

# CALCUL DIFFÉRENTIEL

B. Landelle

## Table des matières

<b>I</b>	<b>Différentielle</b>	<b>2</b>
1	Définitions . . . . .	2
2	Propriétés . . . . .	3
3	Opérations . . . . .	4
<b>II</b>	<b>Dérivée selon un vecteur, dérivée partielle</b>	<b>7</b>
1	Dérivée selon un vecteur . . . . .	7
2	Dérivée partielle, gradient . . . . .	8
3	Matrice jacobienne . . . . .	11
4	Opérations . . . . .	12
<b>III</b>	<b>Fonctions de classe <math>\mathcal{C}^1</math></b>	<b>14</b>
1	Définitions, propriétés . . . . .	14
2	Intégration le long d'un chemin . . . . .	16
3	Vecteurs tangents à une partie . . . . .	17
<b>IV</b>	<b>Optimisation, premier ordre</b>	<b>19</b>
1	Extremums locaux, globaux . . . . .	19
2	Point critique . . . . .	20
3	Optimisation sous contrainte . . . . .	22
<b>V</b>	<b>Fonctions de classe <math>\mathcal{C}^k</math></b>	<b>24</b>
1	Définitions, propriétés . . . . .	24
2	Opérations . . . . .	25
3	Théorème de Schwarz . . . . .	27
<b>VI</b>	<b>Optimisation, deuxième ordre</b>	<b>27</b>
1	Matrice Hessienne . . . . .	27
2	Étude au deuxième ordre . . . . .	28

Dans tout ce qui suit, l'ensemble  $I$  désigne un intervalle ouvert non vide de  $\mathbb{R}$  et les ensembles  $E, F, G, H$  des  $\mathbb{R}$ -ev normés de dimensions finies.

**Notations :** Soit  $g : U \rightarrow \mathbb{R}$  avec  $U \subset E$  et  $a \in U$ . Pour  $f : U \rightarrow F$ , on note  $f(x) = o(g(x))$  si, pour  $x$  au voisinage de  $a$ , on a  $f(x) = g(x)\varepsilon(x)$  avec  $\varepsilon(x) \xrightarrow{x \rightarrow a} 0_F$ . La notation  $o(1)$  pour  $x \rightarrow a$  désigne une fonction de limite nulle dans  $F$  en  $a$ . Dans ce qui suit, on notera abusivement  $o(h)$  pour  $o(\|h\|)$  ou  $\|h\|o(1)$  lorsque  $h \rightarrow 0_E$ .

# I Différentielle

## 1 Définitions

**Définition 1.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$ . On dit que  $f$  différentiable en  $a$  s'il existe  $\ell \in \mathcal{L}(E, F)$  telle que

$$f(a+h) \underset{h \rightarrow 0}{=} f(a) + \ell(h) + o(h)$$

On dit que  $\ell$  est une application linéaire tangente à  $f$  en  $a$ .

**Définition 2.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$ . On appelle développement limité de  $f$  à l'ordre 1 en  $a$  une égalité de la forme

$$f(a+h) = f(a) + \ell(h) + o(h) \quad \text{avec} \quad \ell \in \mathcal{L}(E, F)$$

**Proposition 1.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ . Si  $f$  admet un développement limité en  $a$  à l'ordre 1, alors celui-ci est unique.

*Démonstration.* Soit  $a \in U$  et  $x \in E$ . On a clairement  $tx \xrightarrow{t \rightarrow 0} 0_E$ . Considérons deux développements limités de  $f$  en  $a$  et montrons qu'ils coïncident. On note

$$f(a+h) \underset{h \rightarrow 0}{=} f(a) + \ell_1(h) + o(h) \quad \text{et} \quad f(a+h) \underset{h \rightarrow 0}{=} f(a) + \ell_2(h) + o(h)$$

avec  $\ell_1, \ell_2 \in \mathcal{L}(E, F)$ . Il s'ensuit, en considérant une asymptotique pour  $t \xrightarrow{>} 0$

$$(\ell_1 - \ell_2)(tx) \underset{t \rightarrow 0}{=} o(tx) = t\|x\|o(1)$$

$$\text{d'où} \quad (\ell_1 - \ell_2)(x) \underset{t \rightarrow 0}{=} \|x\|o(1) \xrightarrow{t \rightarrow 0} 0$$

ce qui prouve  $\ell_1 = \ell_2$ . □

**Définition 3.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$  différentiable en  $a$ . L'application linéaire tangente à  $f$  en  $a$  est appelée différentielle de  $f$  en  $a$  et notée  $df(a) \in \mathcal{L}(E, F)$  avec

$$df(a): \begin{cases} E \longrightarrow F \\ h \longmapsto df(a) \cdot h \end{cases}$$

**Remarque :** On peut parler de l'application linéaire tangente (et non d'une application) par unicité de celle-ci.

**Définition 4.** Soit  $U$  ouvert de  $E$  et  $f : U \rightarrow F$ . On dit que  $f$  est différentiable sur  $U$  si  $f$  est différentiable en tout point de  $U$  et on note  $df$  l'application différentielle définie par

$$df: \begin{cases} U \longrightarrow \mathcal{L}(E, F) \\ a \longmapsto df(a) \end{cases}$$

## 2 Propriétés

**Proposition 2.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$  différentiable en  $a$ . Alors, la fonction  $f$  admet le développement limité à l'ordre 1 en  $a$

$$f(a+h) \underset{h \rightarrow 0}{=} f(a) + df(a) \cdot h + o(h)$$

*Démonstration.* Immédiate. □

**Proposition 3.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $\mathcal{L} = (\varepsilon_1, \dots, \varepsilon_m)$  une base de  $F$  et  $f = \sum_{j=1}^m f_j \varepsilon_j : U \rightarrow F$  avec les  $f_j$  fonctions coordonnées. On a

$$f \text{ différentiable en } a \iff f_j \text{ différentiable en } a \text{ pour } j \in \llbracket 1; m \rrbracket$$

et dans ce cas

$$df(a) = \sum_{j=1}^m df_j(a) \varepsilon_j$$

*Démonstration.* On note  $(\varepsilon_j^*)_{1 \leq j \leq m}$  la base duale de  $(\varepsilon_j)_{1 \leq j \leq m}$  qui vérifie  $\varepsilon_j^*(\varepsilon_i) = \delta_{i,j}$  pour tout  $(i, j) \in \llbracket 1; m \rrbracket^2$ . Supposons  $f$  différentiable en  $a$ , on a

$$f(a+h) = f(a) + df(a) \cdot h + o(h)$$

Pour  $j \in \llbracket 1; m \rrbracket$ , en composant par  $\varepsilon_j^*$ , il vient par continuité et donc caractère lipschitzien en zéro de  $\varepsilon_j^*$  :

$$f_j(a+h) = f_j(a) + (\varepsilon_j^* \circ df(a)) \cdot h + o(h)$$

d'où la différentiabilité des fonctions coordonnées. Réciproquement, si pour tout  $j \in \llbracket 1; m \rrbracket$ , on a

$$f_j(a+h) = f_j(a) + df_j(a) \cdot h + o(h)$$

alors

$$f(a+h) = \sum_{i=1}^m f_i(a) \varepsilon_i + \left( \sum_{j=1}^m df_j(a) \varepsilon_j \right) \cdot h + o(h)$$

d'où le résultat. □

**Proposition 4.** Les applications constantes de  $E$  dans  $F$  sont différentiables en tout point de différentielle nulle.

*Démonstration.* Immédiate. □

**Proposition 5.** Soit  $f \in \mathcal{L}(E, F)$ . L'application  $f$  est différentiable en tout point  $a \in E$  avec  $df(a) = f$ .

*Démonstration.* On a  $f(a+h) = f(a) + f(h)$  d'où  $f$  différentiable en tout point  $a \in E$  avec  $df(a) = f$ . □

**Exemple :** Soit  $E = \mathcal{M}_n(\mathbb{R})$  et  $f : E \rightarrow E$  définie par  $f(M) = M^2$  pour  $M \in E$ . Pour  $A \in E$ , on a

$$\forall H \in E \quad f(A+H) = A^2 + AH + HA + H^2 = f(A) + AH + HA + H^2$$

En munissant  $E$  d'une norme sous-multiplicative, on a

$$\forall H \in E \quad \|H^2\| \leq \|H\|^2 = o(\|H\|)$$

d'où  $\forall H \in E \quad f(A+H) = f(A) + \ell(H) + o(H) \quad \text{avec} \quad \ell(H) = AH + HA$


On conclut que  $f$  est différentiable en  $A$  avec  $df(A) : E \rightarrow E, H \mapsto AH + HA$ .

**Proposition 6.** Soit  $f : I \rightarrow F$  et  $a \in I$ . On a

$$f \text{ dérivable en } a \iff f \text{ différentiable en } a$$

et dans ce cas

$$\forall h \in \mathbb{R} \quad df(a) \cdot h = f'(a)h$$

 **Avertissement :** Cette équivalence a lieu pour  $f$  définie sur un intervalle.

*Démonstration.* Supposons  $f$  dérivable en  $a$ . On a

$$f(a+h) = f(a) + f'(a)h + o(h)$$

d'où  $f$  différentiable en  $a$  et  $\forall h \in \mathbb{R} \quad df(a) \cdot h = f'(a)h$

Réciproquement, supposons  $f$  différentiable en  $a$ . On a

$$f(a+h) = f(a) + df(a) \cdot h + o(h)$$

d'où, pour  $h \neq 0$   $\frac{f(a+h) - f(a)}{h} = df(a) \cdot 1 + o(1) \xrightarrow{h \rightarrow 0} df(a) \cdot 1$

Ainsi, la fonction  $f$  est dérivable avec  $f'(a) = df(a) \cdot 1$ . □

**Proposition 7.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$ . Si  $f$  est différentiable en  $a$ , alors  $f$  est continue en  $a$ .

*Démonstration.* Immédiat par continuité de  $df(a)$  comme application linéaire sur un espace de dimension finie. □

### 3 Opérations

**Proposition 8.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $f, g : U \rightarrow F$  et  $\lambda$  réel. Si  $f$  et  $g$  sont différentiables en  $a$ , alors l'application  $f + \lambda g$  est différentiable en  $a$  avec

$$d(f + \lambda g)(a) = df(a) + \lambda dg(a)$$

*Démonstration.* On a

$$(f + \lambda g)(a+h) = (f + \lambda g)(a) + (df(a) + \lambda dg(a)) \cdot h + o(h)$$

avec  $df(a) + \lambda dg(a) \in \mathcal{L}(E, F)$ . Ainsi, l'application  $f + \lambda g$  est différentiable en  $a$  avec  $d(f + \lambda g)(a) = df(a) + \lambda dg(a)$ . □

**Théorème 1.** Soit  $U$  ouvert de  $E$ . L'ensemble des fonctions différentiables de  $U$  dans  $F$  est un sev de  $\mathcal{F}(U, F)$ .

*Démonstration.* L'application nulle est différentiable et on a la stabilité par combinaison linéaire. □

**Théorème 2.** Soit  $U$  ouvert de  $E$ ,  $f : U \rightarrow F$ ,  $g : U \rightarrow G$  et  $B : F \times G \rightarrow H$  une application bilinéaire. Si  $f$  et  $g$  sont différentiables sur  $U$ , alors l'application  $B(f, g)$  est différentiable sur  $U$  avec

$$dB(f, g) = B(df, g) + B(f, dg)$$

*Démonstration.* Soit  $a \in U$ . Par linéarité sur la première variable puis sur la seconde variable en ne détaillant que les termes d'ordre au plus 1, on trouve

$$\begin{aligned} B(f, g)(a + h) &= B(f(a + h), g(a + h)) \\ &= B(f(a) + df(a) \cdot h + o(h), g(a) + dg(a) \cdot h + o(h)) \\ B(f, g)(a + h) &= B(f(a), g(a)) + B(df(a) \cdot h, g(a)) + B(f(a), dg(a) \cdot h) + \\ &\quad B(f(a), o(h)) + B(df(a) \cdot h, \underbrace{dg(a) \cdot h + o(h)}_{=o(1)}) + B(o(h), g(a + h)) \end{aligned}$$

On a  $B(df(a), g(a)) + B(f(a), dg(a)) : h \mapsto B(df(a) \cdot h, g(a)) + B(f(a), dg(a) \cdot h) \in \mathcal{L}(E, H)$ . Il reste à vérifier que les termes qui suivent forment un  $o(h)$ . Comme  $df(a)$  est linéaire sur  $E$  de dimension finie et comme  $B$  est bilinéaire sur  $E^2$  produit d'espaces de dimension finie, il existe  $C_1$  et  $C_2$  positives telles que

$$\forall (x, y, z) \in E \times F \times G \quad \|df(a) \cdot x\| \leq C_1 \|x\| \quad \|B(y, z)\| \leq C_2 \|y\| \|z\|$$

On en déduit que l'expression restant du développement précédent est majorée en norme par

$$C_2 \|f(a)\| \|h\| o(1) + C_2 C_1 \|h\| o(1) + C_2 \|h\| o(1) \|g(a + h)\| = \|h\| o(1)$$

Autrement dit, on a

$$B(f, g)(a + h) = B(f(a), g(a)) + B(df(a) \cdot h, g(a)) + B(f(a), dg(a) \cdot h) + o(h)$$

□

**Corollaire 1.** Soit  $U$  ouvert de  $E$ ,  $f, g : U \rightarrow F$  avec  $F$  une algèbre. Si  $f$  et  $g$  sont différentiables sur  $U$ , alors l'application  $fg$  l'est avec

$$d(fg) = (df)g + f(dg)$$

*Démonstration.* On considère  $B : F^2 \rightarrow F, (x, y) \mapsto xy$ .

