de x et y. Ainsi on a : $\begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \Leftrightarrow \begin{pmatrix} u \\ v \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \dfrac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. Pour simplifier les calculs, nous faisons en sorte d'avoir $ad - bc = 1$.*

Reprenons l'exemple précédent. On doit prendre $a = 2$ et $c = -3$. Reste à choisir b et d tels que $2d + 3b = 1$. Nous avons pris $b = 1$ et $d = -1$ dans le changement de variable proposé, mais nous aurions aussi pu prendre $b = -1$ et $d = 2$. On aurait eu $x = 2u - v$ et $y = -3u + 2v$, ce qui aurait donné $u = 2x + y$ et $v = 3x + 2y$. Nous aurions eu aussi $\dfrac{\partial g}{\partial u} = g$, ce qui aurait donné des solutions de la même forme pour $g : (u, v) \mapsto \lambda(v)e^u$, avec $\lambda \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ et donc $f : (x, y) \mapsto \lambda(3x + 2y)e^{2x+y}$.
Nous voyons l'importance d'un changement de variable bijectif pour retrouver l'expression de f en fonction de x et y.

- Le changement de variable en coordonnées polaires : $x = r\cos(\theta)$ et $y = r\sin(\theta)$.

Dans ce cas on pose $g : (r, \theta) \mapsto f(r\cos(\theta), r\sin(\theta))$. Nous avons :

$$\frac{\partial g}{\partial r}(r, \theta) =$$

$$\frac{\partial g}{\partial \theta}(r, \theta) =$$

Exemple 3.1.3 *1. Exprimer $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ en fonction de $\dfrac{\partial g}{\partial r}$ et $\dfrac{\partial g}{\partial \theta}$.*

2. Trouver toutes les fonctions f de classe $\mathcal{C}^1$ définies sur $\mathbb{R}_+^\star \times \mathbb{R}$ solutions de $(E)\ \ y\dfrac{\partial f}{\partial x} - x\dfrac{\partial f}{\partial y} = y/x$.

Voici un autre exemple de changement de variable :

Exemple 3.1.4

Chercher les fonctions f de classe $\mathcal{C}^1$ sur $U = \{(x,y) \in \mathbb{R}^2,\ x - y > 0\}$ telles que
$(E)\ \dfrac{\partial f}{\partial x}(x,y) - \dfrac{\partial f}{\partial y}(x,y) + 3(x-y)f(x,y) = x^2 - y^2$ *(on posera $u = xy,\ v = x+y$).*

Montrer d'abord que $\phi : (x,y) \mapsto (xy, x+y)$ est une bijection de U sur un ensemble à préciser. Si $u = xy,\ v = x+y$, avec $(x,y) \in U$, alors x et y sont racines distinctes de $X^2 - vX + u$, donc $v^2 - 4u > 0$. On pose $V = \{(u,v) \in \mathbb{R}^2\ v^2 - 4u > 0\}$ et montrons que ϕ réalise une bijection de U dans V.

Soient $(u,v) \in V$. On a $\phi(x,y) = (u,v)$ si et seulement si x et y sont racines de $X^2 - vX + u$ et $x > y$, soit $x = \dfrac{v + \sqrt{v^2 - 4u}}{2}$ et $y = \dfrac{v - \sqrt{v^2-4u}}{2}$ et $\phi^{-1} : (u,v) \mapsto \left(\dfrac{v + \sqrt{v^2-4u}}{2}, \dfrac{v - \sqrt{v^2-4u}}{2}\right)$ qui est de classe $\mathcal{C}^1$ sur V.

Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$. On pose $g = f \circ \phi^{-1}$ de sorte que g soit de classe $\mathcal{C}^1$ sur V et : $\forall (x,y) \in U,\ f(x,y) = g(u,v) = g(xy, x+y)$.

On a : $\left\{\begin{array}{rcl} \dfrac{\partial f}{\partial x}(x,y) & = & y\dfrac{\partial g}{\partial u}(xy, x+y) + \dfrac{\partial g}{\partial v}(xy,x+y) \\ \dfrac{\partial f}{\partial y}(x,y) & = & x\dfrac{\partial g}{\partial u}(xy,x+y) + \dfrac{\partial g}{\partial v}(xy,x+y) \end{array}\right.$.

f vérifie (E) si et seulement si g vérifie $(y-x)\dfrac{\partial g}{\partial u} + 3(x-y)g = (x-y)v$, soit $(E') : \dfrac{\partial g}{\partial u} - 3g = -v$.

Une solution particulière de cette équation à $v \in \mathbb{R}$ fixé est $u \mapsto v/3$ et les solutions homogènes de la forme $u \mapsto A(v)e^{3u}$, avec $A \in \mathcal{F}(\mathbb{R},\mathbb{R})$. Ces solutions sont définies sur $]-\infty, v^2/4[$. On a donc : $\forall (u,v) \in V,\ g(u,v) = A(v)e^{3u} + v/3$. Par ailleurs $(-1,v)$ est dans V, donc : $\forall v \in \mathbb{R},\ A(v) = g(-1,v)e^3 - v/3$, donc A est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

Ainsi $\mathcal{S} = \{(x,y) \mapsto A(x+y)e^{3xy} - (x+y)/3,\ A \in \mathcal{C}^1(\mathbb{R})\}$.

3.1.2 Cas général

Dans ce paragraphe, E et F sont des espaces vectoriels normés de dimension finie et U est un ouvert de E. Nous généralisons les opérations du paragraphe précédent. On se fixe une base $\mathcal{B} = (e_1, ..., e_p)$ de E.

Proposition 3.1.4 (Opérations linéaires sur les fonctions de classe $\mathcal{C}^1$) *Soient f et g deux applications de classe $\mathcal{C}^1$ sur U.*

1. *$f + g$ est de classe $\mathcal{C}^1$ sur U et : $\forall i \in [\![1,p]\!],\ \partial_i(f+g) = \partial_i f + \partial_i g$.*
2. *Pour tout λ dans $\mathbb{R}$, λf est de classe $\mathcal{C}^1$ sur U et : $\forall i \in [\![1,p]\!],\ \partial_i(\lambda f) = \lambda \partial_i f$.*

Démonstration : Revenir à la définition des dérivées partielles via les coordonnées selon une base et utiliser la proposition 3.1.1.

Proposition 3.1.5 (Composition) *Soit V un ouvert de F et G un espace vectoriel normé de dimension finie. Soient $\mathcal{B}' = (e'_1, ..., e'_n)$ une base de F et $\mathcal{B}'' = (e''_1, ..., e''_m)$ une base de G.*
Soient $f : U \to V$ et $g : V \to G$ de classe $\mathcal{C}^1$.
On note $f = \sum_{i=1}^{n} f_i e'_i$ et $g = \sum_{k=1}^{m} g_k e''_k$. On a donc $g \circ f = \sum_{k=1}^{m}\left(g_k \circ \left(\sum_{i=1}^{n} f_i e'_i\right)\right) e''_k$

1. $\forall j \in [\![1,p]\!], \forall k \in [\![1,m]\!], \forall a \in U,\ \partial_j(g_k \circ f)(a) = \sum_{i=1}^{n} (\partial_i g_k)(f(a)).\partial_j f_i(a).$
2. $\forall j \in [\![1,p]\!], \forall a \in U,\ \partial_j(g \circ f)(a) = \sum_{i=1}^{n} \partial_j f_i(a).(\partial_i g)(f(a)) = \sum_{k=1}^{m}\left(\sum_{i=1}^{n}(\partial_i g_k)(f(a)).\partial_j f_i(a)\right) e''_k.$
3. *Autrement dit, matriciellement :*
$$\forall a \in U,\ J_{g\circ f}(a) = J_g(f(a))J_f(a).$$

Démonstration : Revenir à la définition des dérivées partielles via les coordonnées selon une base et utiliser la règle de la chaîne du corollaire 3.1.1

3.2 Opérations sur les différentielles

3.2.1 Combinaison linéaire

Proposition 3.2.1 (Combinaison linéaire d'applications différentiables) *L'ensemble des applications de U dans F, différentiables sur U, est un $\mathbb{R}$-espace vectoriel et pour toute fonction $f : U \to F$ et $g : U \to F$ et $\lambda, \mu \in \mathbb{R}$. Alors $\lambda f + \mu g$ est différentiable sur U et :*
$$d(\lambda f + \mu g) = \lambda df + \mu dg.$$

Démonstration : Soit $a \in U$. Alors :
$f(a+h) = f(a) + df(a).h + o(\|h\|)$, $g(a+h) = g(a) + dg(a).h + o(\|h\|)$, puis :
$(\lambda f+\mu g)(a+h) = (\lambda f+\mu g)(a)+(\lambda df(a).h + \mu dg(a).h)+o(\|h\|)$. Comme $\lambda df(a)+\mu dg(a)$ est linéaire, car $df(a)$ et $dg(a)$ le sont, alors $\lambda f + \mu g$ est différentiable en a et par unicité de la différentielle, on a : $d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a)$.

Corollaire 3.2.1 (Combinaison linéaire d'applications différentiables et dérivées partielles) *La proposition 3.1.4 reste vraie pour des applications différentiables.*

Démonstration : Dans une base $\mathcal{B} = (e_1, ..., e_p)$ de E, on a : $\forall i \in [\![1, p]\!], \forall a \in U,\ \partial_i f(a) = df(a)(e_i)$ et utiliser cette relation dans la proposition précédente.

3.2.2 Composition avec une application multilinéaire

Proposition 3.2.2 (Composition d'applications différentiables avec une application bilinéaire) *Soient F_1 et F_2 deux espaces vectoriels normés de dimension finie et $B : F_1 \times F_2 \to G$ une application bilinéaire à valeurs dans un $\mathbb{R}$-espace vectoriel de dimension finie G, et f, g deux applications définies et différentiables sur U, à valeurs respectivement dans dans F_1 et F_2.*
Alors $B(f, g)$ est différentiable sur U et pour $a \in U$, on a :

$$d(B(f,g))(a) = B(df(a), g(a))+B(f(a), dg(a)),\ \textit{soit : } \forall h \in E,\ d(B(f,g))(a)(h) = B(df(a).h, g(a))+B(f(a), dg(a).h).$$

Démonstration : Soit $a \in U$. Alors

$$f(a+h) = f(a)+df(a).h+\|h\|\varepsilon_1(h), \quad g(a+h) = g(a)+dg(a).h\|h\|\varepsilon_2(h),\ \text{avec } \lim_{h\to 0} \varepsilon_1(h) = \lim_{h\to 0} \varepsilon_2(h) = 0.$$

$$\begin{aligned} B(f,g)(a+h) &= B(f(a+h), g(a+h)) = B(f(a) + df(a).h + \|h\|\varepsilon_1(h), g(a) + dg(a).h + \|h\|\varepsilon_2(h)) \\ &= B(f(a), g(a)) + \Big(B(f(a), dg(a).h) + B(df(a).h, g(a))\Big) + Q(h) \end{aligned}$$

avec : $Q(h) = \|h\| \left(B(f(a), \varepsilon_2(h)) + B(df(a).h, \varepsilon_2(h)) + B(\varepsilon_1(h), g(a)) + B(\varepsilon_1(h), dg(a).h) + B(df(a).h, dg(a).h) + \|h\|B(\varepsilon_1(h), \varepsilon_2(h))\right)$. Pour h dans E, on pose :
$\varepsilon(h) = B(f(a), \varepsilon_2(h)) + B(df(a).h, \varepsilon_2(h)) + B(\varepsilon_1(h), g(a))+$
$B(\varepsilon_1(h), dg(a).h) + B(df(a).h, dg(a).h) + \|h\|B(\varepsilon_1(h), \varepsilon_2(h))$.
Comme nous sommes en dimension finie, alors les applications linéaires $df(a), dg(a)$ et bilinéaires B sont continues, donc :
$\lim_{h\to 0} \varepsilon(h) = B(f(a), 0) + B(df(a)(0), 0) + B(0, g(a)) + B(0, dg(a)(0)) + B(df(a)(0), dg(a)(0)) + 0 = 0$, car $df(a)(0) = dg(a)(0) = 0$. Comme $h \mapsto B(f(a), dg(a).h) + B(df(a).h, g(a))$ est linéaire, on a :
$B(f,g)(a+h) = B(f(a), g(a)) + \Big(B(f(a), dg(a).h) + B(df(a).h, g(a))\Big) + \|h\|\varepsilon(h)$, qui est un développement limité d'ordre 1 de $B(f, g)$ en a , d'où le résultat.