□

**Théorème 3.** Soit  $U$  ouvert de  $E$ ,  $F_1, \dots, F_p$  des  $\mathbb{R}$ -ev normés de dimensions finies,  $f_i : U \rightarrow F_i$  différentiable sur  $U$  et  $M : \prod_{i=1}^p F_i \rightarrow G$  application  $p$ -linéaire avec  $p \geq 2$ . L'application  $M(f_1, \dots, f_p) : U \rightarrow G, x \mapsto M(f_1(x), \dots, f_p(x))$  est différentiable sur  $U$  avec

$$dM(f_1, \dots, f_p) = M(df_1, f_2, \dots, f_p) + M(f_1, df_2, \dots, f_p) + \dots + M(f_1, \dots, f_{p-1}, df_p)$$

*Démonstration.* On procède par récurrence sur  $p$ . On a

$$\begin{aligned} M(f_1(a + h), \dots, f_{p+1}(a + h)) &= \\ M(f_1(a + h), \dots, f_p(a + h), f_{p+1}(a)) &+ M(f_1(a + h), \dots, f_p(a + h), df_{p+1}(a) \cdot h) + \\ &M(f_1(a + h), \dots, f_p(a + h), o(h)) \end{aligned}$$

Puis  $\|M(f_1(a + h), \dots, f_p(a + h), o(h))\| \leq C \|h\| o(1) \prod_{i=1}^p \|f_i(a + h)\| = o(h)$

Par hypothèse de récurrence, il vient

$$\begin{aligned} M(f_1(a + h), \dots, f_p(a + h), f_{p+1}(a)) &= \\ M(f_1, \dots, f_{p+1})(a) &+ M(df_1, f_2, \dots, f_{p+1})(a) \cdot h + \dots + M(f_1, \dots, df_p, f_{p+1})(a) \cdot h + o(h) \end{aligned}$$

Puis, on décompose dans  $\mathcal{B} = (e_1, \dots, e_n)$  base de  $E$  avec  $h = \sum_{i=1}^n h_i e_i$  avec les  $h_i$  réels

$$\begin{aligned} M(f_1(a+h), \dots, f_p(a+h), df_{p+1}(a) \cdot h) &= \sum_{i=1}^n h_i \underbrace{M(f_1(a+h), \dots, f_p(a+h), df_{p+1}(a) \cdot e_i)}_{=M(f_1(a), \dots, f_p(a), df_{p+1}(a) \cdot e_i) + o(1)} \\ &= M(f_1(a), \dots, f_p(a), df_{p+1}(a) \cdot h) + o(h) \end{aligned}$$

Le résultat suit.  $\square$

**Théorème 4.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $V$  ouvert de  $F$ ,  $f : U \rightarrow F$  et  $g : V \rightarrow G$  avec  $f(U) \subset V$ . Si  $f$  et  $g$  sont différentiables respectivement en  $a$  et  $f(a)$ , alors  $g \circ f$  est différentiable en  $a$  avec

$$d(g \circ f)(a) = dg(f(a)) \circ df(a)$$

*Démonstration.* On a

$$\begin{aligned} g \circ f(a+h) &= g(f(a) + df(a) \cdot h + o(h)) \\ &= g \circ f(a) + dg(f(a)) \cdot (df(a) \cdot h + o(h)) + o(df(a) \cdot h + o(h)) \\ &= g \circ f(a) + [dg(f(a)) \circ df(a)] \cdot h + dg(f(a)) \cdot o(h) + o(df(a) \cdot h + o(h)) \end{aligned}$$

Les applications  $df(a)$  et  $dg(f(a))$  sont lipschitziennes de constantes respectives  $C_1, C_2 \geq 0$  d'où

$$\|dg(f(a)) \cdot o(h)\| \leq C_2 \|h\| o(1)$$

$$\text{et } \|o(df(a) \cdot h + o(h))\| = \|df(a) \cdot h + o(h)\| o(1) \leq (C_1 \|h\| + \|h\| o(1)) o(1) = \|h\| o(1)$$

$$\text{Ainsi } g \circ f(a+h) = g \circ f(a) + [dg(f(a)) \circ df(a)] \cdot h + o(h)$$

d'où le résultat.  $\square$

**Corollaire 2.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $\varphi : I \rightarrow \mathbb{R}$  et  $f : U \rightarrow \mathbb{R}$  telle que  $f(U) \subset I$ . Si  $f$  est différentiable en  $a$  et  $\varphi$  dérivable en  $f(a)$ , alors  $\varphi \circ f$  est différentiable en  $a$  avec

$$d(\varphi \circ f)(a) = \varphi' \circ f(a) df(a)$$

*Démonstration.* D'après la proposition 6, on a  $\varphi$  différentiable en  $f(a)$  puis, d'après le théorème 4, il vient pour  $h \in E$

$$d(\varphi \circ f)(a) \cdot h = d\varphi(f(a)) \circ df(a) \cdot h = \varphi' \circ f(a) df(a) \cdot h$$

d'où le résultat.  $\square$

**Remarque :** L'ensemble  $I$  est un intervalle d'où l'usage licite de la dérivabilité.

**Applications :** On en déduit  $d(f^n) = n f^{n-1} df$ ,  $d\left(\frac{1}{f}\right) = -\frac{1}{f^2} df$  si  $f$  ne s'annule pas, etc. ...

Les applications coordonnées sur  $\mathbb{R}^n$  sont clairement différentiables puisque linéaires. Par produit et combinaison linéaire, il s'ensuit que les applications polynomiales le sont aussi et par composition et produit, les fonctions rationnelles le sont sur leur ensemble de définition (nécessairement ouvert ...)

**Remarque :** Dans la notation  $\int_a^b f(x) dx$  avec  $f \in \mathcal{C}^0([a; b], \mathbb{R})$ , il est d'usage de voir le terme  $dx$  comme un élément infinitésimal. Cette notation est compatible avec le point de vue d'une

différentielle lors d'un changement de variables : si on pose  $x = \varphi(t)$  avec  $\varphi \in \mathcal{C}^1(J, I)$  où  $I, J$  sont des intervalles de  $\mathbb{R}$  avec  $(a, b) \in (\text{Im } \varphi)^2$ , en considérant  $t = \text{id}$ , il vient

$$dx = d(\varphi \circ t) = \varphi' \circ t \, dt = \varphi'(t) \, dt$$

**Corollaire 3 (Dérivée le long d'un arc).** Soit  $U$  ouvert de  $E$ ,  $f : U \rightarrow F$  et  $\gamma : I \rightarrow E$  avec  $\gamma(I) \subset U$ . Si  $\gamma$  est dérivable en  $t \in I$  et si  $f$  est différentiable en  $\gamma(t)$ , alors  $f \circ \gamma$  est dérivable en  $t$  et

$$(f \circ \gamma)'(t) = df(\gamma(t)) \cdot \gamma'(t)$$

*Démonstration.* On  $\gamma$  dérivable en  $t$  avec  $d\gamma(t) = h \mapsto \gamma'(t)h$ . D'après le théorème précédent, on a  $f \circ \gamma$  dérivable en  $t$  avec

$$d(f \circ \gamma)(t) = df(\gamma(t)) \circ d\gamma(t) = df(\gamma(t)) \circ (h \mapsto \gamma'(t)h)$$

On conclut avec  $(f \circ \gamma)'(t) = d(f \circ \gamma)(t) \cdot 1$ . □

**Remarque :** En particulier, pour  $\gamma : t \mapsto x + th$ , on a

$$\frac{d}{dt}f(x + th) = df(x + th) \cdot h$$

## II Dérivée selon un vecteur, dérivée partielle

### 1 Dérivée selon un vecteur

**Définition 5.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $v \in E$  et  $f : U \rightarrow F$ . On dit que  $f$  est dérivable en  $a$  selon le vecteur  $v$  si la fonction  $t \mapsto f(a + tv)$  est dérivable en zéro. On note alors

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

qu'on appelle vecteur dérivée de  $f$  en  $a$  selon le vecteur  $v$ .

**Remarque :** Pour  $v \in E$ , la fonction  $t \mapsto f(a + tv)$  est définie sur un intervalle ouvert centré en zéro. C'est immédiat si  $v = 0_E$  et sinon, comme  $a \in U$  avec  $U$  ouvert, il existe  $r > 0$  tel que  $B(a, r) \subset U$  et

$$a + tv \in B(a, r) \iff t \in \left] -\frac{r}{\|v\|} ; \frac{r}{\|v\|} \right[$$

**Définition 6.** Soit  $U$  ouvert de  $E$ ,  $v \in E$  et  $f : U \rightarrow F$ . On dit que  $f$  est dérivable sur  $U$  selon le vecteur  $v$  si  $f$  dérivable en tout point de  $U$  selon le vecteur  $v$  et on note  $D_v f$  l'application dérivée selon  $v$  définie par

$$D_v f : \begin{cases} U \longrightarrow F \\ a \longmapsto D_v f(a) \end{cases}$$

**Théorème 5.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$ . Si  $f$  est différentiable en  $a$ , alors l'application  $f$  est dérivable en  $a$  selon tout vecteur  $v \in E$  et on a

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

*Démonstration.* Soit  $v \in E$ . Comme  $tv \xrightarrow[t \rightarrow 0]{} 0_E$ , on a

$$f(a + tv) \underset{t \rightarrow 0}{=} f(a) + df(a) \cdot tv + |t| \|v\| o(1)$$

Comme  $|t|/t$  est bornée pour  $t \neq 0$ , il vient

$$\frac{f(a + tv) - f(a)}{t} = df(a) \cdot v + o(1)$$

□

**Remarque :** La réciproque est fausse. Considérons la fonction  $f$  définie sur  $\mathbb{R}^2$  par

$$\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = \begin{cases} \frac{x^3}{y} & \text{si } y \neq 0 \\ 0 & \text{sinon} \end{cases}$$

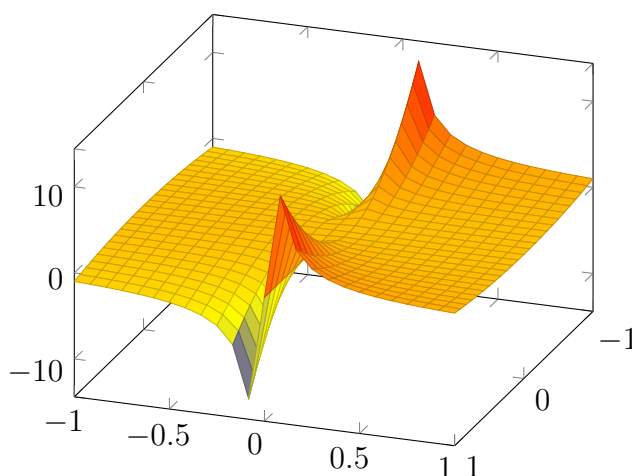


FIGURE 1 – Tracé de la surface d'équation  $z = f(x, y)$

Soit  $y \neq 0$ . On a

$$\frac{f(tx, ty) - f(0, 0)}{t} = t \frac{x^3}{y} \xrightarrow[t \rightarrow 0]{} 0$$

et pour  $y = 0$

$$\frac{f(tx, ty) - f(0, 0)}{t} = 0$$

Donc  $f$  est dérivable selon tout vecteur en  $(0, 0)$ . Cependant, on a

$$(x, x^3) \xrightarrow[x \rightarrow 0]{} (0, 0) \quad \text{et} \quad \forall x \neq 0 \quad f(x, x^3) = 1 \not\xrightarrow[t \rightarrow 0]{} f(0, 0)$$

ce qui prouve que  $f$  n'est pas continue en  $(0, 0)$  et ne peut donc être différentiable en  $(0, 0)$ . Cet exemple illustre également le fait qu'une fonction peut être dérivable selon tout vecteur en un point sans y être continue.

## 2 Dérivée partielle, gradient

Dans ce qui suit, on suppose que  $\mathcal{B} = (e_1, \dots, e_n)$  est une base de  $E$  et  $\mathcal{L} = (\varepsilon_1, \dots, \varepsilon_m)$  une base de  $F$ . Pour  $x \in E$ , on note  $x = \sum_{i=1}^n x_i e_i$  sa décomposition dans  $\mathcal{B}$ . Les  $x_i$  sont les coordonnées de  $x$  dans  $\mathcal{B}$ . Pour  $f : U \rightarrow F$  avec  $U \subset E$ , on s'autorise à confondre les écritures  $f(x)$  et  $f(x_1, \dots, x_n)$ .



**Définition 7.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $f : U \rightarrow F$  et  $i \in \llbracket 1; n \rrbracket$ . On dit que  $f$  admet une  $i$ -ème dérivée partielle en  $a$  dans la base  $\mathcal{B}$  si  $f$  est dérivable en  $a$  selon  $e_i$ , i.e.  $t \mapsto f(a + te_i)$  dérivable en zéro. On note cette  $i$ -ème dérivée partielle  $\partial_i f(a)$  ou encore  $\frac{\partial f}{\partial x_i}(a)$ , c'est-à-dire

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

**Rermarque :** La notation  $\frac{\partial f}{\partial x_i}(a)$  est malheureuse puisqu'il n'y a aucune raison *a priori* que  $x_i$  désigne la  $i$ -ème coordonnée selon  $e_i$ . Mais l'usage s'est répandu d'où la nécessité de connaître et manipuler cette notation.

**Définition 8.** Soit  $U$  ouvert de  $E$ ,  $f : U \rightarrow F$  et  $i \in \llbracket 1; n \rrbracket$ . On dit que  $f$  admet une  $i$ -ème dérivée partielle sur  $U$  dans la base  $\mathcal{B}$  si  $f$  est dérivable en tout point de  $U$  selon  $e_i$ . On note cette  $i$ -ème dérivée partielle  $\partial_i f$  ou encore  $\frac{\partial f}{\partial x_i}$ , c'est-à-dire

$$\partial_i f = \frac{\partial f}{\partial x_i} : \begin{cases} U \longrightarrow F \\ a \longmapsto D_{e_i} f(a) \end{cases}$$

**Proposition 9.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f = \sum_{j=1}^m f_j \varepsilon_j : U \rightarrow F$  avec les  $f_j$  fonctions coordonnées. Pour  $i \in \llbracket 1; n \rrbracket$ , on a

$f$  admet une  $i$ -ème dérivée partielle en  $a \iff f_j$  admet une  $i$ -ème dérivée partielle en  $a$  pour tout  $j \in \llbracket 1; m \rrbracket$

et dans ce cas

$$\partial_i f(a) = \sum_{j=1}^m \partial_i f_j(a) \varepsilon_j$$

*Démonstration.* Conséquence des théorèmes sur les fonctions vectorielles. □

**Théorème 6.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$ . Si  $f$  est différentiable en  $a$ , alors ses dérivées partielles en  $a$  dans la base  $\mathcal{B}$  existent et on a

$$\forall i \in \llbracket 1; n \rrbracket \quad \partial_i f(a) = df(a) \cdot e_i$$

et

$$\forall h = \sum_{i=1}^n h_i e_i \in E \quad df(a) \cdot h = D_h f(a) = \sum_{i=1}^n h_i \partial_i f(a)$$

*Démonstration.* L'existence des dérivées partielles et leur expression fonction de  $df$  est une conséquence du théorème 5. La dernière égalité vient par linéarité de  $df(a)$  pour  $a \in U$  :

$$df(a) \cdot h = df(a) \cdot \left( \sum_{i=1}^n h_i e_i \right) = \sum_{i=1}^n h_i df(a) \cdot e_i = \sum_{i=1}^n h_i \partial_i f(a)$$

□

**Remarques :** (1) On peut aussi écrire  $df(a) = \sum_{i=1}^n \partial_i f(a) e_i^*$  avec  $(e_i^*)_{1 \leq i \leq n}$  base duale de  $\mathcal{B}$

(également couramment noté  $df(a) = \sum_{i=1}^n \partial_i f(a) dx_i$  avec  $(dx_i)_{1 \leq i \leq n}$  base duale de  $\mathcal{B}$ ).

(2) Comme pour le théorème 5, la réciproque est fausse. Il suffit de considérer le contre-exemple correspondant.

**Corollaire 4.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$ . Si  $f$  est différentiable en  $a$ , on a le développement limité à l'ordre 1

$$f(a+h) = f(a) + \sum_{i=1}^n h_i \partial_i f(a) + o(h)$$

*Démonstration.* Immédiate. □

**Proposition 10.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $f : U \rightarrow F$  et  $i \in \llbracket 1; n \rrbracket$ . Si  $f$  admet une  $i$ -ème dérivée partielle en  $a$ , alors on a

$$\partial_i f(a) = \frac{\partial f}{\partial x_i}(a) = \frac{d}{dx_i} [f(a_1, \dots, x_i, \dots, a_n)]|_{x_i=a_i}$$

Plus généralement, si  $f$  admet une  $i$ -ème dérivée partielle sur  $U$ , on a

$$\forall x \in U \quad \partial_i f(x) = \frac{\partial f}{\partial x_i}(x) = \frac{d}{dx_i} [f(x_1, \dots, x_i, \dots, x_n)]$$

*Démonstration.* Conséquence de la définition de la dérivée de  $f$  en  $a$  selon  $e_i$  :

$$\frac{1}{t} [f(a_1, \dots, a_i + t, \dots, a_n) - f(a)] \xrightarrow{t \rightarrow 0} \frac{d}{dx_i} [f(a_1, \dots, x_i, \dots, a_n)]|_{x_i=a_i}$$

□

**Commentaire :** En pratique, pour calculer la  $i$ -ème dérivée partielle d'une fonction (si elle existe), on fige toutes les autres variables et on dérive en  $x_i$ .

**Exemple :** Soit  $f$  définie sur  $\mathbb{R}^2$  par

$$\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = \begin{cases} \frac{x^3}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

L'application est différentiable sur  $\mathbb{R}^2 \setminus \{(0, 0)\}$  (fonction rationnelle). Par exemple, pour calculer  $\frac{\partial f}{\partial x}(x, y)$ , on dérive  $f$  en la variable  $x$  en considérant  $y$  constante. On trouve

$$\forall (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\} \quad \frac{\partial f}{\partial x}(x, y) = \frac{3x^2(x^2 + y^2) - 2x^4}{(x^2 + y^2)^2} = \frac{x^2(x^2 + 3y^2)}{(x^2 + y^2)^2}$$

Pour savoir si  $f$  admet en  $(0, 0)$  des dérivées partielles selon  $x$  et selon  $y$ , on étudie les limites des taux d'accroissements suivants :

$$\begin{aligned} \frac{f(x, 0) - f(0, 0)}{x - 0} = 1 &\xrightarrow{x \rightarrow 0} 1 \implies \frac{\partial f}{\partial x}(0, 0) \text{ existe et } \frac{\partial f}{\partial x}(0, 0) = 1 \\ \frac{f(0, y) - f(0, 0)}{y - 0} = 0 &\xrightarrow{y \rightarrow 0} 0 \implies \frac{\partial f}{\partial y}(0, 0) \text{ existe et } \frac{\partial f}{\partial y}(0, 0) = 0 \end{aligned}$$

**Théorème 7.** Soit  $U$  ouvert de  $E$  euclidien,  $\mathcal{B} = (e_1, \dots, e_n)$  base orthonormée de  $E$ ,  $a \in U$  et  $f : U \rightarrow \mathbb{R}$ . Si  $f$  est différentiable en  $a$ , il existe un unique vecteur de  $E$  noté  $\nabla f(a)$  tel que

$$\forall h \in E \quad df(a) \cdot h = \langle \nabla f(a), h \rangle$$

et

$$\nabla f(a) = \sum_{i=1}^n \partial_i f(a) e_i$$

*Démonstration.* On a  $df(a) \in \mathcal{L}(E, \mathbb{R})$  donc l'existence et unicité de  $\nabla f(a)$  est une conséquence immédiate du théorème de représentation de Riesz. Par ailleurs, on a

$$\nabla f(a) = \sum_{i=1}^n \langle \nabla f(a), e_i \rangle e_i = \sum_{i=1}^n (df(a) \cdot e_i) e_i = \sum_{i=1}^n \partial_i f(a) e_i$$

□

**Définition 9.** Soit  $U$  ouvert de  $E$  euclidien,  $a \in U$  et  $f : U \rightarrow \mathbb{R}$  différentiable en  $a$ . Le vecteur  $\nabla f(a)$  est appelé gradient de l'application  $f$  en  $a$ .

**Exemple :** Soit  $f$  définie sur  $\mathbb{R}^2$  par  $f : (x, y) \mapsto x^2 + 3xy$ . On a

$$\forall (x, y) \in \mathbb{R}^2 \quad \nabla f(x, y) = (2x + 3y, 3x)$$

**Corollaire 5.** Soit  $U$  ouvert de  $E$  euclidien,  $a \in U$  et  $f : U \rightarrow \mathbb{R}$ . Si  $f$  est différentiable en  $a$ , on a le développement limité à l'ordre 1

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

*Démonstration.* Immédiate. □

**Proposition 11.** Soit  $U$  ouvert de  $E$  euclidien,  $a \in U$  et  $f : U \rightarrow \mathbb{R}$  différentiable en  $a$ . Si  $\nabla f(a) \neq 0_E$ , alors le vecteur unitaire selon lequel la dérivée de  $f$  en  $a$  est maximale est positivement colinéaire au gradient.

*Démonstration.* Soit  $h \in E$  normé. D'après l'inégalité de Cauchy-Schwarz, on a

$$D_h f(a) = df(a) \cdot h = \langle \nabla f(a), h \rangle \leq \|\nabla f(a)\|$$

et la dérivée de  $f$  en  $a$  selon  $h$  est donc maximale si l'inégalité est une égalité donc si  $h$  est positivement colinéaire à  $\nabla f(a)$ . □

**Remarque :** Cette propriété justifie l'intérêt de la *méthode du gradient*.

### 3 Matrice jacobienne

Dans ce qui suit, on suppose que  $\mathcal{B} = (e_1, \dots, e_n)$  est une base de  $E$  et  $\mathcal{L} = (\varepsilon_1, \dots, \varepsilon_m)$  une base de  $F$ . Les dérivées partielles sont relatives à la base  $\mathcal{B}$ . Pour  $f : U \rightarrow F$  avec  $U$  ouvert de  $E$ , on note  $f = \sum_{i=1}^m f_i \varepsilon_i$  avec les  $f_i$  applications coordonnées de  $f$  dans  $\mathcal{L}$ .

 **Avertissement :** On choisit les indices pour la cohérence avec ce qui suit.

**Proposition 12.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow F$ . Si  $f$  est différentiable en  $a$ , la matrice de  $df(a)$  dans les bases  $\mathcal{B}$  et  $\mathcal{L}$  est donnée par

$$\text{mat}_{\mathcal{B}, \mathcal{L}} df(a) = (\partial_j f_i(a))_{(i,j) \in \llbracket 1; m \rrbracket \times \llbracket 1; n \rrbracket}$$

*Démonstration.* Conséquence du théorème 6 et de la proposition 3

$$\forall h \in E \quad df(a) \cdot h = \sum_{j=1}^n \partial f_j(a) h_j = \sum_{i=1}^m \left( \sum_{j=1}^n \partial_j f_i(a) h_j \right) \varepsilon_i$$

□

**Définition 10.** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $a \in U$  et  $f = (f_1, \dots, f_m) : U \rightarrow \mathbb{R}^m$ . Si  $f$  est différentiable en  $a$ , on définit la matrice jacobienne de  $f$  en  $a$  notée  $J_f(a)$  comme matrice de  $df(a)$  dans les bases canoniques de  $\mathbb{R}^n$  et  $\mathbb{R}^m$ , c'est-à-dire

$$J_f(a) = (\partial_j f_i(a))_{(i,j) \in \llbracket 1; m \rrbracket \times \llbracket 1; n \rrbracket} = \left( \frac{\partial f_i}{\partial x_j}(a) \right)_{(i,j) \in \llbracket 1; m \rrbracket \times \llbracket 1; n \rrbracket}$$

**Exemples :** 1. Soit  $f(x, y) = (x + y, xy, x^2 + y^2)$  pour tout  $(x, y) \in \mathbb{R}^2$ . L'application  $f$  est à coordonnées polynomiales donc  $f$  différentiable. La matrice jacobienne en  $(x, y)$  est

$$J_f(x, y) = \begin{pmatrix} 1 & 1 \\ y & x \\ 2x & 2y \end{pmatrix}$$

Ainsi, pour  $a = (a_1, a_2) \in \mathbb{R}^2$ , notant  $X^\top = (x \ y)$ , on a

$$J_f(a)X = \begin{pmatrix} 1 & 1 \\ a_2 & a_1 \\ 2a_1 & 2a_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ a_2x + a_1y \\ 2(a_1x + a_2y) \end{pmatrix}$$

autrement dit  $\forall (x, y) \in \mathbb{R}^2 \quad df(a) \cdot (x, y) = (x + y, a_2x + a_1y, 2(a_1x + a_2y))$

2. Soit  $f(x, y) = x^2 + 3xy$  pour tout  $(x, y) \in \mathbb{R}^2$ . L'application  $f$  est polynomiale donc  $f$  différentiable. La matrice jacobienne en  $(x, y)$  est la matrice ligne

$$J_f(x, y) = (2x + 3y \quad 3x)$$

Ainsi, pour  $a = (a_1, a_2) \in \mathbb{R}^2$ , notant  $X^\top = (x \ y)$ , on a

$$J_f(a)X = (2a_1 + 3a_2 \quad 3a_1) \begin{pmatrix} x \\ y \end{pmatrix} = (2a_1 + 3a_2)x + 3a_1y$$

autrement dit  $\forall (x, y) \in \mathbb{R}^2 \quad df(a) \cdot (x, y) = (2a_1 + 3a_2)x + 3a_1y$

## 4 Opérations

On conserve les conventions fixées à la sous-partie précédente.