Exemple 3.2.1 1. *Soit* $f : \left\{ \begin{array}{ccc} GL_n(\mathbb{R}) & \to & \mathcal{M}_n(\mathbb{R}) \\ M & \mapsto & M^{-1} \end{array} \right.$.

(a) *Montrer que f est de classe $\mathcal{C}^1$ sur l'ouvert $GL_n(\mathbb{R})$ de $\mathcal{M}_n(\mathbb{R})$. Que peut-on en déduire ?*
(b) *Que vaut $Id_{\mathcal{M}_n(\mathbb{R})} \times f$? En déduire df.*

2. *Pour $M \in \mathcal{M}_n(\mathbb{R})$ telle que $d^\circ \mu_M = n$. On pose* $f(M) = (tr(M), ..., tr(M^n))$.
   (a) *Montrer que f est différentiable et calculer sa différentielle.*
   (b) *Montre que $\mathrm{Im}\,(df(M))$ n'est inclus dans aucun hyperplan de $\mathbb{R}^n$. En déduire $\mathrm{rg}\,(df(M))$.*

   (a)

   (b) *Si $\mathrm{Im}\,(df(M))$ est inclus dans un hyperplan de $\mathbb{R}^n$, alors il existe $a_0, ..., a_{n-1} \in \mathbb{R}$, non tous nuls tels que $\forall (x_0, ..., x_{n-1}) \in \mathrm{Im}\,(df(M))$, $\sum_{i=0}^{n-1} a_i x_i = 0$, donc :*
   $\forall H \in \mathcal{M}_n(\mathbb{R}),\ 0 = \sum_{i=0}^{n-1} a_i(i+1)tr(HM^i) = tr\left(H \sum_{i=0}^{n-1} a_i(i+1)M^i\right)$ *et donc en prenant H^T au lieu de H, on a :*
   $\forall H \in \mathcal{M}_n(\mathbb{R}),\ tr\left(H^T \sum_{i=0}^{n-1} a_i(i+1)M^i\right) = 0$. *En se rappelant que $(A, B) \mapsto tr(A^T B)$ définit un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$ et donc :*
   $\forall H \in \mathcal{M}_n(\mathbb{R}),\ \left(H \middle| \sum_{i=0}^{n-1} a_i(i+1)M^i\right) = 0$, *donc $\sum_{i=0}^{n-1} a_i(i+1)M^i$ est dans $\mathcal{M}_n(\mathbb{R})^\perp = \{0\}$, donc $\sum_{i=0}^{n-1} a_i(i+1)M^i = 0$, donc $P = \sum_{i=0}^{n-1} a_i(i+1)X^i$ qui est non nul et de degré au plus $n-1$ annule M, ce qui est contradictoire avec $d^\circ(\mu_M) = n$. Ainsi on ne peut pas avoir*
   $\mathrm{rg}\,(df(M)) \leq n-1$ *(sinon on considère $(e_1, ..., e_p)$ une base de $\mathrm{Im}\,(df(M))$ que l'on complète en une base $(e_1, ..., e_n)$ de $\mathbb{R}^n$, puis $\mathrm{Im}\,(df(M))$ est inclus dans l'hyperplan $\mathrm{Vect}(e_1, ..., e_{n-1})$). Ainsi $\mathrm{rg}\,(df(M)) = n$.*

Proposition 3.2.3 (Composition d'applications différentiables avec une application multilinéaire) *Soient $F_1, ..., F_r$ des espaces vectoriels normés de dimension finie et $M : F_1 \times ... \times F_r \to G$ une application multilinéaire à valeurs dans un $\mathbb{R}$-espace vectoriel de dimension finie G, et $f_1, ..., f_r$ des applications définies et différentiables sur U, à valeurs respectivement dans dans $F_1, ..., F_r$.*
Alors $M(f_1, ..., f_r)$ est différentiable sur U et pour $a \in U$, on a :

$$d(M(f_1, ..., f_r))(a) \ = \sum_{i=1}^{r} M(f_1(a), ..., df_i(a), ..., f_r(a)) \text{ soit :}$$

$$\forall h \in E,\ d(M(f_1, ..., f_r))(a).h \ = \sum_{i=1}^{r} M(f_1(a), ..., df_i(a).h, ..., f_r(a)).$$

Démonstration : On reprend la preuve précédente en développant brutalement $M(f_1(a+h), ..., f_r(a+h))$.

3.2.3 Composition

Proposition 3.2.4 (Composition d'applications différentiables) *Soient V un ouvert de F, G un espace vectoriel normé de dimension finie, $f : U \to F$ et $g : V \to G$ différentiables telles que : $f(U) \subset V$. L'application $g \circ f$ est différentiable sur U et :*

$$\forall a \in U,\ d(g \circ f)(a) = dg(f(a)) \circ df(a),\ \textit{soit} : \forall a \in U, \forall h \in E,\ d(g \circ f)(a).h = dg(f(a))(df(a).h).$$

Démonstration : Posons $b = f(a)$. On a :

$$\begin{aligned} f(a+h) &= f(a) + df(a).h + \|h\|\varepsilon_1(h), \quad \text{avec } \lim_{h\to 0} \varepsilon_1(h) = 0, \\ g(b+k) &= g(b) + dg(b).k + \|k\|\varepsilon_2(k), \quad \text{avec } \lim_{k\to 0} \varepsilon_2(k) = 0 \end{aligned}$$

D'où :

$$g \circ f\,(a+h) \;=\; g \circ f(a) + dg_b(df(a).h + \|h\|\varepsilon_1(h)) + \|df(a).h + \|h\|\varepsilon_1(h)\|\varepsilon_2(df(a).h + \|h\|\varepsilon_1(h)).$$

Or : $dg(b)(df(a).h + \|h\|\varepsilon_1(h))) = dg(b) \circ df(a)(h) + \|h\| dg(b)(\varepsilon_1(h))$. Or par continuité de l'application linéaire $dg(b)$ (on est en dimension finie), on a : $\lim_{h\to 0} dg(b)(\varepsilon_1(h)) = dg(b)(0) = 0$. Ainsi on a : $dg(b)(df(a).h + \|h\|\varepsilon_1(h))) = (dg(b) \circ df(a)).(h) + o(\|h\|)$.

Comme $df(a)$ est continue (on est en dimension finie), il existe $C \in \mathbb{R}_+$ tel que :
$\forall h \in E, ;\ \|df(a).h\| \leq C\|h\|$. Ainsi :
$\|df(a).h + \|h\|\varepsilon_1(h)\| \times \|\varepsilon_2(df(a).h + \|h\|\varepsilon_1(h))\| \leq \|h\|(C + \|\varepsilon_1(h)\|) \times \|\varepsilon_2(df(a).h + \|h\|\varepsilon_1(h))\|$.
Or $\lim_{h\to 0} \varepsilon_2(df(a).h + \|h\|\varepsilon_1(h)) = 0$, par composition de limites donc
$\|df(a).h + \|h\|\varepsilon_1(h)\|\varepsilon_2(df(a).h + \|h\|\varepsilon_1(h))\| = o(\|h\|)$. Ainsi :
$g \circ f(a+h) \;=\; g \circ f\,(a) + (dg(b) \circ df(a)).h + o(\|h\|)$.
On reconnaît un développement limité d'ordre 1 de $g \circ f$ en , d'où le résultat.

Exemple 3.2.2 *Soient U un ouvert de E et $f : U \to \mathbb{R}$, une application différentiable sur U ne s'annulant jamais. Montrer que $1/f$ est différentiable sur U et déterminer $d(1/f)$.*

Corollaire 3.2.2 (Composition d'applications différentibles et dérivées partielles) *La proposition 3.1.5 reste valable pour des applications différentiables.*

Démonstration : La traduction matricielle de $d(g \circ f)(a) = dg(f(a)) \circ df(a)$ donne la relation suivante entre les matrices jacobiennes : $J_{g\circ f}(a) = J_g(f(a)).J_f(a)$. Nous avons ensuite les relations sur les dérivées partielles en détaillant les coefficients de ces matrices et en posant le produit matriciel à l'aide de ses coefficients.

3.3 Dérivée le long d'un arc et applications

Proposition 3.3.1 (Dérivée le long d'un arc) *Soient I un intervalle d'intérieur non vide de $\mathbb{R}$ et $\gamma : I \to U$ et $f : U \to F$. Si γ est dérivable en t_0 et si f est différentiable en $a = \gamma(t_0)$, alors $f \circ \gamma : I \to F$ est dérivable en t_0 et*

$$(f \circ \gamma)'(t_0) =$$

Si $E = \mathbb{R}^p$ et $\gamma = (x_1, ..., x_p)$, alors : $(f \circ \gamma)'(t_0) = \dfrac{d}{dt} f(x_1(t), ..., x_p(t))|_{t=t_0} =$

Démonstration : On applique le théorème de composition des applications différentiables. On a : $\gamma(t_0) \in U \subset E$, $df(\gamma(t_0)) \in \mathcal{L}(E, F)$, $\gamma'(t_0) \in E$. Plus précisément, la proposition 2.1.6 donne :

$$(f \circ \gamma)'(t_0) \;=\; d(f \circ \gamma)(t_0).1 \;=\; (df(\gamma(t_0)) \circ d\gamma(t_0)).1 \;=\; df(\gamma(t_0))\Big(d\gamma(t_0).1\Big) \;=\; df(\gamma(t_0))\Big(\gamma'(t_0)\Big).$$

Corollaire 3.3.1 (Dérivée le long de $t \mapsto a + th$) *Soient $a \in U$ et $h \in E$ tel que : $a + h \in U$. Alors :*
$\forall t \in [0, 1],\ \dfrac{d}{dt}(f(a + th)) =$

Si $E = \mathbb{R}^p$, $h = (h_1, ..., h_p)$, $a = (a_1, ..., a_p)$, alors :
$\forall t \in [0, 1],\ \dfrac{d}{dt}(f(a + th)) =$

Démonstration : on utilise la proposition précédente, avec $\gamma : t \mapsto a + th$ et $\gamma'(t) = h$.

Exemple 3.3.1 *Une application $f : E \to \mathbb{R}$ est dite convexe si :*
$\forall a, b \in E, \forall t \in [0, 1],\ f(ta + (1 - t)b) \leq tf(a) + (1 - t)f(b) \quad (*)$.
Soit $f : E \to \mathbb{R}$ de classe $\mathcal{C}^1$. Montrer que f est convexe si et seulement si : $\forall a, b \in E,\ df(b)(a - b) \leq f(a) - f(b) \ (*)$.