**Proposition 13.** Soit  $U$  ouvert de  $E$ ,  $f, g : U \rightarrow F$ ,  $a \in U$  et  $\lambda$  réel. Si  $f$  et  $g$  admettent des dérivées partielles en  $a$ , alors  $f + \lambda g$  également et

$$\forall j \in \llbracket 1; n \rrbracket \quad \partial_j(f + \lambda g)(a) = \partial_j f(a) + \lambda \partial_j g(a)$$

*Démonstration.* Immédiate. □

**Proposition 14.** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $f, g : U \rightarrow \mathbb{R}^m$  et  $a \in U$  et  $\lambda$  réel. Si  $f$  et  $g$  sont différentiables en  $a$ , alors

$$J_{f+\lambda g}(a) = J_f(a) + \lambda J_g(a)$$

*Démonstration.* Soit comme conséquence de ce qui précède, soit comme conséquence de la linéarité de la différentiation d. □

**Proposition 15.** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $f : U \rightarrow F$ ,  $g : U \rightarrow G$  et  $B : F \times G \rightarrow H$  une application bilinéaire. Si  $f$  et  $g$  admettent des dérivées partielles en  $a$ , alors  $B(f, g)$  également et

$$\forall j \in \llbracket 1; n \rrbracket \quad \partial_j B(f, g)(a) = B(\partial_j f, g)(a) + B(f, \partial_j g)(a)$$

*Démonstration.* Immédiat par propriétés sur des fonctions d'une variable réelle dérivable en un point appliquées à  $t \mapsto B(f(a + te_j), g(a + te_j))$  avec  $j \in \llbracket 1; n \rrbracket$ .  $\square$

**Théorème 8 (Règle de la chaîne ou dérivation composée).** Soit  $U$  ouvert de  $E$ ,  $a \in U$ ,  $V$  ouvert de  $F$ ,  $f : U \rightarrow F$  et  $g : V \rightarrow G$  avec  $f(U) \subset V$ . Si  $f$  et  $g$  sont différentiables respectivement en  $a$  et  $f(a)$ , alors  $g \circ f$  admet des dérivées partielles en  $a$  et

$$\forall j \in \llbracket 1; n \rrbracket \quad \partial_j g \circ f(a) = \sum_{i=1}^m \partial_i g(f(a)) \partial_j f_i(a)$$

*Démonstration.* On a  $d(g \circ f)(a) = dg(f(a)) \circ df(a)$

$$\begin{aligned} \text{et} \quad \forall j \in \llbracket 1; n \rrbracket \quad \partial_j(g \circ f)(a) &= d(g \circ f)(a) \cdot e_j \\ &= dg(f(a)) \cdot (df(a) \cdot e_j) = dg(f(a)) \cdot \partial_j f(a) \end{aligned}$$

Par ailleurs, on a 
$$\partial_j f(a) = \sum_{i=1}^m \partial_j f_i(a) \varepsilon_i$$

d'où 
$$\partial_j(g \circ f)(a) = \sum_{i=1}^m \partial_j f_i(a) dg(f(a)) \cdot \varepsilon_i = \sum_{i=1}^m \partial_j f_i(a) \partial_i g(f(a))$$

$\square$

**Remarque :** Ce résultat est souvent énoncé sous la forme suivante

$$\frac{d}{du_i} [f(x_1(u_1, \dots, u_m), \dots, x_n(u_1, \dots, u_m))] = \sum_{k=1}^n \partial_k f(x(u)) \partial_i x_k(u)$$

Avec des approximations au premier ordre, on peut retrouver intuitivement ce résultat :

$$f(x(u_1, \dots, u_i + \delta u_i, \dots, u_m)) \simeq f(x(u) + \partial_i x(u) \delta u_i) \simeq f(x(u)) + \delta u_i \sum_{k=1}^n \partial_k f(x(u)) \partial_i x_k(u)$$

**Corollaire 6 (Règle de la chaîne).** Soit  $U$  ouvert de  $E$ ,  $I$  intervalle ouvert de  $\mathbb{R}$ ,  $t \in I$ ,  $x : I \rightarrow E$  et  $f : U \rightarrow F$  avec  $x(I) \subset U$ . Si  $x$  est dérivable  $t$  et si  $f$  est différentiable en  $x(t)$ , alors  $f \circ x$  est dérivable en  $t$  et

$$(f \circ x)'(t) = \sum_{j=1}^n x'_j(t) \partial_j f(x(t))$$

*Démonstration.* Conséquence directe du résultat précédent (pour une fonction d'une seule variable, l'existence de dérivées partielles est simplement la dérivabilité).  $\square$

**Théorème 9.** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $a \in U$ ,  $V$  ouvert de  $\mathbb{R}^m$ ,  $f : U \rightarrow \mathbb{R}^m$  et  $g : V \rightarrow \mathbb{R}^p$  avec  $f(U) \subset V$ . Si  $f$  est différentiable en  $a$  et  $g$  différentiable en  $f(a)$ , alors

$$J_{g \circ f}(a) = J_g(f(a)) J_f(a)$$

*Démonstration.* Soit comme conséquence du théorème 8, soit comme écriture matricielle du théorème 4.  $\square$

### III Fonctions de classe $\mathcal{C}^1$

Dans ce qui qui, on note  $\mathcal{B} = (e_1, \dots, e_n)$  une base de  $E$ . Les dérivées partielles sont relatives à la base  $\mathcal{B}$ . Pour  $U$  ouvert de  $E$ , une application  $f : U \rightarrow F$  est dite de classe  $\mathcal{C}^0$  si elle est continue sur  $U$ . On note  $\mathcal{C}^0(U, F)$  l'ensemble des fonctions de classe  $\mathcal{C}^0$  sur  $U$ .

#### 1 Définitions, propriétés

**Définition 11.** Soit  $U$  ouvert de  $E$  et  $f : U \rightarrow F$ . L'application  $f$  est dite de classe  $\mathcal{C}^1$  sur l'ouvert  $U$  si elle est différentiable sur  $U$  et si  $df$  est continue sur  $U$ .

**Notations :** On note  $\mathcal{C}^1(U, F)$  l'ensemble des fonctions de classe  $\mathcal{C}^1$  de  $U$  dans  $F$ .


**Théorème 10.** Soit  $U$  ouvert de  $E$  et  $f : U \rightarrow F$ . La fonction  $f$  est de classe  $\mathcal{C}^1$  sur  $U$  si et seulement si ses dérivées partielles dans une base existent en tout point de  $U$  et sont continues sur  $U$ .

*Démonstration.* En annexe. □

**Commentaire :** Établir la seule différentiabilité d'une application n'est pas trivial. Ce théorème fournit un critère simple à vérifier en pratique pour un résultat plus fort que la différentiabilité, d'où sa très grande utilité.

**Corollaire 7.** Les applications constantes et linéaires de  $E$  dans  $F$  sont de classe  $\mathcal{C}^1$ .

*Démonstration.* Immédiat pour les applications constantes dont les dérivées partielles sont nulles. Si  $f \in \mathcal{L}(E, F)$ , on trouve  $\partial_i f(a) = f(e_i)$  pour tout  $a \in E$  et  $i \in \llbracket 1; n \rrbracket$  donc les dérivées partielles sont constantes et par conséquent continues. □

 On admet momentanément que la combinaison linéaire, le produit et la composition de fonctions de  $\mathcal{C}^1$  est encore de classe  $\mathcal{C}^1$  (ces résultats seront établis plus généralement pour des fonctions de classe  $\mathcal{C}^k$  avec  $k \in \mathbb{N} \cup \{\infty\}$ ).

**Corollaire 8.** Les fonctions polynomiales sur  $\mathbb{R}^n$  sont de classe  $\mathcal{C}^1$  et les fonctions rationnelles sur  $\mathbb{R}^n$  sont de classe  $\mathcal{C}^1$  sur leur ensemble de définition (qui est nécessairement un ouvert de  $\mathbb{R}^n \dots$ ).

*Démonstration.* Les applications coordonnées sont linéaires donc de classe  $\mathcal{C}^1$ . Par produit et combinaison linéaire, les fonctions polynomiales sont de classe  $\mathcal{C}^1$  sur  $\mathbb{R}^n$ . Considérons  $R : x \mapsto P/Q(x)$ . La fonction  $R$  est bien définie sur  $Q^{-1}(\mathbb{R}^*)$ , ouvert comme image réciproque d'un ouvert par une application continue. Par composition et produit, on conclut que  $R$  est de classe  $\mathcal{C}^1$  sur  $Q^{-1}(\mathbb{R}^*)$ . □

**Proposition 16.** Soit  $U$  ouvert de  $E$ . On a

$$\mathcal{C}^1(U, F) \subset \mathcal{C}^0(U, F)$$

*Démonstration.* Le caractère  $\mathcal{C}^1$  implique différentiable qui implique continue. □

**Exemples :** 1. Soit  $f$  définie sur  $\mathbb{R}^2$  par

$$\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = \begin{cases} \frac{x^3}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

L'application est de classe  $\mathcal{C}^1$  sur  $\mathbb{R}^2 \setminus \{(0, 0)\}$  (fonction rationnelle) et on a

$$\forall (x, y) \in \mathbb{R}^2 \quad \frac{\partial f}{\partial x}(x, y) = \begin{cases} \frac{3x^2(x^2 + y^2) - 2x^4}{(x^2 + y^2)^2} = \frac{x^2(x^2 + 3y^2)}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 1 & \text{sinon} \end{cases}$$

Or  $\forall y \neq 0 \quad \frac{\partial f}{\partial x}(0, y) = 0 \xrightarrow{y \rightarrow 0} 1 = \frac{\partial f}{\partial x}(0, 0)$

Ainsi

$$\boxed{f \notin \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})}$$

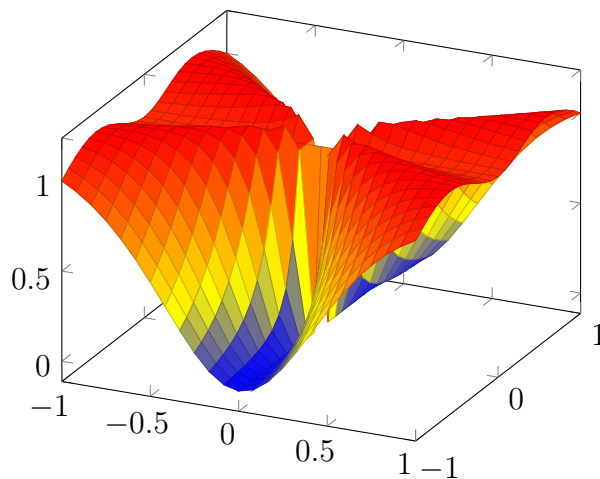


FIGURE 2 – Graphe de  $z = \frac{\partial f}{\partial x}(x, y)$

2. Soit  $f$  définie sur  $\mathbb{R}^2$  par

$$\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = \begin{cases} \frac{x^2 y^2}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

L'application est de classe  $\mathcal{C}^1$  sur  $\mathbb{R}^2 \setminus \{(0, 0)\}$  (fonction rationnelle) et on a

$$\forall (x, y) \in \mathbb{R}^2 \quad \frac{\partial f}{\partial x}(x, y) = \begin{cases} \frac{2xy^4}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

Puis  $\left| \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(0, 0) \right| \leq 2|x| \xrightarrow{(x, y) \rightarrow (0, 0)} 0$

Par un argument de symétrie, on aura la même régularité pour la dérivée partielle en  $y$  et on conclut

$$\boxed{f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})}$$

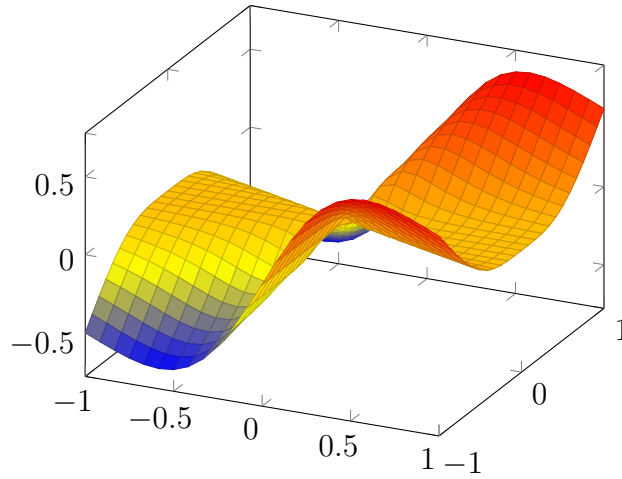


FIGURE 3 – Graphe de  $z = \frac{\partial f}{\partial x}(x, y)$

## 2 Intégration le long d'un chemin

**Théorème 11.** Soit  $U$  ouvert de  $E$ ,  $f \in \mathcal{C}^1(U, F)$  et  $\gamma \in \mathcal{C}^1(I, E)$  avec  $I$  intervalle ouvert contenant  $[0; 1]$  et  $\gamma(I) \subset U$ . Notant  $a = \gamma(0)$  et  $b = \gamma(1)$ , on a

$$f(b) - f(a) = \int_0^1 df(\gamma(t)) \cdot \gamma'(t) dt$$

*Démonstration.* D'après le corollaire 3, la fonction  $f \circ \gamma$  est dérivable sur  $I$  avec  $(f \circ \gamma)'(t) = df(\gamma(t)) \cdot \gamma'(t)$  pour  $t \in I$ . D'après le théorème fondamental d'intégration appliqué à la fonction continue  $(f \circ \gamma)'$ , il vient

$$\int_0^1 (f \circ \gamma)'(t) dt = f(\gamma(1)) - f(\gamma(0)) = f(b) - f(a)$$

□

**Corollaire 9.** Soit  $U$  ouvert de  $E$ ,  $f \in \mathcal{C}^1(U, F)$ . Pour  $[a; b] \subset U$ , on a

$$f(b) - f(a) = \int_0^1 df(a + t(b - a)) \cdot (b - a) dt$$

*Démonstration.* Si  $a = b$ , le résultat est trivial. On suppose  $a \neq b$ . Par ouverture de  $U$ , on dispose de  $\varepsilon > 0$  tel que  $B(a, \varepsilon) \subset U$  et  $B(b, \varepsilon) \subset U$ . Avec  $\delta = \varepsilon/\|b - a\|$  et  $I = ]-\delta; 1 + \delta[$ , on définit  $\gamma : I \rightarrow U, t \mapsto a + t(b - a)$ . Il suffit alors d'appliquer le résultat précédent. □

**Corollaire 10.** Soit  $U$  ouvert connexe par arcs de  $E$  et  $f \in \mathcal{C}^1(U, F)$ . On a

$$f \text{ constante} \iff df = 0 \iff \forall i \in \llbracket 1; n \rrbracket \quad \partial_i f = 0$$

*Démonstration.* Pour la première équivalence, le sens direct est immédiat. Montrons la réciproque dans le cas où  $U$  est convexe (cas général hors-programme). Pour  $(a, b) \in U^2$ , on a  $[a; b] \subset U$  et le résultat découle alors du corollaire précédent. La dernière équivalence peut se voir comme conséquence du théorème 6. □



### 3 Vecteurs tangents à une partie

**Définition 12.** Si  $X$  est une partie de  $E$  et  $x \in X$ , un vecteur  $v \in E$  est dit *tangent à  $X$  en  $x$*  s'il existe  $\varepsilon > 0$  et un arc  $\gamma : ]-\varepsilon; \varepsilon[ \rightarrow X$ , dérivable en 0 tel que  $\gamma(0) = x$  et  $\gamma'(0) = v$ .