Théorème 3.3.1 (Caractérisation des fonctions constantes par la différentielle) *Soit U un ouvert connexe par arcs de E et $f : U \to F$ de classe $\mathcal{C}^1$. La fonction f est constante sur U si et seulement si $df = 0$.*

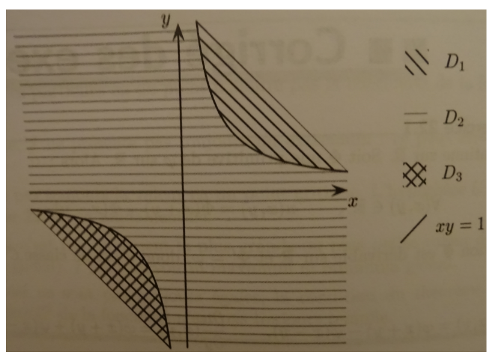
Démonstration : Si f est constante, nous avons déjà montré que $df = 0$.
Supposons que $df = 0$. Montrons le résultat dans le cas où U est convexe.

Traitons maintenant le cas général où U est connexe par arcs.
Soit $x \in U$ et $x_0 \in U$. Soit $\gamma : [0,1] \to U$ continue telle que $\gamma(0) = x_0$ et $\gamma(1) = x$. Soit $t_0 = \sup(A)$, avec $A = \{t \in [0,1],\ \forall u \in [0,t],\ f(\gamma(u)) = f(x_0)\}$. A est non vide ($t_0 \in A$) et majorée par 1, donc $\sup(A)$ est bien définie. Par ailleurs soit $(s_n)_{n\geq 0}$ une suite de A qui converge vers t_0. On a : $\forall n \in \mathbb{N},\ f(\gamma(s_n)) = f(x_0)$, puis par continuité de $f \circ \gamma$, on a : $f(\gamma(t_0)) = f(x_0)$. De plus U est un ouvert de E, donc il existe $\eta > 0$ tel que : $B(\gamma(t_0), \eta) \subset U$. Comme γ est continue, alors $\gamma^{-1}(B(\gamma(t_0), \eta))$ est un ouvert relatif de $[0,1]$ contenant t_0, donc il existe $\alpha \in \mathbb{R}_+^*$ tel que : $[t_0 - \alpha, t_0 + \alpha] \cap [0,1] \subset \gamma^{-1}(B(\gamma(t_0), \eta))$, soit : $\forall t \in [t_0 - \alpha, t_0 + \alpha] \cap [0,1],\ \gamma(t) \in B(\gamma(t_0), \eta)$. Or $B(\gamma(t_0), \eta)$ est un ouvert convexe sur lequel df est nulle, donc f y est constante grâce à ce qui précède. Ainsi : $\forall t \in [t_0 - \alpha, t_0 + \alpha] \cap [0,1],\ f(\gamma(t)) = f(\gamma(t_0)) = f(x_0)$. Si $t_0 < 1$, on a une contradiction de la maximalité de t_0, donc $t_0 = 1$, puis $f(x) = f(x_0)$. Ainsi f est constante sur U.

Exemple 3.3.2 *Soit* $f : (x,y) \mapsto \operatorname{Arctan}(x) + \operatorname{Arctan}(y) - \operatorname{Arctan}\left(\dfrac{x+y}{1-xy}\right)$.

1. *Montrer que f est de classe $\mathcal{C}^1$ sur un domaine D que l'on précisera et que l'on représentera graphiquement.*
2. *Exprimer $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sur D, puis simplifier l'expression de f.*

1. *Sur l'ouvert $D = \{(x,y) \in \mathbb{R}^2,\ xy \neq 1\}$, f est de classe $\mathcal{C}^1$ par opérations/compositions.*



2. *Soit $(x,y) \in D$. On a :*
$\dfrac{\partial f}{\partial x}(x,y) = \dfrac{1}{1+x^2} - \dfrac{(1-xy)-(x+y)(-y)}{(1-xy)^2} \times \dfrac{1}{1+\frac{(x+y)^2}{(1-xy)^2}} = \dfrac{1}{1+x^2} - \dfrac{1+y^2}{(1-xy)^2+(x+y)^2} = \dfrac{1}{1+x^2} - \dfrac{1+y^2}{1+x^2+y^2+x^2y^2} = \dfrac{1}{1+x^2} - \dfrac{1+y^2}{(1+x^2)(1+y^2)} = 0.$
De même : $\dfrac{\partial f}{\partial y} = 0$ *et donc* $df(x,y) = 0$.
On constates que D est la réunion disjointe de trois ouverts connexes par arcs :
$D_1 = \left\{(x,y) \in (\mathbb{R}_+^*)^2,\ y > \dfrac{1}{x}\right\}$, $D_2 = \left\{(x,y) \in \mathbb{R}^2,\ xy < 1\right\}$ *et* $D_3 = \left\{(x,y) \in (\mathbb{R}_-^*)^2,\ y < \dfrac{1}{x}\right\}$.
Grâce au théorème précédent, f est constante sur chacun de ces ouverts (ils sont connexes par arcs).
Pour $x \in]1, +\infty[$, on a $(x,1)$ dans D_1 et :
$f(x,1) = \operatorname{Arctan}(x) + \operatorname{Arctan}(1) - \operatorname{Arctan}\left(\dfrac{x+1}{1-x}\right) \underset{x \to +\infty}{\to} \dfrac{\pi}{2} + \dfrac{\pi}{4} - \operatorname{Arctan}(-1) = \pi.$
Ainsi : $\forall (x,y) \in D_1,\ f(x,y) = \pi$.
Comme $f(0,0) = 0$, *alors :* $\forall (x,y) \in D_2,\ f(x,y) = 0$.
Enfin : $\forall (x,y) \in D_3,\ f(x,y) = -f\underbrace{(-x,-y)}_{\in D_1} = -\pi$.

4 Applications de classe $\mathcal{C}^k$

4.1 Définitions et opérations

U désignera un ouvert de $\mathbb{R}^p$.

Définition 4.1.1 (Dérivées partielles successives) *Soient $k \in \mathbb{N}^*$, $j_1, ..., j_k \in [\![1, p]\!]$ et une application $f : \left\{ \begin{array}{ccc} U & \to & \mathbb{R} \\ (x_1, ..., x_p) & \mapsto & f(x_1, ..., x_p) \end{array} \right.$. On dit que f admet une dérivée k-ème successivement selon les variables $x_{j_1}, ..., x_{j_k}$, lorsque $\dfrac{\partial}{\partial x_{j_k}} \left(\dfrac{\partial}{\partial x_{j_{k-1}}} \left(... \left(\dfrac{\partial f}{\partial x_{j_1}} \right) ... \right) \right)$ existe sur U.*
Dans ce cas, la fonction $\dfrac{\partial}{\partial x_{j_k}} \left(\dfrac{\partial}{\partial x_{j_{k-1}}} \left(... \left(\dfrac{\partial f}{\partial x_{j_1}} \right) ... \right) \right)$ est noté $\dfrac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} ... \partial x_{j_1}}$ ou $\partial_{j_k} \partial_{j_{k-1}} ... \partial_{j_1} f$ ou $\partial_{j_1, ..., j_k} f$.
Une telle fonction s'appelle dérivée partielle d'ordre k.

Définition 4.1.2 (Applications $\mathcal{C}^k$)

1. *Une application $f : U \to F$ est dite de classe $\mathcal{C}^k$ si toutes ses dérivées partielles d'ordre k existent et sont continues sur U.*
2. *On note $\mathcal{C}^k(U, F)$ l'ensemble des applications de classe $\mathcal{C}^k$ de U dans F.*
3. *Une application est dite de classe $\mathcal{C}^\infty$ sur U si elle est dans $\mathcal{C}^k(U, F)$ pour tout k de $\mathbb{N}$.*
4. *On note $\mathcal{C}^\infty(U, F)$ l'ensemble des applications de classe $\mathcal{C}^\infty$ de U dans F.*

Exemple 4.1.1 *Les fonctions polynomiales sont dans $\mathcal{C}^\infty(\mathbb{R}^p, \mathbb{R})$.*

Proposition 4.1.1 (Structure de $\mathcal{C}^k(U, F)$) *Soit $k \in \mathbb{N} \cup \{\infty\}$. Alors :*

1. *L'ensemble $\mathcal{C}^k(U, F)$ est un espace vectoriel : $\forall f, g \in \mathcal{C}^k(U, F), \forall \lambda, \mu \in \mathbb{R},\ \lambda f + \mu g \in \mathcal{C}^k(U, F)$.*
2. *Si $F = \mathbb{R}$, alors l'espace vectoriel $\mathcal{C}^k(U, \mathbb{R})$ possède de plus une structure d'anneau, donc de $\mathbb{R}$-algèbre. On a donc : $\forall f, g \in \mathcal{C}^k(U, \mathbb{R}),\ fg \in \mathcal{C}^k(U, \mathbb{R})$.*
3. *La composition d'applications de classe $\mathcal{C}^k$ est de classe $\mathcal{C}^k$.*

Démonstration : Admis.

Exemple 4.1.2

1. *Soit $f(x, y) = \begin{cases} \dfrac{\mathrm{ch}\,(xy) - \cos(xy)}{x^2 y^2} & \text{si } xy \neq 0 \\ 1 & \text{si } xy = 0 \end{cases}$. Montrer que f est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$.*

2. *Déterminer les fonctions de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ telles que : $\dfrac{\partial^2 f}{\partial y \partial x} = 0$.*

3. *Soient* $k \in \mathbb{N}\backslash\{0,1\}$ *et* $f : \mathbb{R} \to \mathbb{R}$ *de classe* $\mathcal{C}^k$. *On pose* $F : \left\{ \begin{array}{lll} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \left\{ \begin{array}{lll} \dfrac{f(y)-f(x)}{y-x} & si & x \neq y \\ f'(x) & si & x = y \end{array} \right. \end{array} \right.$

Montrer que F *est de classe* $\mathcal{C}^{k-1}$ *sur* $\mathbb{R}^2$.

Pour $x \neq y$, *on a :* $F(x,y) = \dfrac{1}{y-x}\displaystyle\int_x^y f'(t)dt = \int_0^1 f'((1-s)x+sy)ds$, *en posant* $t = (1-s)x + sy$, *soit* $s = \dfrac{t-x}{y-x}$. *Cette dernière égalité reste valable pour* $x = y$.

Montrons que $\dfrac{\partial F}{\partial x}$ *existe.*

On fixe $y \in \mathbb{R}$.

- *Soit* $s \in [0,1]$. $x \mapsto f'((1-s)x+sy)$ *est de classe* $\mathcal{C}^1$ *sur* $\mathbb{R}$.
- *Soit* $x \in \mathbb{R}$. $s \mapsto f'((1-s)x+sy)$ *est continue par morceaux et donc intégrable sur le segment* $[0,1]$.
- *Soit* $[a,b] \subset \mathbb{R}$. *La fonction* $(x,s) \mapsto \dfrac{\partial}{\partial x}(f'((1-s)x+sy)) = (1-s)f''((1-s)x+sy)$ *est continue sur le compact* $[a,b]\times[0,1]$, *donc il existe* $M \in \mathbb{R}_6+$ *tel que :* $\forall (x,s) \in [a,b]\times[0,1]$, $|f''((1-s)x+sy)| \leq M$. *De plus* $s \mapsto M$ *est continue par morceaux et intégrable sur* $[0,1]$.

Ainsi $x \mapsto F(x,y)$ *est de classe* $\mathcal{C}^1$ *sur* $\mathbb{R}$ *et :* $\forall (x,y) \in \mathbb{R}^2$, $\dfrac{\partial F}{\partial x}(x,y) = \displaystyle\int_0^1 (1-s)f''((1-s)x+sy)ds$.

Montrons maintenant que $\dfrac{\partial F}{\partial x}$ *est continue sur* $\mathbb{R}^2$.