**Notation :** On note  $T_x X$  l'ensemble des vecteurs tangents à  $X$  en  $x$  qu'on appelle *espace tangent à  $X$  en  $x$*  (appellation qui ne figure pas dans le programme officiel).

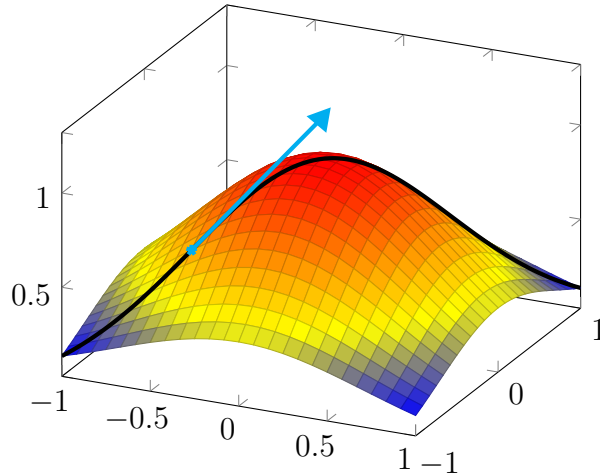


FIGURE 4 – Vecteur tangent au graphe de  $z = e^{-x^2-y^2}$

**Exemples :** 1. Soit  $X = x + F$  avec  $x \in E$  et  $F$  un sev de  $E$ . Soit  $v \in T_x X$ ,  $\varepsilon > 0$  et  $\gamma : ]-\varepsilon; \varepsilon[ \rightarrow X$  un arc associé à  $v$ . Pour  $h \neq 0$ , on a

$$\frac{1}{h} (\gamma(h) - \gamma(0)) = \frac{1}{h} (\gamma(h) - x - (\gamma(0) - x)) \xrightarrow{h \rightarrow 0} \gamma'(0) = v$$

et est à valeurs dans  $F$  fermé d'où  $v \in F$ . Réciproquement, pour  $v \in F$ , l'arc  $\gamma : t \mapsto x + tv$  est dérivable en 0 avec  $\gamma(0) = x$  et  $\gamma'(0) = v$  d'où  $T_x X = F$ .

2. Soit  $E$  euclidien,  $a \in E$ ,  $r > 0$ ,  $X = S(a, r)$  et  $x \in X$ . Soit  $v \in T_x X$ ,  $\varepsilon > 0$  et  $\gamma : ]-\varepsilon; \varepsilon[ \rightarrow X$  un arc associé à  $v$ . On a  $\langle \gamma(t) - a, \gamma(t) - a \rangle = r^2$  pour tout  $t \in ]-\varepsilon; \varepsilon[$  et par dérivation en 0, il vient

$$\langle \gamma'(0), \gamma(0) - a \rangle = \langle v, x - a \rangle = 0$$

ce qui prouve  $T_x X \subset \text{Vect}(x - a)^\perp$ . On peut montrer l'inclusion réciproque avec quelques efforts. Soit  $v \in \text{Vect}(x - a)^\perp$  non nul (le cas  $v = 0_E$  est trivial). On note  $e_1 = \frac{x - a}{\|x - a\|}$  et  $e_2 = \frac{v}{\|v\|}$  qui forme une famille orthonormée de  $E$ . On pose

$$\forall t \in \mathbb{R} \quad \gamma(t) = a + r \left( \cos \left( \frac{\|v\|}{r} t \right) e_1 + \sin \left( \frac{\|v\|}{r} t \right) e_2 \right)$$

et on vérifie que  $\gamma$  est à valeurs dans  $S(a, r)$  avec  $\gamma(0) = x$  et  $\gamma'(0) = v$ .

3. On munit l'espace  $\mathbb{R}^3$  de sa structure euclidienne canonique pour la définition et la proposition qui suivent.

**Définition 13.** Soit  $U$  ouvert de  $\mathbb{R}^2$  et  $f : U \rightarrow \mathbb{R}$ . On appelle *graphe de  $f$*  l'ensemble défini par

$$\{(x, y, z) \in U \times \mathbb{R} \mid z = f(x, y)\} \quad \text{ou} \quad \{(x, y, f(x, y)), (x, y) \in U\}$$

**Proposition 17.** Soit  $U$  ouvert de  $\mathbb{R}^2$ ,  $f : U \rightarrow \mathbb{R}$  différentiable en  $M_0 = (x_0, y_0) \in U$ . On note  $A_0 = (x_0, y_0, f(M_0))$ ,  $\vec{n}_0 = \left( \frac{\partial f}{\partial x}(M_0), \frac{\partial f}{\partial y}(M_0), -1 \right)$ . Notant  $X$  le graphe de  $f$ , on a

$$T_{A_0}X \subset \text{Vect}(\vec{n}_0)^\perp$$

*Démonstration.* Soit  $v$  un vecteur tangent au graphe en  $A_0$ . Il existe alors  $\varepsilon > 0$  et  $\gamma : ]-\varepsilon; \varepsilon[ \rightarrow X$  associé à  $v$ . On a  $\gamma(t) = (x(t), y(t), f(x(t), y(t)))$  pour  $t \in ]-\varepsilon; \varepsilon[$  avec  $x, y$  dérivables en 0 et  $f$  différentiable en  $M_0$  avec  $(x, y)(0) = M_0$ . Par dérivation (avec la règle de la chaîne), il vient

$$\gamma'(0) = \left( x'(0), y'(0), x'(0) \frac{\partial f}{\partial x}(M_0) + y'(0) \frac{\partial f}{\partial y}(M_0) \right)$$

et clairement

$$\gamma'(0) \in \text{Vect} \left( \frac{\partial f}{\partial x}(M_0), \frac{\partial f}{\partial y}(M_0), -1 \right)^\perp$$

□

**Remarques :** (1) Le plan affine  $A_0 + \text{Vect}(\vec{n}_0)^\perp$  est appelé *plan tangent* à la surface d'équation  $z = f(x, y)$  en  $A_0$  et on a

$$(x, y, z) \in A_0 + \text{Vect}(\vec{n}_0)^\perp \iff z = f(M_0) + \frac{\partial f}{\partial x}(M_0)(x - x_0) + \frac{\partial f}{\partial y}(M_0)(y - y_0)$$

(2) On peut démontrer que l'inclusion est une égalité mais ceci requiert un théorème difficile.

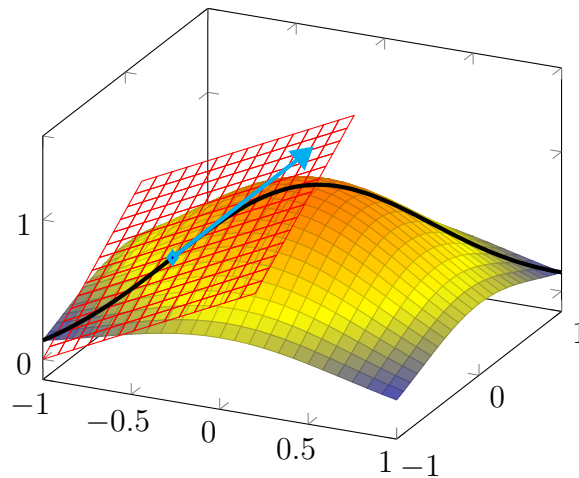


FIGURE 5 – Plan tangent au graphe de  $z = e^{-x^2-y^2}$  et contenant un vecteur tangent

**Théorème 12.** Soit  $U$  ouvert de  $E$ ,  $f \in \mathcal{C}^1(U, \mathbb{R})$ ,  $\lambda$  réel et  $X = f^{-1}(\{\lambda\})$  une ligne de niveau de  $f$ . Pour  $x \in X$  tel que  $df(x) \neq 0_{\mathcal{L}(E, \mathbb{R})}$ , on a

$$T_x X = \text{Ker } df(x)$$

Si l'espace  $E$  est euclidien, on a

$$T_x X = \nabla f(x)^\perp$$

*Démonstration partielle.* Soit  $v \in T_x X$ ,  $\varepsilon > 0$  et  $\gamma : ]-\varepsilon; \varepsilon[ \rightarrow X$  un arc associé à  $v$ . On a

$$\forall t \in ]-\varepsilon; \varepsilon[ \quad f(\gamma(t)) = \lambda$$

La fonction  $f \circ \gamma$  est dérivable en 0 avec

$$(f \circ \gamma)'(0) = df(\gamma(0)) \cdot \gamma'(0) = df(x) \cdot v = 0$$

ce qui prouve l'inclusion  $T_x X \subset \text{Ker } df(x)$ . L'inclusion réciproque requiert un théorème difficile et hors-programme (théorème des fonctions implicites). Si  $E$  est euclidien, on a  $df(x) \cdot h = \langle \nabla f(x), h \rangle$  pour tout  $h \in E$  et le résultat suit.  $\square$

**Remarque :** Pour  $E = \mathbb{R}^n$  muni de sa structure euclidienne canonique, on a

$$v \in \nabla f(x)^\perp \iff \sum_{i=1}^n \partial_i f(x) v_i = 0$$

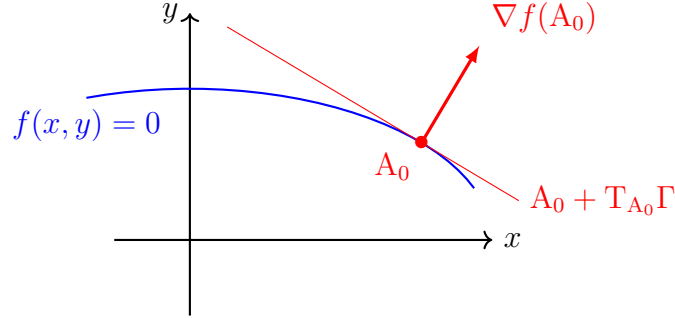


FIGURE 6 – Tangente en  $A_0$  dirigée par  $T_{A_0} \Gamma$  avec  $\Gamma : f(x, y) = 0$

**Exemple :** Soit  $U$  ouvert de  $\mathbb{R}^3$ ,  $f \in \mathcal{C}^1(U, \mathbb{R})$  et  $S$  la surface de  $\mathbb{R}^3$  décrite par l'équation  $f(x, y, z) = 0$ . En un point  $A_0 = (x_0, y_0, z_0) \in U$  dit *régulier*, c'est-à-dire tel que  $\nabla f(A_0) \neq 0$ , on définit le *plan tangent* à  $S$  en  $A_0$  comme le plan affine  $A_0 + T_{A_0} S$ . On a

$$M \in A_0 + T_{A_0} S \iff \langle \nabla f(A_0), \overrightarrow{A_0 M} \rangle = 0$$

En particulier, pour le cas d'un graphe décrit par  $z = f(x, y)$  en  $A_0 = (M_0, f(M_0))$ , on retrouve le vecteur  $\vec{n}_0$  introduit précédemment en considérant  $g(x, y, z) = z - f(x, y)$  avec

$$\nabla g(A_0) = -\vec{n}_0 = \left( -\frac{\partial f}{\partial x}(M_0), -\frac{\partial f}{\partial y}(M_0), 1 \right)$$

## IV Optimisation, premier ordre

### 1 Extremums locaux, globaux

**Définition 14.** Soit  $A$  une partie de  $E$ ,  $a \in A$  et  $f : A \rightarrow \mathbb{R}$ . On dit que  $f$  admet un minimum local, respectivement maximum local, en  $a$  s'il existe un voisinage  $\mathcal{V}$  de  $a$  tel que

$$\forall x \in \mathcal{V} \cap A \quad f(x) \geq f(a)$$

respectivement

$$\forall x \in \mathcal{V} \cap A \quad f(x) \leq f(a)$$

Un extremum local est un maximum ou minimum local.