- *Soit* $s \in [0,1]$. $(x,y) \mapsto (1-s)f''((1-s)x+sy)$ *est continue sur* $\mathbb{R}^2$.
- *Soit* $(x_0,y_0) \in \mathbb{R}^2$. *Ainsi* $\overline{B}((x_0,y_0),1)$ *est un voisinage de* (x_0y_0). *Sur le compact* $\overline{B}((x_0,y_0),1)\times[0,1]$ *la fonction* $((x,y),s) \mapsto (1-s)f''((1-s)x+sy)$ *est continue, donc il existe* $K \in \mathbb{R}_+$ *tel que :* $\forall ((x,y),s) \in \overline{B}((x_0,y_0),1)\times[0,1]$, $|(1-s)f''((1-s)x+sy)| \leq K$ *et* $s \mapsto K$ *est continue par morceaux et intégrable sur* $[0,1]$.

.

Ainsi $(x,y) \mapsto \dfrac{\partial F}{\partial x}(x,y) = \displaystyle\int_0^1 (1-s)f''((1-s)x+sy)ds$ *est continue sur* $\mathbb{R}^2$.

On a le même résultat pour $\dfrac{\partial F}{\partial y}$.

Pour conclure on raisonne par récurrence sur k.
Le cas $k=2$ *a été traité.*
Soit $k \geq 2$ *et on suppose le résultat vrai pour* k. *Soit* $f : \mathbb{R} \to \mathbb{R}$ *de classe* $\mathcal{C}^{k+1}$, *on montre que* F *est de classe* $\mathcal{C}^k$ *de la même manière, en montrant que pour* $i,j \in \mathbb{N}$ *tels que* $i+j=k$ *alors :* $\dfrac{\partial^{i+j}F}{\partial x^i \partial y^j}(x,y) = \displaystyle\int_0^1 (1-s)^i s^j f^{(i+j+1)}((1-s)x+sy)ds$.

Théorème 4.1.1 (Théorème de Schwarz) *On suppose* $f : U \to F$ *de classe* $\mathcal{C}^2$ *et* $i,j \in [\![1,p]\!]$. *Alors :*

$$\forall a \in U, \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$$

Démonstration : (HORS PROGRAMME) Sans perte de généralité, on peut supposer que $n=2$. Notons les variables (x,y). Soit $(x_0,y_0) \in U$ et montrons que :
$\dfrac{\partial^2 f}{\partial x \partial y}(x_0,y_0) = \dfrac{\partial^2 f}{\partial y \partial x}(x_0,y_0)$.
Pour $h \in \mathbb{R}$ petit pour avoir $(x_0+h, y_0+h) \in U$, on pose :

$$F(h) = f(x_0+h, y_0+h) - f(x_0+h, y_0) - f(x_0, y_0+h) - f(x_0,y_0).$$

h étant fixé, possons $g : y \mapsto f(x_0+h, y) - f(x_0, y)$, qui est de classe $\mathcal{C}^1$ sur un intervalle ouvert contenant y_0. De sorte que $F(h) = g(y_0+h) - g(y_0)$. Soit $\varphi : \left\{ \begin{array}{lll} [0;1] & \to & F \\ t & \to & g(y_0+th) \end{array} \right.$ On a :

$$F(h) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt = h\int_0^1 g'(y_0+th)\,dt = h\int_0^1 \left(\frac{\partial f}{\partial y}(x_0+h, y_0+th) - \frac{\partial f}{\partial y}(x_0, y_0+th)\right)dt$$

t étant fixé, soit $\Psi_t : \left\{ \begin{array}{lll} [0;1] & \to & F \\ u & \mapsto & \dfrac{\partial f}{\partial y}(x_0+uh, y_0+th) \end{array} \right.$ qui est de classe $\mathcal{C}^1$.

$$F(h) = h\int_0^1 (\Psi_t(1) - \Psi_t(0))\,dt = h\int_0^1\int_0^1 \Psi_t'(u)dudt = h^2\int_0^1\int_0^1 \frac{\partial^2 f}{\partial x \partial y}(x+uh, y_0+th)\,dudt.$$

Par continuité de $\dfrac{\partial^2 f}{\partial x \partial y}$ sur U : pour ε dans $\mathbb{R}_+^*$, il existe $\alpha \in \mathbb{R}_+^*$ tel que :
$|h| \leq \alpha$ et $|k| \leq \alpha \Rightarrow \left|\dfrac{\partial^2 f}{\partial x \partial y}(x_0+h, y_0+k) - \dfrac{\partial^2 f}{\partial y \partial x}(x_0,y_0)\right| \leq \varepsilon$.
Si $|h| \leq \varepsilon$(et $h \neq 0$), alors on a :

$$\left|\frac{F(h)}{h^2} - \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\right| \leq \int_0^1\int_0^1 \left|\frac{\partial^2 f}{\partial x \partial y}(x_0+h, y_0+th) - \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\right| dudt \leq \varepsilon$$

Ainsi, $\displaystyle\lim_{h\to 0}\frac{F(h)}{h^2} = \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)$. De même en posant $l : x \mapsto f(x, y_0+h) - f(x,y_0)$, on obtient $\displaystyle\lim_{h\to 0}\frac{F(h)}{h^2} = \frac{\partial^2 f}{\partial y \partial x}(x_0,y_0)$ et on conclut par unicité de la limite.

Remarque 4.1.1 *Si f est de classe $\mathcal{C}^k$, avec $k \geq 2$, alors les dérivées partielles d'ordre k ne dépendent pas de l'ordre de dérivation.*

Exemple 4.1.3 *1. Soit* $f : \left\{ \begin{array}{lll} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \left\{ \begin{array}{ll} xy\dfrac{x^2-y^2}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{array} \right. \end{array} \right.$

Dans l'exemple 2.3.2, nous avons vu que : $\dfrac{\partial f}{\partial x} : \left\{ \begin{array}{lll} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \left\{ \begin{array}{ll} \dfrac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{array} \right. \end{array} \right.$

et : $\dfrac{\partial f}{\partial y} : \left\{ \begin{array}{lll} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \left\{ \begin{array}{ll} \dfrac{x^5-4x^3y^2-xy^4}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{array} \right. \end{array} \right.$.

Montrer que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$. Montrer que $\dfrac{\partial^2 f}{\partial x \partial y}(0,0)$ et $\dfrac{\partial^2 f}{\partial y \partial x}(0,0)$ existent. Que peut-on conclure ?

2. Soit f une fonction de classe $\mathcal{C}^2$ de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et si x est un élément de $\mathbb{R}^n$, on note $f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix}$, *en identifiant $\mathcal{M}_{n,1}(\mathbb{R})$ avec $\mathbb{R}^n$.*

Soient $i, j \in [\![1,n]\!]$, alors $D_jf_i(x)$ désignera $\dfrac{\partial f_i}{\partial x_j}(x)$.

Avec les notations précédentes, on a : $J_f(x) = [D_jf_i(x)]_{1\leq i,j\leq n}$.

Soient $i, j, k \in [\![1,n]\!]$, alors $D_{i,j}f_k(x)$ désignera $\dfrac{\partial^2 f_k}{\partial x_i \partial x_j}(x)$, ou encore $f_{i,j,k}(x)$.

On considère la proposition

$(\mathcal{P})$ *: Pour tout x de $\mathbb{R}^n$, la matrice jacobienne $J_f(x)$ de f est orthogonale.*

Pour x dans $\mathbb{R}^n$ et i, j, k dans $[\![1,n]\!]$, on note $\alpha_{i,j,k}(x) = \displaystyle\sum_{p=1}^{n} \frac{\partial f_p}{\partial x_i}(x).\frac{\partial^2 f_p}{\partial x_j \partial x_k}(x)$.

(a) On suppose $(\mathcal{P})$.

i. Montrer que pour tous i, j et k dans $[\![1,n]\!]$, on a : $\alpha_{i,j,k} = \alpha_{i,k,j} = -\alpha_{k,j,i}$.

ii. *En déduire que pour tous i, j et k dans $[\![1, n]\!]$, on a :* $\alpha_{i,j,k} = 0$.

iii. *Montrer qu'il existe une matrice orthogonale A et un élément b de $\mathbb{R}^n$ tels que :* $\forall x \in \mathbb{R}^n,\ f(x) = Ax + b$.

(b) *Donner une condition nécessaire et suffisante sur f pour que la proposition $(\mathcal{P})$ soit réalisée.*

3. *Équation de d'Alembert sur la propagation d'une onde :* *déterminer les fonctions de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ telles que :* $\dfrac{\partial^2 f}{\partial x^2} - \dfrac{1}{c^2}\dfrac{\partial^2 f}{\partial t^2} = 0$, *avec* $c > 0$. *On posera le changement de variable* $u = x - ct$ *et* $v = x + ct$.

Le changement de variable est bijectif de $\mathbb{R}^2$ dans $\mathbb{R}^2$, car $\left\{\begin{array}{lll} u & = & x - ct \\ v & = & x + ct \end{array}\right. \Leftrightarrow \left\{\begin{array}{lll} t & = & \dfrac{v-u}{2c} \\ x & = & \dfrac{u+v}{2} \end{array}\right.$.

Soit f une de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ et on pose $g : (u, v) \mapsto f\left(\underbrace{\dfrac{u+v}{2}}_{x}, \underbrace{\dfrac{v-u}{2c}}_{t}\right)$ *qui est définie sur $\mathbb{R}^2$ et de classe $\mathcal{C}^2$.*

On a : $f(x, t) = g(\underbrace{x - ct}_{u}, \underbrace{x + ct}_{v})$.

On a : $\dfrac{\partial f}{\partial x}(x, t) = \dfrac{\partial g}{\partial u}(x - ct, x + ct) + \dfrac{\partial g}{\partial v}(x - ct, x + ct)$, *puis :*
$\dfrac{\partial^2 f}{\partial x^2}(x, t) = \dfrac{\partial^2 g}{\partial u^2}(x - ct, x + ct) + \dfrac{\partial^2 g}{\partial v \partial u}(x - ct, x + ct) + \dfrac{\partial^2 g}{\partial u \partial v}(x - ct, x + ct) + \dfrac{\partial^2 g}{\partial v^2}(x - ct, x + ct) = \dfrac{\partial^2 g}{\partial u^2}(u, v) + 2\dfrac{\partial^2 g}{\partial v \partial v}(u, v) + \dfrac{\partial^2 g}{\partial v^2}(u, v)$ *grâce au théorème de Schwarz, car g est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$.*

On a : $\dfrac{\partial f}{\partial t}(x, t) = -c\dfrac{\partial g}{\partial u}(x - ct, x + ct) + c\dfrac{\partial g}{\partial v}(x - ct, x + ct)$, *puis :* $\dfrac{\partial^2 f}{\partial t^2}(x, t) =$
$c^2\dfrac{\partial^2 g}{\partial u^2}(x - ct, x + ct) - c^2\dfrac{\partial^2 g}{\partial v \partial u}(x - ct, x + ct) - c^2\dfrac{\partial^2 g}{\partial u \partial v}(x - ct, x + ct) + c^2\dfrac{\partial^2 g}{\partial v^2}(x - ct, x + ct) = c^2\dfrac{\partial^2 g}{\partial u^2}(u, v) - 2c^2\dfrac{\partial^2 g}{\partial v \partial u}(u, v) + c^2\dfrac{\partial^2 g}{\partial v^2}(u, v)$

(theorème de Schwarz).

Ainsi $0 = \dfrac{\partial^2 f}{\partial x^2} - \dfrac{1}{c^2}\dfrac{\partial^2 f}{\partial t^2} = 4\dfrac{\partial^2 g}{\partial v \partial u} \Leftrightarrow \forall (u,v) \in \mathbb{R}^2,\ \dfrac{\partial^2 g}{\partial v \partial u}(u,v) = 0.$

Grâce à l'exemple 4.1.2, cela est équivalent à l'existence de H, K *dans* $\mathcal{C}^2(\mathbb{R},\mathbb{R})$ *telles que :* $\forall (u,v) \in \mathbb{R}^2,\ g(u,v) = H(v) + K(u)$.