**Vocabulaire :** On rappelle qu'un voisinage  $\mathcal{V}$  de  $a$  est un ensemble contenant un ouvert contenant  $a$ .

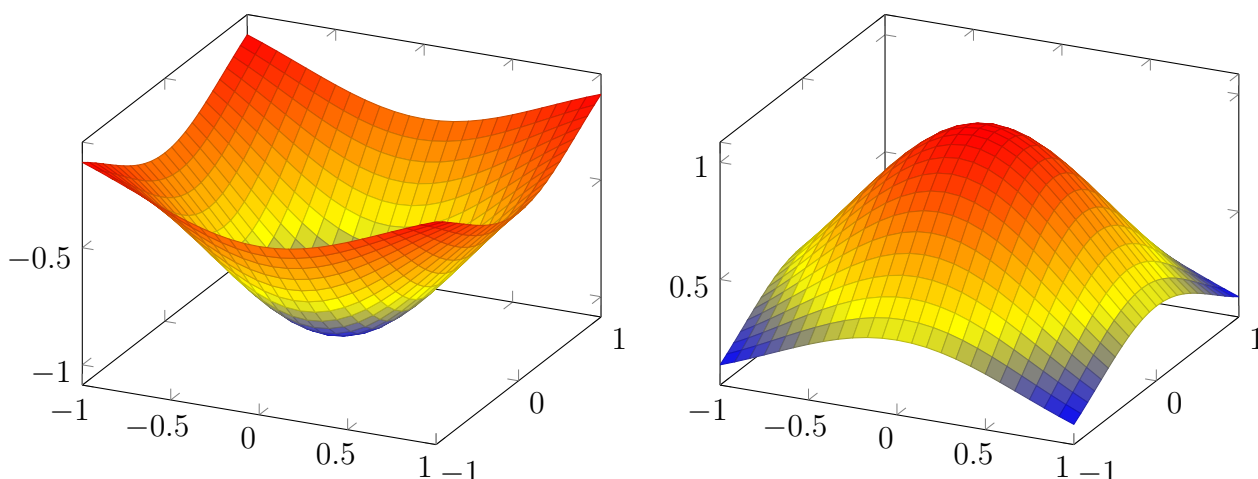


FIGURE 7 – Minimum et maximum local

**Définition 15.** Soit  $A$  une partie de  $E$ ,  $a \in A$  et  $f : A \rightarrow \mathbb{R}$ . On dit que  $f$  admet un minimum global, respectivement maximum global, en  $a$  si

$$\forall x \in A \quad f(x) \geq f(a)$$

respectivement

$$\forall x \in A \quad f(x) \leq f(a)$$

Un extremum global est un maximum ou minimum global.

**Remarque :** Un extremum global est nécessairement local.

**Vocabulaire :** Si une des inégalités précédentes est stricte pour  $x \in \mathcal{V} \cap A \setminus \{a\}$  ou pour  $x \in A \setminus \{a\}$ , on parle d'extremum *strict*.

## 2 Point critique

**Définition 16.** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow \mathbb{R}$  différentiable en  $a$ . On dit que le point  $a$  est point critique de  $f$  si  $df(a) = 0_{\mathcal{L}(E, \mathbb{R})}$ .

**Proposition 18.** On suppose  $E$  muni d'une base quelconque  $\mathcal{B} = (e_1, \dots, e_n)$ . Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow \mathbb{R}$  différentiable en  $a$ . On a

$$df(a) = 0 \iff \forall i \in \llbracket 1; n \rrbracket \quad \partial_i f(a) = 0$$

Si  $E$  est euclidien, on a  $df(a) = 0 \iff \nabla f(a) = 0$

*Démonstration.* Conséquence immédiate des théorèmes 6 et 7. □

**Théorème 13 (Condition nécessaire d'extremum local).** Soit  $U$  ouvert de  $E$ ,  $a \in U$  et  $f : U \rightarrow \mathbb{R}$  différentiable. Si  $f$  admet un extremum local en  $a$ , alors le point  $a$  est point critique.

*Démonstration.* Soit  $a$  un extremum local de  $f$  et  $\mathcal{V}$  un voisinage ouvert de  $a$  tel que l'inégalité est satisfaite sur  $\mathcal{V} \cap U$ . L'ensemble  $\mathcal{V} \cap U$  est un ouvert de  $E$  comme intersection finie d'ouverts et contient  $a$  donc contient une boule ouverte  $B(a, \varepsilon)$  avec  $\varepsilon > 0$ . Soit  $v$  vecteur non nul de  $E$ . Il existe un intervalle  $J = ]-\delta; \delta[$  tel que  $a + tv \in B(a, \varepsilon)$  (on choisit  $\delta = \varepsilon/\|v\|$ ). On pose

$$\forall t \in J \quad g(t) = f(a + tv)$$

D'après le corollaire 3, la fonction  $g = f \circ (t \mapsto a + tv)$  est dérivable en 0. Or, la fonction  $g$  admet un extremum local en 0 point intérieur de  $J$  donc  $g'(0) = 0$  et on a

$$g'(0) = D_v f(a) = df(a) \cdot v$$

Ainsi, la différentielle s'annule en tout vecteur d'où  $df(a) = 0$ . □

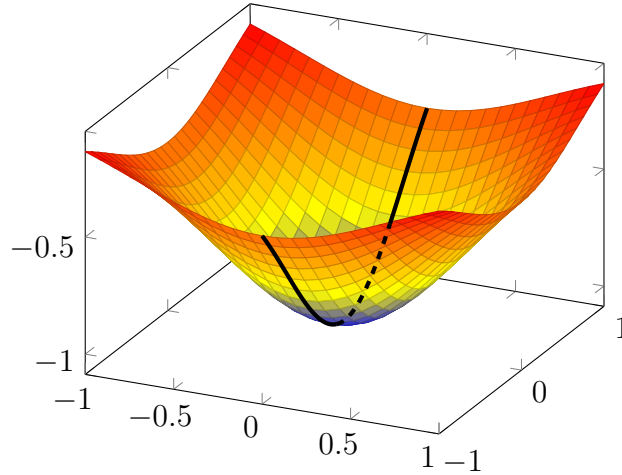


FIGURE 8 – Extremum local suivant une courbe coordonnée

**⚠ Avertissements :** Il est indispensable de travailler sur un ouvert. Ce résultat est important puisqu'il permet de localiser les potentiels extremums locaux. Enfin, la réciproque est fausse.

**Contre-exemples :** 1. Pour une fonction d'une variable réelle, on peut considérer  $t \mapsto t^3$  par exemple.

2. Sur  $\mathbb{R}^2$ , on pose  $f : (x, y) \mapsto xy$ . Le point  $(0, 0)$  est point critique de  $f$  mais n'est pas extremum local de  $f$ . On dit qu'il s'agit d'un *point col* ou *point selle* (en référence à un col de montagne ou une selle de cheval).

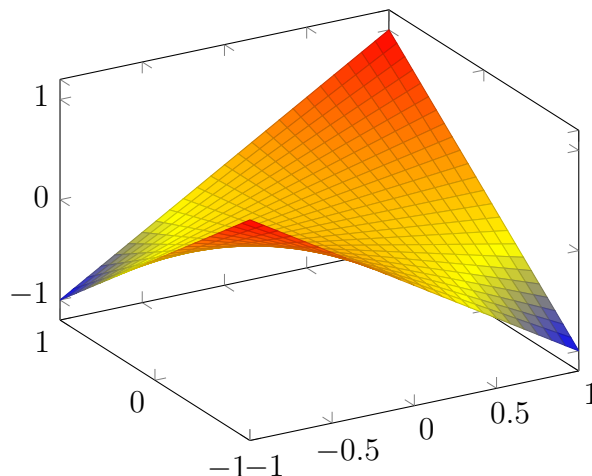


FIGURE 9 – Point col ou point selle

**Exemples :** 1. Soit  $f : (x, y) \mapsto 3x^2 + 3xy + y^2$ . On a

$$\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = \left(\frac{3}{2}x + y\right)^2 + \frac{3}{4}x^2 \geq 0 = f(0, 0)$$

2. Soit  $f : (x, y) \mapsto x^2 + y^2 + x^3$ . On a  $(0, 0)$  point critique puis

$$f(x, y) - f(0, 0) = (x^2 + y^2)(1 + o(1)) \geq 0 \text{ localement et } f(x, 0) \xrightarrow{x \rightarrow -\infty} -\infty$$

Le point  $(0, 0)$  est un minimum local non global.

3. Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto xy(1 - x - y)$  et  $A = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, x + y \leq 1\}$ .

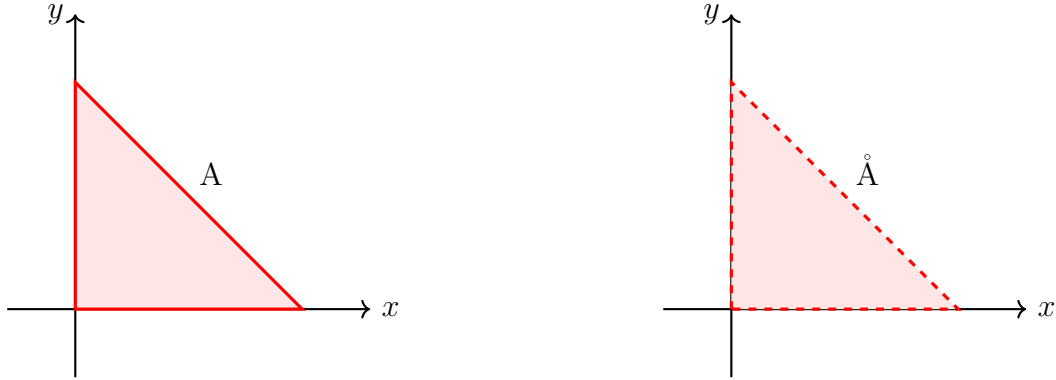


FIGURE 10 – Domaines  $A$  et  $\overset{\circ}{A}$

La fonction  $f$  est continue car polynomiale. L'ensemble  $A$  est un fermé borné de  $\mathbb{R}^2$ , espace de dimension finie et par conséquent  $A$  est compact. D'après le théorème des bornes atteintes, la fonction  $f$  admet un minimum et un maximum sur  $A$ . Les extremums de  $f$  sont atteints soit sur  $\partial A$ , soit dans  $\overset{\circ}{A}$ . On trouve en munissant par exemple  $\mathbb{R}^2$  de  $\|\cdot\|_\infty$  l'intérieur  $\overset{\circ}{A} = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0, x + y < 1\}$  et on voit que  $f(x, y) > 0$  pour  $(x, y) \in \overset{\circ}{A}$ . On observe que  $f(x, y) = 0$  pour  $(x, y) \in \partial A$  puisque

$$\forall t \in [0; 1] \quad f(0, t) = f(t, 0) = f(t, 1 - t) = 0$$

On en déduit que la fonction  $f$  atteint son minimum sur  $A$  en tout point de  $\partial A$  et son maximum sur  $A$  dans  $\overset{\circ}{A}$ . La fonction  $f$  est de classe  $\mathcal{C}^1$  sur  $\overset{\circ}{A}$  car polynomiale et elle atteint son maximum sur  $A$  dans l'ouvert  $\overset{\circ}{A}$  donc en un point critique. On trouve pour  $(x, y) \in \overset{\circ}{A}$

$$\nabla f(x, y) = (0, 0) \iff \begin{cases} y(1 - y - 2x) = 0 \\ x(1 - x - 2y) = 0 \end{cases} \iff (x, y) = \left(\frac{1}{3}, \frac{1}{3}\right)$$

### 3 Optimisation sous contrainte

**Proposition 19.** Soit  $U$  ouvert de  $E$ ,  $X \subset U$ ,  $f : U \rightarrow \mathbb{R}$ . Si la restriction  $f|_X$  admet un extremum local en  $x \in X$  avec  $f$  différentiable en  $x$ , alors

$$T_x X \subset \text{Ker } df(x)$$

*Démonstration.* Soit  $v \in T_x X$ ,  $\varepsilon > 0$  et  $\gamma : ]-\varepsilon; \varepsilon[ \rightarrow X$  un arc associé. La fonction  $f|_X$  admet un extremum local en  $x$  et par conséquent, la fonction  $f \circ \gamma$  admet un extremum local en 0, point intérieur à  $]-\varepsilon; \varepsilon[$ . D'après le corollaire 3, la fonction  $f \circ \gamma$  est dérivable en 0 et on a  $(f \circ \gamma)'(0) = 0$ , autrement dit

$$df(\gamma(0)) \cdot \gamma'(0) = df(x) \cdot v = 0$$

ce qui prouve le résultat attendu. □

**Proposition 20.** Soit  $E$  un  $\mathbb{K}$ -ev et  $\varphi, \psi$  des formes linéaires sur  $E$ . On a

$$\text{Ker } \varphi \subset \text{Ker } \psi \iff \psi \in \text{Vect}(\varphi)$$

*Démonstration.* Le sens indirect est immédiat. Supposons  $\text{Ker } \varphi \subset \text{Ker } \psi$  et aussi  $\varphi \neq 0_{\mathcal{L}(E, \mathbb{R})}$  (sinon, c'est immédiat). Soit  $x_0 \in E \setminus \text{Ker } \varphi$ . On a  $E = \text{Ker } \varphi \oplus \text{Vect}(x_0)$  puis, pour  $x \in E$ , on dispose d'un unique couple  $(u, \lambda) \in \text{Ker } \varphi \times \mathbb{K}$  tel que  $x = u + \lambda x_0$ . Il vient

$$\psi(x) = \psi(u) + \lambda\psi(x_0) = \lambda\psi(x_0)$$

Or 
$$\varphi(x) = \varphi(u) + \lambda\varphi(x_0) = \lambda\varphi(x_0) \quad \text{et} \quad \varphi(x_0) \neq 0$$

ce qui prouve 
$$\psi(x) = \frac{\psi(x_0)}{\varphi(x_0)}\varphi(x)$$

et le résultat suit.  $\square$

**Théorème 14 (Optimisation sous contrainte).** Soit  $U$  ouvert de  $E$  et  $f, g : U \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$  et  $X = g^{-1}(\{0\})$ . Pour  $x \in X$  tel que  $dg(x) \neq 0_{\mathcal{L}(E, \mathbb{R})}$ , si la restriction  $f|_X$  admet un extremum local en  $x$ , alors  $df(x)$  est colinéaire à  $dg(x)$ .