Ainsi $\mathcal{S} = \{(x,t) \mapsto H(x+ct) + K(x-ct),\ H, K \in \mathcal{C}^2(\mathbb{R},\mathbb{R})\}$.

5 Cas des applications à valeurs dans $\mathbb{R}$

5.1 Gradient

5.1.1 Gradient sur $\mathbb{R}^p$

Soient U un ouvert de $\mathbb{R}^p$ et $f : U \to \mathbb{R}$ une application différentiable en a.

Définition 5.1.1 (Gradient sur $\mathbb{R}^p$) *Le gradient de f en a est le vecteur*

$$\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_p f(a) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_p}(a) \end{pmatrix} \quad ou \quad \nabla f(a) = (\partial_1 f(a), ..., \partial_p f(a)) \,.$$

Proposition 5.1.1 (Différentielle et Gradient) *On munit $\mathbb{R}^p$ de sa structure euclidienne usuelle. Soit $f \in \mathcal{C}^1(U,\mathbb{R})$ et $a \in U$. Alors, pour tout $h \in \mathbb{R}^p$,*

$$df(a) \cdot h = (\nabla f(a)|h).$$

Démonstration :

Remarque 5.1.1 *Grâce à la proposition précédente, on remarque que pour un petit déplacement h par rapport à a, avec $\|h\|$ fixé, la variation entre $f(a+h)$ et $f(a)$ (c'est-à-dire $f(a+h) - f(a)$ qui vaut à l'ordre un $(\nabla f(a)|h)$) est :*

- *maximale lorsque*

- *minimale lorsque*

Ainsi $\nabla f(a)$ donne la direction selon laquelle f varie le plus vite. Nous verrons plus d'interprétations du gradient dans le chapitre sur les courbes.

Exemple 5.1.1 *Soit* $f : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \to & \mathbb{R} \\ M & \mapsto & \det(M) \end{cases}$. *Déterminer $\nabla(f)$ et exprimer la différentielle df.*

5.1.2 Cas général

Dans ce paragraphe E est un espace euclidien et $f : U \to \mathbb{R}$ est une application différentiable en a. L'application $df(a) : E \to \mathbb{R}$ est donc une forme linéaire. Le théorème de représentation de Riesz nous donne :

Définition 5.1.2 (Gradient) *Il existe un unique vecteur $v \in E$ tel que :*

$$\forall h \in E,\ df(a)(h) = (h|v) = (v|h).$$

Le vecteur v est le gradient de f en a noté $\nabla f(a)$. Ainsi : $\forall h \in E,\ df(a)(h) = (h|\nabla f(a))$.

Faisons le lien comme dans le paragraphe précédent avec les dérivées partielles :

Proposition 5.1.2 (Composantes du gradient en base orthonormée) *Soit $\mathcal{B} = (e_1, ..., e_p)$ une base orthonormée de E. Alors :*

$$\nabla f(a) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a) e_i = \sum_{i=1}^{p} \partial_i f(a) e_i.$$

Démonstration : Soit $i \in [\![1, n]\!]$. Comme $\mathcal{B}$ est une base orthonormée, $\nabla f(a) = \sum_{i=1}^{n} (\nabla f(a)|e_i) e_i$ ($\nabla f(a)$ est un vecteur de E que l'on vient de décomposer dans $\mathcal{B}$). Soit $i \in [\![1, n]\!]$. On a par définition du gradient : $(\nabla f(a)|e_i) = df(a)(e_i) = D_{e_i}(a) = \partial_i f(a)$.

Exemple 5.1.2 *Soit E un espace euclidien et f un endomorphisme autoadjoint de E et $u \in E$. Soit g définie sur E par $g : x \mapsto \frac{1}{2}(f(x)|x) - (u|x)$.*

1. *Déterminer $\nabla g(a)$.*
2. *Soit $\mathcal{B} = (e_1, ..., e_n)$ une base orthonormée de E et $A = [a_{i,j}]_{1 \leq i,j \leq n} = Mat_{\mathcal{B}}(f)$. Déterminer $\frac{\partial g}{\partial x_k}(x)$ pour $x \in E$ et $k \in [\![1, p]\!]$.*

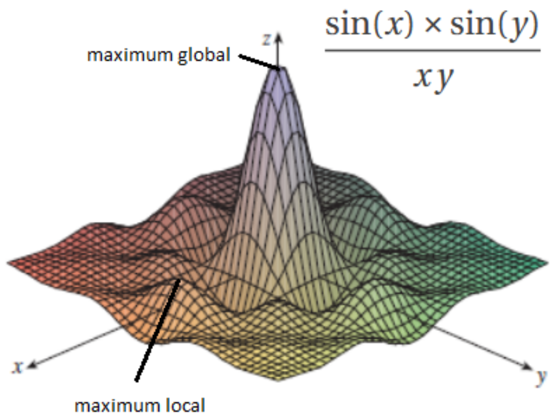
5.2 Extremums

5.2.1 Extremums et points critiques

Définition 5.2.1 (Extremum local / global) *Soient A une partie de E, $f : A \to \mathbb{R}$ une application et $a \in A$.*

1. *f présente un maximum local en a s'il existe une boule ouverte $\mathcal{B}(a, r)$ telle que pour tout $x \in \mathcal{B}(a, r) \cap A$, $f(x) \leq f(a)$. Ce maximum vaut $f(a)$.*
2. *f présente un minimum local en a s'il existe une boule ouverte $\mathcal{B}(a, r)$ telle que pour tout $x \in \mathcal{B}(a, r) \cap A$, $f(x) \geq f(a)$ Ce minimum vaut $f(a)$.*
3. *f présente un maximum global en a si : $\forall x \in A,\ f(x) \leq f(a)$.*

4. *f présente un minimum global en a si : $\forall x \in A,\ f(x) \geq f(a)$.*
5. *Un extremum est un maximum ou un minimum.*



Exemple 5.2.1 *Soient E un espace euclidien, $\|.\|$ la norme de cet espace, φ une forme linéaire non nulle sur E et f définie sur E par : $\forall x \in E,\ f(x) = \varphi(x)e^{-\|x\|^2}$. Étudier les extrema de f. On pourra se placer dans une base simplifiant les expressions.*

Définition 5.2.2 (Point critique) *Soient U un ouvert de E, $f : U \to \mathbb{R}$ une application différentiable en $a \in U$. On dit que a est un point critique de f si $df(a) = 0$.*
Autrement dit si on se fixe une base $\mathcal{B} = (e_1, ..., e_p)$ de E, alors : $\forall i \in [\![1, p]\!],\ \dfrac{\partial f}{\partial x_i}(a) = 0$, soit $\nabla f(a) = 0$ si E est un espace euclidien avec $\mathcal{B}$ une base orthonormale .

Exemple 5.2.2 *Soit E un espace euclidien, $f \in \mathcal{S}^{++}(E)$ et $u \in E$. Soit g définie sur E par $g : x \mapsto \dfrac{1}{2}(f(x)|x) - (u|x)$. Montrer que g a un unique point critique.*

Proposition 5.2.1 (Condition nécessaire d'existence d'un extremum) *Soient U un ouvert de E, $f : U \to \mathbb{R}$ une application différentiable en $a \in U$. Si f présente un extremum local en a, alors a est un point critique de f.*

Démonstration :

Remarque 5.2.1 1. *ATTENTION : on doit être sur un ouvert ! Par exemple si $f : \left\{ \begin{array}{ccc} [0,1] & \to & [0,1] \\ x & \mapsto & x \end{array} \right.$, f admet des extrema en 0 et 1, mais $f' = 1$ et ne s'annule jamais.*

2. *ATTENTION : être un point critique n'implique pas être un extremum local ou global. Contrexemple : soit $f : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array} \right.$ On a $f'(0) = 0$, mais f n'admet pas d'extremum en 0.*
3. *IMPORTANT : pour chercher des extrema, on se place sur un ouvert et on cherche les points critiques.*
   Ensuite il faut vérifier que si a est un point critique, alors f admet un extremum local ou global en a. Pour vérifier cela, on étudie le signe de $f(a+h) - f(a)$ avec h voisin de 0. Pour avoir un minimum local, cette quantité doit être positive au voisinage de a et pour un maximum, elle doit être négative.
   Si f est définie sur une partie de $\mathbb{R}^p$, alors $a = (a_1, ..., a_p)$ et on étudie le signe de $f(a_1 + h_1, ..., a_p + h_p) - f(a_1, ..., a_p)$ avec $h = (h_1, ..., h_p)$ voisin de 0.
   Dans le paragraphe suivant nous donneront des méthodes efficaces pour savoir si on a un minimum ou un maximum local.

Exemple 5.2.3 1. *Soit E un espace euclidien, $f \in \mathcal{S}^{++}(E)$ et $u \in E$. Soit g définie sur E par $g : x \mapsto \frac{1}{2}(f(x)|x) - (u|x)$. Montrer que g admet un minimum global.*

2. *Soit f une fonction convexe de classe $\mathcal{C}^1$ de $\mathbb{R}^n$ dans $\mathbb{R}$. Montrer que $a \in \mathbb{R}^n$ est un point critique de f si et seulement si c'est un minimum global.*

   Le sens retour découle de la proposition précédente car $\mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$.

   Pour le sens direct, grâce à la formule de l'exemple 3.3.1, si a est un point critique, alors : $\forall x \in \mathbb{R}^n,\ 0 = df(a)(x-a) \leq f(x) - f(a)$.

5.2.2 Extremum et fonctions de classe $\mathcal{C}^2$

Dans ce paragraphe, U désignera un ouvert de $\mathbb{R}^p$ et $f : U \to \mathbb{R}$ une application de classe $\mathcal{C}^2$.
On rappelle un point que l'on a vu en exemple dans le chapitre 13. Si A est dans $S_n(\mathbb{R})$, avec $\lambda_1 \leq ... \leq \lambda_n$ ses valeurs propres comptées avec leur multiplicité, alors :
$\forall X \in \mathcal{M}_{n,1}(\mathbb{R}),\ \lambda_1 \|X\|^2 \leq X^T A X \leq \lambda_n \|X\|^2$.

Définition 5.2.3 (Hessienne) *On appelle matrice hessienne de f en x dans U, la matrice*

$$H_f(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{1\leq i,j\leq n}.$$

Remarque 5.2.2

1. *(IMPORTANTE) On a aussi* $H_f(x) = \left[\frac{\partial^2 f}{\partial x_j \partial x_i}\right]_{1\leq i,j\leq n}$ *qui est donc symétrique, grâce*
2. *$H_f(x)$ est donc diagonalisable en tant que matrice symétrique réelle.*
3. *On a :* $H_f(x) = J_{\nabla f(x)}$.

Exemple 5.2.4 *Soit* $\mathcal{B} = (e_1, ..., e_n)$ *la base canonique de* $\mathbb{R}^n$ *muni de son produit scalaire usuel,* $f \in \mathcal{S}^{++}(\mathbb{R}^n)$ *et* $u \in \mathbb{R}^n$. *Soit g définie sur $\mathbb{R}^n$ par* $g : x \mapsto \frac{1}{2}(f(x)|x) - (u|x)$.
On pose $A = [a_{i,j}]_{1\leq i,j\leq n} = Mat_{\mathcal{B}}(f)$. *Soit* $x \in \mathbb{R}^n$. *Déterminer* $H_g(x)$.

Proposition 5.2.2 (Formule de Taylor-Young à l'ordre 2) *On a :*

$$f(x+h) \underset{h\to 0}{=} f(x)+df(x).h+\frac{1}{2}\langle H_f(x)h, h\rangle+o(\|h\|^2) \underset{h\to 0}{=} f(x)+\langle \nabla f(x), h\rangle+\frac{1}{2}\langle H_f(x)h, h\rangle+o(\|h\|^2) \underset{h\to 0}{=}$$

$$f(x) + \nabla f(x)^T h + \frac{1}{2}h^T H_f(x)h + o(\|h\|^2),$$

soit

$$f(x+h) = f(x)+df(x).h+\frac{1}{2}\langle H_f(x)h, h\rangle+\|h\|^2\varepsilon(h) = f(x)+\langle \nabla f(x), h\rangle+\frac{1}{2}\langle H_f(x)h, h\rangle+\|h\|^2\varepsilon(h) =$$

$$f(x) + \nabla f(x)^T h + \frac{1}{2}h^T H_f(x)h + \|h\|^2\varepsilon(h),$$

avec $\lim_{h\to 0} \varepsilon(h) = 0$.

Démonstration : (Hors-programme)

Soit $g : (h,t) \mapsto (1-t).|||H_f(a+th) - H_f(a)|||$ définie sur $\overline{B}(0,\alpha) \times [0,1]$.
Soit $t \in [0,1]$. La fonction $h \mapsto g(h,t)$ est continue sur $\overline{B}(0,\alpha)$.
Soit $h \in \overline{B}(0,\alpha)$. La fonction $t \mapsto g(h,t)$ est continue par morceaux sur $[0,1]$.
L'ensemble $\overline{B}(0,\alpha) \times [0,1]$ est compact (produit de deux espaces compacts qui sont respectivement des fermés bornés d'un espace vectoriel de dimension finie). De plus, la fonction g est continue par opération sur ce compact, donc elle y est bornée. Il existe donc $M \in \mathbb{R}_+$ tel que :
$\forall (h,t) \in \overline{B}(0,\alpha) \times [0,1],\ |g(h,t)| \leq M$. Par ailleurs la fonction $t \mapsto M$ est continue et intégrable sur le segment $[0,1]$.
Ainsi l'application $h \mapsto \displaystyle\int_0^1 g(h,t)dt = \int_0^1 (1-t).|||H_f(a+th) - H_f(a)|||dt$ est continue sur $\overline{B}(0,\alpha)$ et donc $\displaystyle\lim_{h\to 0} \int_0^1 g(h,t)dt = \int_0^1 g(0,t)dt = \int_0^1 0 = 0$. Ceci clôt la preuve.

Remarque 5.2.3 *1. En utilisant les coordonnées, on a pour $a \in U$ et $h = (h_1, ..., h_n)$ qui tend vers $(0,...,0)$:*

$$f(a+h) =$$

2. En utilisant cette preuve, si on a $U = \mathbb{R}^n$ et : $\forall x \in \mathbb{R}^n,\ Sp(H_f(x)) \subset [a, +\infty[$, alors :
$\forall x \in \mathbb{R}^n,\ f(x) \geq f(0) + (\nabla(f)(0)|x) + \dfrac{a\|x\|^2}{2}$.

On reprend la preuve précédente et on a :

$\forall x \in \mathbb{R}^n,\ f(x) = f(0) + (\nabla(f)(0)|x) + \int_0^1 (1-t)x^T H_f(tx)xdt$.

Comme $H_f(tx)$ est une matrice symétrique, nous avons vu dans le chapitre 13 que si λ est sa plus petite valeur propre, alors : $x^T H_f(tx)x \geq \lambda x^T x \geq a\|x\|^2$

puis : $\int_0^1 (1-t)x^T H_f(tx)xdt \geq \int_0^1 (1-t)a\|x\|^2 dt = a\|x\|^2/2$.

3. Soit $f : E \to \mathbb{R}$ de classe $\mathcal{C}^2$.
Grâce à cette démonstration, on peut prouver f est convexe si et seulement si :
$\forall a \in E,\ H_f(a) \in S_n^+(\mathbb{R})$.

Soient $a, h \in E$. Dans la démonstration de la proposition 5.2.2, on a vu que si on pose $\forall t \in \mathbb{R},\ \varphi(t) = f(a+th)$, alors φ est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ et :

$\forall t \in \mathbb{R},\ \varphi''(t) = h^T H_f(a+th)h$.

Supposons f convexe. On a alors :

$\forall x, y \in \mathbb{R}, \forall t \in [0,1],\ \varphi(tx+(1-t)y) = f(a+(tx+(1-t)y)h) = f(t(a+xh)+(1-t)(a+yh)) \leq tf(a+xh)+(1-t)f(a+yh) = t\varphi(x)+(1-t)\varphi(y)$.

Ainsi φ est convexe et en particulier $\varphi''(0) \geq 0$, puis : $\forall h \in E,\ h^T H_f(a)h \geq 0$. Comme $H_f(a)$ est symétrique, alors : $H_f(a) \in S_n^+(\mathbb{R})$.

On suppose que : $\forall a \in E,\ H_f(a) \in S_n^+(\mathbb{R})$.

Soient $x, y \in E$. On pose $h = x - y$ et grâce au résultat rappelé, $\varphi : t \mapsto f(y+th)$ est convexe, car : $\varphi'' \geq 0$.

En particulier, on a : $\forall t \in [0,1],\ \varphi(t) = \varphi(t.1+(1-t).0) \leq t\varphi(1)+(1-t)\varphi(0)$, soit : $f(tx+(1-t)t) \leq tf(x)+(1-t)f(y)$, donc f est convexe.

Proposition 5.2.3 (Condition nécessaire d'extremum) *Si f admet un minimum (resp. maximum) local en $x_0 \in U$, alors x_0 est un point critique de f et $H_f(x_0)$ est dans $S_n^+(\mathbb{R})$, c'est-à-dire que les valeurs propres de $H_f(x_0)$ sont positives (resp. $-H_f(x_0)$ est dans $S_n^+(\mathbb{R})$, c'est-à-dire que les valeurs propres de $H_f(x_0)$ sont négatives).*

Démonstration : On suppose que $f(x_0)$ est un minimum local.

Proposition 5.2.4 (Condition suffisante d'extremum) *Si $x_0 \in U$ est un point critique de f et si $H_f(x_0)$ (resp. $-H_f(x_0)$) est dans $S_n^{++}(\mathbb{R})$, alors f atteint un minimum (resp. maximum) local en x_0 et localement, c'est le seul point en lequel ce minimum (resp. maximum) est atteint.*

Démonstration : La formule de Taylor-Young nous donne :
$f(x_0+h) - f(x_0) = \frac{1}{2}h^T H_f(x_0)h + \|h\|^2\varepsilon(h)$, avec $\varepsilon : U \to \mathbb{R}$ telle que $\lim\limits_{0} \varepsilon = 0$. Grâce au rappel donné au début de ce paragraphe, si on note $m = \min(Sp(H_f(x_0))$, on a :
$f(x_0+h)-f(x_0) \geq \frac{m}{2}\|h\|^2 + \|h\|^2\varepsilon(h) = \|h\|^2\left(\frac{m}{2}+\varepsilon(h)\right)$. Par définition de la limite, il existe $r \in \mathbb{R}_+^*$ tel que $B(x_0,r) \subset U$ et : $\forall h \in B(0,r),\ |\varepsilon(h)| \leq m/4$. Ainsi :
$\forall h \in B(0,r),\ f(x_0+h) - f(x_0) \geq \frac{m}{4}\|h\|^2 \geq 0$, ce qui fait de $f(x_0)$ un minimum local et
$f(x_0+h) - f(x_0) = 0$ si et seulement si $h = 0$, pour h dans $B(0,r)$, ce qui fait que $f(x_0)$ n'est atteint qu'en x_0 sur $B(x_0,r)$.

Exemple 5.2.5 *Soit $\mathcal{B} = (e_1, ..., e_n)$ la base canonique de $\mathbb{R}^n$ muni de son produit scalaire usuel, $f \in \mathcal{S}^{++}(\mathbb{R}^n)$ et $u \in \mathbb{R}^n$. Soit g définie sur $\mathbb{R}^n$ par $g : x \mapsto \frac{1}{2}(f(x)|x) - (u|x)$.*
On pose $A = [a_{i,j}]_{1\leq i,j\leq n} = Mat_{\mathcal{B}}(f)$. On a vu dans l'exemple 5.2.4 que : $\forall x \in \mathbb{R}^n,\ H_g(x) = A$ et A est définie positive (car $\mathcal{B}$ est une base orthonormée). Ainsi on retrouve le fait qu'en $a = f^{-1}(u)$, on ait un minimum local (vu dans l'exemple 5.2.3).

Corollaire 5.2.1 (Condition suffisante d'extremum pour $n = 2$) *Soit x_0 un point critique de*
$f : \begin{cases} U & \to & \mathbb{R} \\ (x,y) & \mapsto & f(x,y) \end{cases}$, *avec U un ouvert de $\mathbb{R}^2$.*
On note $r = \frac{\partial^2 f}{\partial x^2}(x_0)$, $t = \frac{\partial^2 f}{\partial y^2}(x_0)$ et $s = \frac{\partial^2 f}{\partial x \partial y}(x_0)$. Alors

- *Si $\det(H_f(x_0)) = rt - s^2 > 0$ et $tr(H_f(x_0)) = r + t > 0$, alors f admet un minimum local en x_0 ($H_h(x_0)$ a deux valeurs propres strictement positives).*
- *Si $\det(H_f(x_0)) = rt - s^2 > 0$ et $tr(H_f(x_0)) = r + t < 0$, alors f admet un maximum local en x_0 ($H_h(x_0)$ a deux valeurs propres strictement négatives).*
- *Si $\det(H_f(x_0)) = rt - s^2 < 0$, alors f ne présente pas d'extremum local en x_0 ($H_h(x_0)$ a deux valeurs propres de signe contraire).*

Démonstration : Soient λ_1 et λ_2 les deux valeurs propres réelles de $H_f(x_0)$. On a $\det(H_f(x_0)) = \lambda_1\lambda_2$. Dans les deux premiers cas, les valeurs propres sont de même signe. De plus dans ces cas, comme $tr(H_f(x_0)) = \lambda_1+\lambda_2$, alors la trace donne le signe de ces valeurs propres et donc $H_f(x_0)$ est dans $S_n^{++}(\mathbb{R})$ si elles sont strictement positives et $-H_f(x_0)$ est dans $S_n^{++}(\mathbb{R})$ si elles sont strictement négatives. On est donc dans le cadre de la proposition précédente.
Si $\det(H_f(x_0)) < 0$, on a donc deux valeurs propres de signe opposé, par exemple $\lambda_1 > 0$ et $\lambda_2 < 0$. Soit u_i un vecteur propre unitaire associé à la valeur propre λ_i pour i dans $[\![1,2]\!]$. Pour α petit , comme $\|u_i\| = 1$, on a :

$$f(x_0+\alpha u_i) - f(x_0) = \alpha^2\frac{1}{2}u_i^T H_f(x_0)u_i + \alpha^2\varepsilon(\alpha u_i) = \alpha^2\lambda_i\frac{1}{2}\underbrace{u_i^T u_i}_{\|u_i\|=1} + \alpha^2\varepsilon(\alpha u_i) = \alpha^2\left(\frac{\lambda_i}{2}+\varepsilon(\alpha u_i)\right).$$

Comme $\lim\limits_{\alpha\to 0}\varepsilon(\alpha u) = 0$, alors $f(x_0+\alpha u_i) - f(x_0) \underset{\alpha\to 0^+}{\sim} \frac{\alpha^2\lambda_i}{2}$ et donc $f(x_0+\alpha u_1) - f(x_0)$ est positif pour α petit, donc $f(x_0)$ n'est pas un maximum local et $f(x_0+\alpha u_2) - f(x_0)$ est négatif pour α petit, donc $f(x_0)$ n'est pas un minimum local.