*Démonstration.* D'après le théorème 12, on a  $T_x X = \text{Ker } dg(x)$ . Puis, d'après la proposition 19, on a  $T_x X \subset \text{Ker } df(x)$ . Le résultat suit d'après la proposition précédente.  $\square$

**Remarque :** Si l'espace  $E$  est euclidien, la condition  $dg(x) \neq 0$  peut s'écrire  $\nabla g(x) \neq 0$  et on conclut  $\nabla f(x)$  colinéaire à  $\nabla g(x)$ .

**Exemple :** Soit  $s > 0$ . On note  $U = ]0; +\infty[^n$  et

$$\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n \quad f(x) = \prod_{i=1}^n x_i \quad \text{et} \quad g(x) = \sum_{i=1}^n x_i - s$$

Sur le domaine  $K = [0; +\infty[^n \cap g^{-1}(\{0\})$  qui est un fermé borné donc un compact de l'espace de dimension finie  $\mathbb{R}^n$ , la fonction continue  $f$  y atteint son maximum et celui prend une valeur strictement positive puisque  $f(x) > 0$  pour tout  $x \in U \cap g^{-1}(\{0\})$ . Ce maximum est donc atteint dans l'ouvert  $U$ . On peut alors considérer les restrictions de  $f$  et  $g$  à  $U$  (qu'on notera simplement  $f$  et  $g$ ) et appliquer le théorème précédent puisque les fonctions polynomiales  $f$  et  $g$  sont de classe  $\mathcal{C}^1$  sur  $U$  avec  $\nabla g(x) \neq 0$  pour  $x \in U$ . On note désormais  $X = g^{-1}(\{0\})$  et  $x \in X$  un point en lequel la restriction  $f|_X$  admet un maximum global donc local. Le point  $x$  solution est tel que

$$\nabla f(x) = \lambda \nabla g(x) \quad \text{et} \quad g(x) = 0$$

avec  $\lambda$  réel. On trouve  $x_1 = \dots = x_n = \frac{s}{n}$  et par conséquent

$$\forall u = (u_1, \dots, u_n) \in U \cap g^{-1}(\{0\}) \quad f(u) \leq f(x) = \left(\frac{s}{n}\right)^n = \left(\frac{1}{n} \sum_{i=1}^n u_i\right)^n$$

On retrouve alors l'inégalité arithmético-géométrique.

On peut aussi obtenir les conditions du théorème d'optimisation sous contrainte en considérant le *lagrangien* :

$$\forall (x, \lambda) \in U \times \mathbb{R} \quad L(x, \lambda) = f(x) - \lambda g(x)$$

Un point critique du lagrangien vérifie les conditions d'un extremum sous contrainte puisque pour  $(x, \lambda) \in U \times \mathbb{R}$ , on a

$$dL(x, \lambda) = 0 \iff df(x) = \lambda dg(x) \quad \text{et} \quad g(x) = 0$$

On peut montrer qu'un extremum *libre* pour le Lagrangien est un *extremum lié* pour  $f$ , i.e. un extremum sous la contrainte  $g(x) = 0$ . En effet, soit  $(a, \alpha) \in U \times \mathbb{R}$  un extremum libre du lagrangien, par exemple un minimum local de  $L$ . En particulier, le point  $(a, \alpha)$  est point critique de  $L$  d'où  $g(a) = 0$ . Alors, pour  $(x, \lambda)$  dans un voisinage de  $(a, \alpha)$  avec  $g(x) = 0$ , il vient

$$f(x) = L(x, \lambda) \geq L(a, \alpha) = f(a)$$

## V Fonctions de classe $\mathcal{C}^k$

Dans ce qui suit, on suppose que  $\mathcal{B} = (e_1, \dots, e_n)$  est une base de  $E$  et  $\mathcal{L} = (\varepsilon_1, \dots, \varepsilon_m)$  une base de  $F$ .

### 1 Définitions, propriétés

**Définition 17.** Soit  $U$  ouvert de  $E$ ,  $f : U \rightarrow F$ ,  $a \in U$ ,  $k$  entier non nul et  $(i_1, \dots, i_k) \in \llbracket 1; n \rrbracket^k$ . On dit que  $f$  admet une dérivée partielle  $k$ -ème ou d'ordre  $k$  en  $a$  par rapport aux places  $i_1, \dots, i_k$  (ou  $x_{i_1}, \dots, x_{i_k}$ ) successivement si :

- $\partial_{i_1} f, \partial_{i_2}(\partial_{i_1} f), \dots, \partial_{i_{k-1}}(\dots(\partial_{i_1} f)\dots)$  existent sur un voisinage de  $a$  ;
- $\partial_{i_k}(\partial_{i_{k-1}}(\dots(\partial_{i_1} f)\dots))(a)$  existe.

Dans ce cas, l'élément  $\partial_{i_k}(\partial_{i_{k-1}}(\dots(\partial_{i_1} f)\dots))(a)$  est appelé dérivée partielle  $k$ -ème ou d'ordre  $k$  de  $f$  en  $a$  par rapport aux places  $i_1, \dots, i_k$  (ou  $x_{i_1}, \dots, x_{i_k}$ ) successivement et noté

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

Si cette  $k$ -ème dérivée partielle existe en tout point de  $U$ , on définit l'application dérivée partielle  $k$ -ème par rapport aux places  $i_1, \dots, i_k$  (ou  $x_{i_1}, \dots, x_{i_k}$ ) successivement notée  $\partial_{i_k} \dots \partial_{i_1} f$  ou

$$\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} \text{ par}$$

$$\partial_{i_k} \dots \partial_{i_1} f : \begin{cases} U \longrightarrow F \\ a \longmapsto \partial_{i_k} \dots \partial_{i_1} f(a) \end{cases}$$

**Définition 18.** Soit  $U$  ouvert de  $E$  et  $k$  entier non nul. Une application  $f : U \rightarrow F$  est dite de classe  $\mathcal{C}^k$  sur l'ouvert  $U$  si ses dérivées partielles d'ordre  $k$  existent et sont continues sur  $U$ . Une application est dite de classe  $\mathcal{C}^\infty$  si elle est de classe  $\mathcal{C}^k$  pour tout  $k$  entier.

**Notations :** Pour  $k \in \mathbb{N} \cup \{\infty\}$ , on note  $\mathcal{C}^k(U, F)$  l'ensemble des fonctions de classe  $\mathcal{C}^k$  sur  $U$ . On a les relations

$$\forall k \in \mathbb{N} \quad \mathcal{C}^{k+1}(U, F) \subset \mathcal{C}^k(U, F) \quad \mathcal{C}^\infty(U, F) = \bigcap_{k=0}^{+\infty} \mathcal{C}^k(U, F)$$

**Remarques :** (1) Pour  $k = 1$ , on ne retrouve pas la définition du caractère  $\mathcal{C}^1$  mais la caractérisation fournie par le théorème 11 ce qui assure une cohérence de la définition.  
(2) Pour  $n = 1$ , la définition coïncide avec celle d'une fonction d'une variable réelle.

**Proposition 21.** Les applications constantes de  $E$  dans  $F$  sont de classe  $\mathcal{C}^\infty$  sur  $E$ .

*Démonstration.* Les applications constantes sont différentiables de différentielles nulles donc de dérivées partielles à tout ordre nulles d'où le résultat.  $\square$



**Proposition 22.** Les applications linéaires de  $E$  dans  $F$  sont de classe  $\mathcal{C}^\infty$ .

*Démonstration.* Soit  $f \in \mathcal{L}(E, F)$ . On a  $f$  différentiable avec ses dérivées partielles constantes  $\partial_i f = f(e_i)$  pour tout  $i \in \llbracket 1; n \rrbracket$ . On conclut d'après le résultat précédent.  $\square$

**Théorème 15.** Soit  $k \in \mathbb{N} \cup \{\infty\}$ ,  $U$  ouvert de  $E$  et  $f : U \rightarrow F$  avec  $f = \sum_{j=1}^m f_j \varepsilon_j$ . On a

$$f \in \mathcal{C}^k(U, F) \iff \forall j \in \llbracket 1; m \rrbracket \quad f_j \in \mathcal{C}^k(U, \mathbb{R})$$

*Démonstration.* Par propriétés sur les fonctions vectorielles.  $\square$

## 2 Opérations

Dans ce qui suit, on a  $k \in \mathbb{N} \cup \{\infty\}$ .

**Théorème 16.** Soit  $U$  ouvert de  $E$ ,  $f, g : U \rightarrow F$  et  $\lambda$  réel. Si  $f$  et  $g$  sont de classe  $\mathcal{C}^k$ , alors  $f + \lambda g$  l'est aussi.

*Démonstration.* On procède par récurrence sur  $k$ . Le cas  $k = 0$  est vraie. Supposons le résultat vrai pour  $k$  entier. Soient  $f, g$  de classe  $\mathcal{C}^{k+1}$ . D'après la proposition 13, l'application  $f + \lambda g$  admet des dérivées partielles avec

$$\forall i \in \llbracket 1; n \rrbracket \quad \partial_i(f + \lambda g) = \partial_i f + \lambda \partial_i g$$

Les fonctions  $\partial_i f$  et  $\partial_i g$  sont de classe  $\mathcal{C}^k$  d'où  $\partial_i(f + \lambda g)$  également par hypothèse de récurrence. Le résultat suit. Le cas  $k = \infty$  s'en déduit.  $\square$

**Corollaire 11.** Soit  $U$  ouvert de  $E$ . L'ensemble  $\mathcal{C}^k(U, F)$  est un  $\mathbb{R}$ -ev.

*Démonstration.* L'ensemble  $\mathcal{C}^k(U, F)$  contient la fonction nulle et est stable par combinaison linéaire donc est un sev de  $\mathcal{F}(U, F)$ .  $\square$

**Théorème 17.** Soit  $U$  ouvert de  $E$ ,  $f : U \rightarrow F$ ,  $g : U \rightarrow G$  et  $B : F \times G \rightarrow H$  bilinéaire. Si  $f$  et  $g$  sont de classe  $\mathcal{C}^k$ , alors  $B(f, g)$  l'est également.

*Démonstration.* On procède par récurrence sur  $k$ . Le cas  $k = 0$  est vraie car  $B$  est continue. Supposons le résultat vrai pour  $k$  entier. Soient  $f, g$  de classe  $\mathcal{C}^{k+1}$ . D'après la proposition 15, l'application  $B(f, g)$  admet des dérivées partielles avec

$$\forall i \in \llbracket 1; n \rrbracket \quad \partial_i B(f, g) = B(\partial_i f, g) + B(f, \partial_i g)$$

Il suffit ensuite d'appliquer l'hypothèse de récurrence et le théorème 16. Le cas  $k = \infty$  s'en déduit.  $\square$

**Corollaire 12.** Soit  $U$  ouvert de  $E$ ,  $f : U \rightarrow F$ ,  $g : U \rightarrow \mathbb{R}$ . Si  $f$  et  $g$  sont de classe  $\mathcal{C}^k$ , alors  $gf$  est de classe  $\mathcal{C}^k$ .

*Démonstration.* On considère  $B : \mathbb{R} \times F \rightarrow F$ ,  $(x, y) \mapsto xy$  et on applique le théorème précédent.  $\square$

**Corollaire 13.** Soit  $U$  ouvert de  $E$  et  $F$  une algèbre. L'ensemble  $\mathcal{C}^k(U, F)$  est également une algèbre.

*Démonstration.* On vérifie que  $\mathcal{C}^k(\mathbf{U}, \mathbf{F})$  est une sous-algèbre de  $\mathcal{F}(\mathbf{U}, \mathbf{F})$ . En particulier, on considère  $B : \mathbb{F}^2 \rightarrow \mathbb{F}, (x, y) \mapsto xy$  et on applique le théorème précédent.  $\square$

**Remarque :** On utilise ce résultat typiquement avec  $\mathbf{F} = \mathbb{R}$  ou  $\mathbf{F} = \mathcal{M}_n(\mathbb{R})$ .

**Théorème 18.** Soit  $\mathbf{U}$  ouvert de  $\mathbf{E}$ ,  $f : \mathbf{U} \rightarrow \mathbf{F}$ ,  $\mathbf{V}$  ouvert de  $\mathbf{F}$ ,  $g : \mathbf{V} \rightarrow \mathbf{G}$  avec  $f(\mathbf{U}) \subset \mathbf{V}$ . Si  $f$  et  $g$  sont de classe  $\mathcal{C}^k$ , alors  $g \circ f$  est de classe  $\mathcal{C}^k$ .