Remarque 5.2.4 *Si $rt - s^2 = 0$, alors on ne peut rien conclure. On verra dans le paragraphe suivant comment faire dans ce cas.*

Exemple 5.2.6 *Soit f la fonction définie sur $\mathbb{R}^2$ par :*
$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = 2x^3 + 6xy - 3y^2 + 2.$

1. *f admet-elle des extrema locaux sur $\mathbb{R}^2$? Si oui, les déterminer.*
2. *f admet-elle des extrema globaux sur $\mathbb{R}^2$?*

5.2.3 Cas particuliers d'extremums

a) Cas où la Hessienne ne permet pas de conclure

Exemple 5.2.7 *Chercher les extrema locaux sur $\mathbb{R}^2$ de $f : (x,y) \mapsto x^4 + y^4 + y^3$.*

b) Extremum sur un compact

Lorsque nous faisons l'étude d'extrema sur un compact, la situation est un peu particulière car pour chercher les points critiques, nous devons imperativement nous placer sur un ouvert. Voici un exemple qui montre comment s'en sortir :

Exemple 5.2.8 *Montrer que* $f : (x, y) \mapsto 2x^3 + 6xy - 3y^2 + 2$ *admet un maximum global sur* $K = [0, 1] \times [0, 1]$ *et le déterminer.*

6 Vecteurs tangents à une partie et optimisation sous contrainte

6.1 Surfaces de $\mathbb{R}^3$

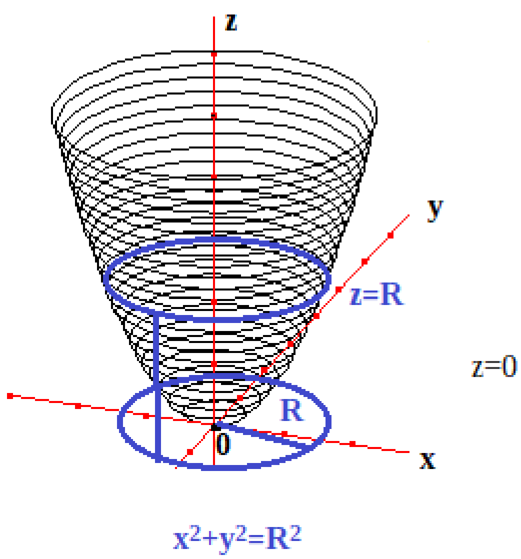
Définition 6.1.1 (Surfaces de $\mathbb{R}^3$) *Si F est une application de classe $\mathcal{C}^1$ d'un ouvert U de $\mathbb{R}^3$ dans $\mathbb{R}$, alors l'ensemble $\{(x,y,z) \in U,\ F(x,y,z) = 0\}$ est appelé surface de $\mathbb{R}^3$.*

Exemple 6.1.1
1. *$0 = ax + by + cz - d$, avec $(a,b,c) \neq (0,0,0)$ représente l'équation d'un plan de l'espace.*
2. *$0 = x^2 + y^2 + z^2 - 2z - 8$ représente la sphère de rayon $A(0,0,1)$ et de rayon 3, car on a : $x^2 + y^2 + (z-1)^2 = 3^2$, soit $\|(x,y,z) - (0,0,1)\| = 3$.*

Voici un cas particulier de surface :

Définition 6.1.2 (Graphe d'une fonction numérique de $\mathbb{R}^2$) *Une fonction $f : \begin{cases} \mathbb{R}^2 & \to \mathbb{R} \\ (x,y) & \mapsto f(x,y) \end{cases}$ de classe $\mathcal{C}^1$ se représente dans l'espace à l'aide de points de coordonnées (x,y,z) tels que $z = f(x,y)$. Cela définit une surface qui est incluse dans $\mathbb{R}^3$, en considérant $\{(x,y,z) \in \mathbb{R}^3,\ f(x,y) - z = 0\}$.*

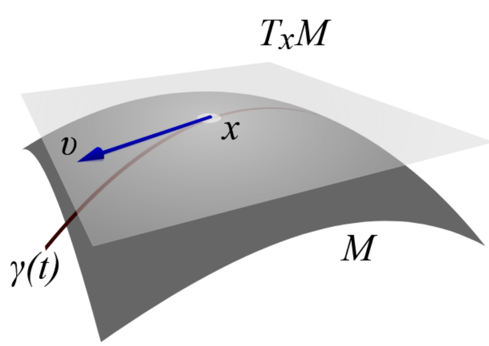
Exemple 6.1.2 *Représenter $f : (x,y) \mapsto x^2 + y^2$ sur $\mathbb{R}^2$.*



6.2 Vecteurs tangents

Dans ce paragraphe, E désigne un $\mathbb{R}$-espace vectoriel de dimension finie et M est une partie de E.

Définition 6.2.1 (Vecteur tangent à M) *Soient $x \in M$ et $v \in E$. On dit que v est un vecteur tangent à M en x s'il existe un réel strictement positif ε et une application $\gamma :]-\varepsilon, \varepsilon[\to M$ dérivable en 0, telle que : $\gamma(0) = x$ et $\gamma'(0) = v$.*
On note $T_x M$ l'ensemble des vecteurs tangents à M en x.



Exemple 6.2.1 *Déterminer $T_{I_n}(O_n(\mathbb{R}))$.*

Proposition 6.2.1 (Vecteurs tangents d'une partie définie implicitement) *Soit $g : \Omega \to \mathbb{R}$, une application de classe $\mathcal{C}^1$, avec Ω un ouvert de E. Soit $M = \{x \in \Omega,\ g(x) = 0\}$. Soit $x \in M$. Si $dg(x) \neq 0$, alors $T_xM = \mathrm{Ker}\,(dg(x))$.*
En particulier si E est muni d'une structure euclidienne, si $\nabla g(x) \neq 0$, alors $T_xM = (\nabla g(x))^{\perp}$.

Démonstration : Nous admettons le premier point.
Le deuxième point vient du fait que : $\forall h \in E,\ dg(x).h = \langle \nabla g(x), h\rangle$.

Remarque 6.2.1 *On retrouve le fait que le gradient soit orthogonal aux lignes de niveau. En effet soit $f : \Omega \to \mathbb{R}$ une application de classe $\mathcal{C}^1$, avec Ω un ouvert de E. Soit $\lambda \in \mathbb{R}$. On considère la ligne de niveau $X_\lambda = \{x \in \Omega,\ f(x) = \lambda\}$. Soient $x \in X_\lambda$ et $v \in T_xX_\lambda$. Alors v et $\nabla f(x)$ sont orthogonaux. En reprenant la proposition précédente, on considère $g = f - \lambda$ et donc $X_\lambda = \{x \in \Omega,\ g(x) = 0\}$. Ainsi $T_xX_\lambda = (\nabla g(x))^{\perp} = (\nabla f(x))^{\perp}$, car $\nabla g = \nabla f$, car λ est constante.*

Exemple 6.2.2

1. *Nous rappelons que nous avons vu dans le chapitre 15, que si $f = \det : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$, alors : $\forall H \in \mathcal{M}_n(\mathbb{R}),\ df(I_n)(H) = tr(H)$. Soit $SL_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}),\ \det(M) = 1\}$. Déterminer $T_{I_n}(SL_n(\mathbb{R}))$.*

2. *On considère la sphère S d'équation $x^2 + y^2 + z^2 = R^2$, avec $R > 0$. Soit $u = (u_1, u_2, u_3) \in S$. Déterminer T_uS.*

Corollaire 6.2.1 (Vecteurs tangents à un graphe d'une fonction de deux variables) *Soient U un ouvert de $\mathbb{R}^2$ et $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Soit $\mathcal{S}$ la surface (de $\mathbb{R}^3$) définie par l'équation : $z = f(x, y)$ qui est le graphe de f.*
Soit $a = (x_0, y_0, z_0) \in \mathcal{S}$. Alors

$$T_a\mathcal{S} = \left\{(u, v, w) \in \mathbb{R}^3,\ w = u\frac{\partial f}{\partial x}(x_0, y_0) + v\frac{\partial f}{\partial y}(x_0, y_0)\right\}.$$

Démonstration :

Définition 6.2.2 (Plan tangent à une surface de $\mathbb{R}^3$) *Soient U un ouvert de $\mathbb{R}^3$ et $g : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. On pose $\mathcal{S}$ la surface définie par l'équation : $g(x, y, z) = 0$ et on considère $a \in \mathcal{S}$. Le plan tangent à $\mathcal{S}$ en a est le sous-espace affine $a + T_a\mathcal{S}$ de $\mathbb{R}^3$. Ce plan passe donc par a et $T_a\mathcal{S}$ est sa direction.*

Corollaire 6.2.2 (Équation des plans tangents au graphe d'une fonction de deux variables) *Soient U un ouvert de $\mathbb{R}^2$ et $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. On pose $\mathcal{S}$ la surface définie par l'équation : $z = f(x, y)$.*
Soit $a = (x_0, y_0, z_0) \in \mathcal{S}$. Alors l'équation du plan tangent à $\mathcal{S}$ en a est :

Démonstration : On a : $m(x, y, z) \in \mathcal{S} \Leftrightarrow \overrightarrow{am}(x - x_0, y - y_0, z - z_0) \in T_a\mathcal{S}$, car a est dans $\mathcal{S}$.

Exemple 6.2.3
1. *Le plan tangent de la spère S d'équation $x^2 + y^2 + z^2 = R^2$ (avec $R > 0$) en $u = (u_1, u_2, u_3) \in S$ est*

2. *Soit $\mathcal{S}$ la surface d'équation $z = x^2 + y^2$. Soit $a = (1, -1, 2) \in \mathcal{S}$. Quelle est l'équation du plan tangent à $\mathcal{S}$ en a ?*

6.3 Optimisation sous contrainte d'égalité

Proposition 6.3.1 (Extremum sur une partie M) *Soit Ω un ouvert de E. Soit M une partie de Ω et $f : \Omega \to \mathbb{R}$ une application différentiable en un point x de M. Si $f|_M$ admet un extremum local en x, alors : $\forall v \in T_xM,\ df(x)(v) = 0$.*

Démonstration :

Théorème 6.3.1 (Optimisation sous contrainte) *Soient Ω un ouvert et f et g deux fonctions de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$. On pose $M = \{x \in \Omega,\ g(x) = 0\}$. Soit $x \in M$ tel que : $dg(x) \neq 0$.*
Si $f|_M$ admet un extremum local en x, alors il existe $\lambda \in \mathbb{R}$ tel que $df(x) = \lambda dg(x)$.
En particulier, si E est un espace euclidien, dans ces conditions il existe $\lambda \in \mathbb{R}$ tel que $\nabla f(x) = \lambda \nabla g(x)$.

Démonstration : $df(x)$ et $dg(x)$ sont deux formes linéaires. Or nous avons $T_xM = \mathrm{Ker}\,(dg(x))$, donc grâce à la proposition précédente, $\mathrm{Ker}\,(dg(x)) \subset \mathrm{Ker}\,(df(x))$.
Si $\mathrm{Ker}\,(df(x)) = E$, alors $df(x) = 0$, puis $df(x) = 0.dg(x)$.
Si $\mathrm{Ker}\,(df(x)) \neq E$, alors $\dim(\mathrm{Ker}\,(dg(x))) = \dim(\mathrm{Ker}\,(df(x))) = \dim(E) - 1$, puis $\mathrm{Ker}\,(dg(x)) = \mathrm{Ker}\,(df(x))$. Ainsi il existe $\lambda \in \mathbb{R}^*$ tel que : $df(x) = \lambda dg(x)$, car nous avons un hyperplan défini par les formes linéaires non nulles $df(x)$ et $dg(x)$.
Dans le cas euclidien, cela signifie que : $\forall h \in E,\ \langle \nabla f(x), h\rangle = df(x).h = \lambda dg(x).h = \lambda \langle \nabla g(x), h\rangle$, puis : $\forall h \in E,\ \langle \nabla f(x) - \lambda \nabla g(x), h\rangle$. Ainsi $\nabla f(x) - \lambda \nabla g(x)$ est dans $E^\perp = \{0\}$ et donc : $\nabla f(x) - \lambda \nabla g(x) = 0$.

Exemple 6.3.1
1. *Soit f l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f : (x, y) \mapsto 4x^2 + 12xy - y^2$.*
Soit $C = \left\{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 13\right\}$

(a) *Montrer que sur C, f admet un minimum et un maximum.*
Soit $(u,v) \in C$ un point où f atteint un de ses extremums.

(b) *Justifier qu'il existe un réel λ tel que le système (S) suivant soit vérifié :*
$(S) : \begin{cases} 4u + 6v = \lambda u \\ 6u - v = \lambda v \end{cases}$

(c) *Montrer que $(\lambda - 4)(\lambda + 1) - 36 = 0$. En déduire les valeurs possibles de λ.*

(d) *Déterminer les valeurs possibles de (u,v), puis donner le maximum et le minimum de f sur C.*

(a) *C est la sphère de $\mathbb{R}^2$ de centre $(0,0)$ et de rayon $\sqrt{13}$ pour la norme $\|\cdot\|_2$ usuelle.*
C est donc une partie fermée et bornée de $\mathbb{R}^2$. Comme $\mathbb{R}^2$ est de dimension finie, C est un compact de $\mathbb{R}^2$. On a aussi $C \neq \emptyset$. f est une application polynomiale donc f est continue sur le compact C. Donc f atteint un maximum et un minimum sur C.

(b) *Soit $g : (x,y) \mapsto x^2 + y^2 - 13$ qui est de classe C^1 sur l'ouvert $\mathbb{R}^2$ en tant que fonction polynomiale, tout comme f. De plus $\nabla g : (x,y) \mapsto (2x, 2y)$, de sorte que $\nabla g(x,y) = (0,0)$ si et seulement si $(x,y) = (0,0)$. On en déduit que ∇g ne s'annule pas sur C.*
Il découle alors, du théorème d'optimisation sous une contrainte, qu'il existe un scalaire λ tel que $\begin{pmatrix} 8u + 12v \\ 12u - 2v \end{pmatrix} = \nabla f(u,v) = \lambda \nabla g(u,v) = \lambda \begin{pmatrix} 2u \\ 2v \end{pmatrix}$,
soit $\begin{cases} 4u & +6v & = \lambda u \\ 6u & -v & = \lambda v \end{cases}$

(c) $(S) \Leftrightarrow \begin{cases} (4-\lambda)u & +6v & = 0 \\ 6u & -(1+\lambda)v & = 0 \end{cases}$
Comme $(0,0) \notin C$ et que f possède effectivement des extremums sur C, ce système linéaire a nécessairement au moins une solution non nulle. La matrice associée à ce système n'est pas inversible. Son déterminant $(\lambda - 4)(\lambda + 1) - 36$ est donc nul, donc : $\lambda^2 - 3\lambda - 40 = 0$.
Les solutions de cette équation sont 8 et -5.

(d) • *Si $\lambda = 8$ on obtient, d'après (S), $v = \frac{2}{3}u$. Comme $(u,v) \in C$, il vient*
$u^2 + v^2 = u^2 + \frac{4}{9}u^2 = \frac{13}{9}u^2 = 13$, *donc $u = \pm 3$, donc $(u,v) \in \{(3,2), (-3,-2)\}$.*
• *Si $\lambda = -5$ on obtient, d'après (S), $v = -\frac{3}{2}u$. Comme $(u,v) \in C$, il vient $\frac{13}{4}u^2 = 13$, donc $u = \pm 2$, donc $(u,v) \in \{(2,-3), (-2,3)\}$.*
Or $f(3,2) = f(-3,-2) = 104$ et $f(2,-3) = f(-2,3) = -65$.
Comme f atteint effectivement un maximum et un minimum sur C et qu'elle ne peut les atteindre qu'en ces points, on a donc : $\max_C f = 104$ et $\min_C f = -65$.

2. *Soit E un espace euclidien et $u \in S(E)$. On pose $f : x \mapsto \langle u(x), x \rangle$. On a vu dans le chapitre précédent que f est différentiable et : $\forall x \in E,\ \nabla f(x) = 2u(x)$. Montrer que f admet un minimum sur $S(0,1)$ et que celui-ci est une valeur propre de u.*

3. *Soient $\alpha_1, ..., \alpha_n \in \mathbb{R}_+^*$ tels que $\sum_{i=1}^{n} \alpha_i = 1$ et $f : \begin{cases} (\mathbb{R}_+)^n & \to & \mathbb{R} \\ (x_1, ..., x_n) & \mapsto & x_1^{\alpha_1} ... x_n^{\alpha_n} \end{cases}$.*

On définit $\Gamma = \{x = (x_1, ..., x_n) \in (\mathbb{R}_+)^n,\ \sum_{i=1}^{n} \alpha_i x_i = 1\}$.

(a) *Montrer qu'il existe $a = (a_1, ..., a_n) \in \Gamma$ tel que $f(a) = \sup_{x \in \Gamma} f(x)$ $(*)$, avec les a_i dans $\mathbb{R}_+^*$.*

(b) *Déterminer tous les points a vérifiant $(*)$ et en déduire que : $\forall x \in \Gamma,\ ,\ x_1^{\alpha_1} ... x_n^{\alpha_n} \leq 1$.*

(c) *Montrer que : $\forall x \in (\mathbb{R}_+)^n,\ x_1^{\alpha_1} ... x_n^{\alpha_n} \leq \sum_{i=1}^{n} \alpha_i x_i$.*

(a) *La fonction $g : (\mathbb{R}_+)^n \to \mathbb{R}, x \mapsto \alpha_1 x_1 + \cdots + \alpha_n x_n - 1$ étant polynomiale, est continue. Donc $\Gamma = g^{-1}(\{0\})$ est un fermé de $\mathbb{R}^n$.*

D'autre part, on a : $\forall i \in [\![1,n]\!], 0 \leqslant \alpha_i x_i \leqslant 1$. D'où :

$\forall i \in [\![1,n]\!], 0 \leqslant x_i \leqslant \frac{1}{\alpha_i}$ (car $\alpha_i > 0$).

Il s'ensuit que Γ est aussi borné, donc c'est un compact de $\mathbb{R}^n$ (on est en dimension finie). Sur Γ, la fonction f est continue par théorème sur les *opérations sur les fonctions continues ; d'où l'existence de $a \in \Gamma$ tel que $f(a) = \sup_{x\in\Gamma} f(x) = \max_{x\in\Gamma} f(x)$.*

Si un des a_i est nul, alors $f(a) = 0$. Or $f(1, \dots, 1) = 1$, ce qui contredit le fait que f ait un maximum en a. Donc $\forall i \in [\![1, n]\!]$, $a_i > 0$.

(b)

(c)

Remarque 6.3.1 *Parfois, on peut éviter le recours à ce théorème lorsque l'on peut paramétrer M. Reprenons le premier exemple avec f l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par*
$f : (x, y) \mapsto 4x^2 + 12xy - y^2$.
Soit $C = \left\{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 13\right\}$. On cherche les extremums de f sur C.
On a $C = \left\{\left(\sqrt{13}\cos(\theta), \sqrt{13}\sin(\theta)\right),\ \theta \in \mathbb{R}\right\}$.
On a : $\forall\theta \in \mathbb{R},\ f\left(\sqrt{13}\cos(\theta), \sqrt{13}\sin(\theta)\right) = 13\left(4\cos^2(\theta) + 12\cos(\theta)\sin(\theta) - \sin^2(\theta)\right) =$
$13\left(2(1 + \cos(2\theta)) + 6\sin(2\theta) - \dfrac{1 - \cos(2\theta)}{2}\right) = 13\left(\dfrac{3}{2} + \dfrac{5}{2}\cos(\theta) + 6\sin(2\theta)\right) =$
$13\left(\dfrac{3}{2} + \sqrt{\left(\dfrac{5}{2}\right)^2 + 6^2}.\cos(\theta - \varphi)\right) = \dfrac{39}{2} + \dfrac{169}{2}\cos(\theta - \varphi)$.
Il est donc clair que $\min_C f = -65$ et $\max_C f = 104$.

7 Compléments : Laplacien et fonctions harmoniques

Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{C}$ de classe $\mathcal{C}^2$, avec $n \geq 2$. On appelle Laplacien de f la fonction $\Delta(f) = \displaystyle\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$.
On dit qu'une telle fonction est harmonique si $\Delta(f) = 0$.

1. Soit $g : \mathbb{R}_+^* \to \mathbb{R}$ de classe $\mathcal{C}^2$. Soit f définie sur $\mathbb{R}^n \setminus \{0\}$ définie par :

   $\forall(x_1, ..., x_n) \in \mathbb{R}^n \setminus \{0\},\ f(x_1, ..., x_n) = g\left(\sqrt{\displaystyle\sum_{i=1}^{n} x_i^2}\right)$. Déterminer toutes les fonctions g telles que $\Delta(f) = 0$.
2. Dans cette question, on se place dans $\mathbb{R}^2$.

   (a) On pose $g(r, \theta) = f(r\cos\theta, r\sin\theta)$. Exprimer le laplacien de f en coordonnées polaires (c'est-à-dire à l'aide de g).

(b) On suppose que $U = \mathbb{R}^2$ et que f est harmonique. Montrer que $r \mapsto \displaystyle\int_0^{2\pi} f(r\cos(\theta), r\sin(\theta))d\theta$ est constante sur r.

(c) Soit $\sum a_n z^n$ une série entière de rayon de convergence $R > 0$ et on pose $U = \{(x,y) \in \mathbb{R}^2,\ |x+iy| < R\}$.
Montrer que $g : (x,y) \mapsto \displaystyle\sum_{n=0}^{+\infty} a_n(x+iy)^n$ est de classe $\mathcal{C}^2$ sur U et que g est harmonique.

3. **(Principe du maximum)** On se place sur $\mathbb{R}^n$, on note $B = B(0,1), \overline{B} = \overline{B}(0,1)$ et $S = S(0,1)$. Soit $f : \overline{B} \to \mathbb{R}$ continue et de classe $\mathcal{C}^2$ sur B.

   (a) Comme f est continue sur le compact $\overline{B}$, alors admet un maximum en un point $x_0 \in \overline{B}$. On suppose de plus que : $\forall x \in B, \Delta f(x) > 0$.
   Montrer que x_0 appartient à S et en déduire que : $\forall x \in B,\ f(x) < \sup\limits_{y \in S} f(y)$.

   (b) On suppose que f harmonique sur B.
   Pour tout $\varepsilon > 0$, on pose $g_\varepsilon(x) = f(x) + \varepsilon\|x\|^2$, qui est continue sur $\overline{B}$, de classe $\mathcal{C}^2$ sur B.

      i. Montrer que : $\forall x \in B,\ \Delta g_\varepsilon(x) > 0$.
      ii. En déduire que $\forall x \in B,\ f(x) \leq \sup\limits_{y \in S} f(y)$.
      iii. Montrer que si f est nulle sur S, alors f est nulle.