*Démonstration.* On procède par récurrence. C'est immédiat pour  $k = 0$ . Supposons le résultat vrai pour  $k$  entier et supposons  $f, g$  de classe  $\mathcal{C}^{k+1}$ . D'après le théorème 8, l'application  $g \circ f$  admet des dérivées partielles avec

$$\forall j \in \llbracket 1; n \rrbracket \quad \partial_j(g \circ f) = \sum_{i=1}^m (\partial_j f_i) \times (\partial_i g) \circ f$$

Le résultat suit par hypothèse de récurrence avec le théorème 16 et le corollaire 12. Le cas  $k = \infty$  s'en déduit.  $\square$

**Théorème 19.** Les fonctions polynomiales sur  $\mathbb{R}^n$  sont de classe  $\mathcal{C}^\infty$  et les fonctions rationnelles sur  $\mathbb{R}^n$  sont de classe  $\mathcal{C}^\infty$  sur leur ensemble de définition.

*Démonstration.* Les applications coordonnées  $x \mapsto x_i$  sont linéaires donc de classe  $\mathcal{C}^\infty$ . Par produit (corollaire 13) et combinaison linéaire (théorème 16), le résultat suit pour les fonctions polynomiales. Par composition (théorème 18) avec la fonction inverse puis produit, le résultat suit pour les fonctions rationnelles.  $\square$

**Remarque :** Le résultat s'étend aux fonctions polynomiales sur  $\mathcal{M}_n(\mathbb{R})$  puisque  $\mathcal{M}_n(\mathbb{R})$  est isomorphe à  $\mathbb{R}^{n^2}$ .

**Exemples :** 1. Le déterminant  $\det : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$  est de classe  $\mathcal{C}^\infty$  car polynomial.

2. Soit  $f$  définie sur  $\mathbb{R}^2$  par

$$\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = \begin{cases} \frac{\cos(x) - \cos(y)}{x - y} & \text{si } x \neq y \\ -\sin(x) & \text{sinon} \end{cases}$$

Par trigonométrie, on a

$$\forall (x, y) \in \mathbb{R}^2 \quad \cos(x) - \cos(y) = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$$

On pose  $\forall t \in \mathbb{R} \quad \varphi(t) = \begin{cases} \frac{\sin(t)}{t} & \text{si } t \neq 0 \\ 1 & \text{sinon} \end{cases}$

La fonction  $\varphi$  est de classe  $\mathcal{C}^\infty$  sur  $\mathbb{R}$  car développable en série entière avec

$$\forall t \in \mathbb{R} \quad \varphi(t) = \sum_{n=0}^{+\infty} \frac{(-1)^n t^{2n}}{(2n+1)!}$$

Or, on a  $\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = -\varphi\left(\frac{x-y}{2}\right) \sin\left(\frac{x+y}{2}\right)$

Par composition

$$f \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R})$$

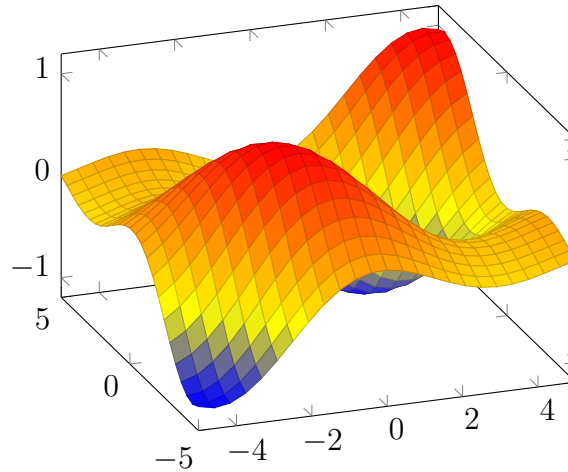


FIGURE 11 – Graphe de  $z = f(x, y)$

### 3 Théorème de Schwarz

**Théorème 20 (Théorème de Schwarz).** Soit  $U$  ouvert de  $E$  et  $f \in \mathcal{C}^2(U, F)$ . On a

$$\forall (i, j) \in \llbracket 1; n \rrbracket^2 \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

ou synthétiquement

$$\partial_i \partial_j f = \partial_j \partial_i f$$

*Démonstration.* En annexe. □

**Corollaire 14.** Soit  $U$  ouvert de  $E$ ,  $k$  entier non nul et  $f \in \mathcal{C}^k(U, F)$ . Alors, pour  $(i_1, \dots, i_k) \in \llbracket 1; n \rrbracket^k$ , on a

$$\forall \sigma \in S_k \quad \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} = \frac{\partial^k f}{\partial x_{i_{\sigma(k)}} \dots \partial x_{i_{\sigma(1)}}}$$

ou synthétiquement

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

autrement dit, les dérivées partielles  $k$ -èmes ne dépendent pas de l'ordre de dérivation.

*Démonstration.* Toute permutation peut s'écrire comme produit de transpositions de type  $(1 \ 2), (2 \ 3), \dots, (k-1 \ k)$ . Le résultat suit. □

## VI Optimisation, deuxième ordre

Dans ce qui suit, l'espace  $\mathbb{R}^n$  est muni de sa structure euclidienne canonique.

### 1 Matrice Hessienne

**Définition 19.** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $a \in U$  et  $f \in \mathcal{C}^2(U, \mathbb{R})$ . On appelle matrice hessienne de  $f$  en  $a$  la matrice notée  $H_f(a)$  définie par

$$H_f(a) = (\partial_i \partial_j f(a))_{1 \leq i, j \leq n} \quad \text{ou} \quad \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{1 \leq i, j \leq n}$$

**Proposition 23.** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $a \in U$  et  $f \in \mathcal{C}^2(U, \mathbb{R})$ . On a  $H_f(a) \in \mathcal{S}_n(\mathbb{R})$ .

*Démonstration.* Conséquence immédiate du théorème de Schwarz.  $\square$

**Théorème 21 (Théorème de Taylor-Young à l'ordre 2).** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $a \in U$  et  $f \in \mathcal{C}^2(U, \mathbb{R})$ . On a, en confondant  $H_f(a)$  et l'application linéaire qui lui est canoniquement associée

$$f(a+h) \underset{h \rightarrow 0_{\mathbb{R}^n}}{=} f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle H_f(a) \cdot h, h \rangle + o(\|h\|^2)$$

qu'on peut aussi écrire, en confondant  $h$  avec un vecteur de  $\mathcal{M}_{n,1}(\mathbb{R})$

$$f(a+h) \underset{h \rightarrow 0_{\mathbb{R}^n}}{=} f(a) + \nabla f(a)^\top h + \frac{1}{2} h^\top H_f(a) h + o(\|h\|^2)$$

*Démonstration.* En annexe.  $\square$

## 2 Étude au deuxième ordre

**Théorème 22.** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $a \in U$  et  $f \in \mathcal{C}^2(U, \mathbb{R})$ . Si la fonction  $f$  admet un minimum local en  $a$ , respectivement maximum local, alors le point  $a$  est point critique de  $f$  et  $H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$ , respectivement  $-H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$ .

*Démonstration.* Quitte à considérer  $-f$ , on peut supposer que le point  $a$  est un minimum local de  $f$ . D'après le théorème 13, le point  $a$  est point critique de  $f$ . La matrice hessienne  $H_f(a)$  symétrique réelle est diagonalisable d'après le théorème spectral. Soit  $\lambda \in \text{Sp}(H_f(a))$  et  $v \in E_\lambda(H_f(a))$  avec  $\|v\| = 1$ . Supposons  $\lambda < 0$ . Considérant  $tv$  avec  $t \rightarrow 0$ , il vient

$$f(a+tv) - f(a) \underset{t \rightarrow 0}{=} \frac{t^2}{2} \langle H_f(a) \cdot v, v \rangle + t^2 \|v\|^2 o(1) \underset{t \rightarrow 0}{=} \frac{t^2}{2} (\lambda + o(1))$$

Comme  $\lambda + o(1) \xrightarrow{t \rightarrow 0} \lambda < 0$ , on peut choisir  $t$  assez proche de zéro pour avoir  $f(a+tv) < f(a)$  ce qui contredit le fait que le point  $a$  soit minimum local. On en déduit  $\lambda \geq 0$  et par caractérisation spectrale des matrices de  $\mathcal{S}_n^+(\mathbb{R})$ , le résultat suit.  $\square$

**Théorème 23.** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $a \in U$  et  $f \in \mathcal{C}^2(U, \mathbb{R})$ . Si le point  $a$  est point critique de  $f$  et si  $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$ , respectivement  $-H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$ , alors le point  $a$  est un minimum local strict, respectivement maximum local strict.

*Démonstration.* Quitte à considérer  $-f$ , on peut supposer  $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$ . On note  $E = \mathbb{R}^n$ .

Pour  $h \in E$ , on a 
$$f(a+h) - f(a) \underset{h \rightarrow 0_E}{=} \frac{1}{2} \langle H_f(a) \cdot h, h \rangle + o(\|h\|^2)$$

Soit  $\mathcal{B} = (e_1, \dots, e_n)$  une base orthonormée de  $E$  constituée de vecteurs propres de  $H_f(a)$  associés aux valeurs propres  $\lambda_1 \leq \dots \leq \lambda_n$ . Par caractérisation spectrale des matrices de  $\mathcal{S}_n^{++}(\mathbb{R})$ , on a  $\text{Sp}(H_f(a)) \subset ]0; +\infty[$  d'où  $\lambda_1 > 0$  puis

$$\forall h = \sum_{i=1}^n h_i e_i \in E \quad \langle H_f(a) \cdot h, h \rangle = \sum_{i=1}^n \lambda_i h_i^2 \geq \lambda_1 \sum_{i=1}^n h_i^2 = \lambda_1 \|h\|^2$$

Ainsi, on obtient 
$$f(a+h) - f(a) \geq \frac{1}{2} \|h\|^2 (\lambda_1 + o(1))$$

Comme  $\lambda_1 + o(1) \xrightarrow{h \rightarrow 0_E} \lambda_1$ , on peut trouver un voisinage  $\mathcal{V}$  de  $0_E$  tel que  $\lambda_1 + o(1) \geq \frac{\lambda_1}{2} > 0$  pour  $h \in \mathcal{V}$  d'où

$$\forall h \in \mathcal{V} \setminus \{0_E\} \quad f(a+h) - f(a) > 0$$

ce qui prouve que le point  $a$  est minimal local strict.  $\square$

**Corollaire 15.** Soit  $U$  ouvert de  $\mathbb{R}^2$ ,  $a \in U$ ,  $f \in \mathcal{C}^2(U, \mathbb{R})$ . On définit les notations de Monge

$$r = \frac{\partial^2 f}{\partial x^2}(a) \quad s = \frac{\partial^2 f}{\partial x \partial y}(a) \quad t = \frac{\partial^2 f}{\partial y^2}(a)$$

On suppose que le point  $a$  est point critique de  $f$ . On a

1. Si  $rt - s^2 > 0$ , le point  $a$  est un extremum local de  $f$  ;  
— Si  $rt - s^2 > 0$  et  $r + t > 0$ , le point  $a$  est un minimum local strict de  $f$  ;  
— Si  $rt - s^2 > 0$  et  $r + t < 0$ , le point  $a$  est un maximum local strict de  $f$  ;
2. Si  $rt - s^2 < 0$ , le point  $a$  n'est pas un extremum local de  $f$  (on dit que  $a$  est point col ou point selle) ;
3. Si  $rt - s^2 = 0$ , on ne peut rien dire.

*Démonstration.* D'après le théorème spectral, la matrice  $H_f(a) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$  est orthogonalement semblable à  $\text{diag}(\lambda_1, \lambda_2)$ . On a

$$\det(H_f(a)) = rt - s^2 = \lambda_1 \lambda_2 \quad \text{et} \quad \text{Tr}(H_f(a)) = r + t = \lambda_1 + \lambda_2$$

Si  $rt - s^2 > 0$ , alors les valeurs propres  $\lambda_1$  et  $\lambda_2$  sont non nulles et de même signe et donc du signe de  $\lambda_1 + \lambda_2$ . Le résultat suit pour ce cas d'après le théorème 23. Si  $rt - s^2 < 0$ , alors les valeurs propres  $\lambda_1, \lambda_2$  sont non nulles et de signe opposé. D'après le théorème 22, le point  $a$  n'est ni un minimum, ni un maximum local. En considérant les applications définies sur  $\mathbb{R}^2$  par  $(x, y) \mapsto x^4 + y^4$ ,  $(x, y) \mapsto -(x^4 + y^4)$  ou  $(x, y) \mapsto x^4 - y^4$  qui admettent toutes  $(0, 0)$  comme unique point critique, on constate que toutes les configurations sont possibles sous la condition  $rt - s^2 = 0$ .

□

**Remarque :** Le critère portant sur  $r + t$  peut être remplacé par un critère portant sur  $r$ . En effet, si  $rt - s^2 > 0$ , alors les valeurs propres  $\lambda_1$  et  $\lambda_2$  sont non nulles de même signe d'où un signe constant pour  $\langle H_f(a) \cdot h, h \rangle$  et en particulier pour  $h = (1, 0)$ . Le résultat suit.

# Annexes

## Caractérisation des fonctions de classe $\mathcal{C}^1$

**Théorème 11.** Soit  $U$  ouvert de  $E$  et  $f : U \rightarrow F$ . La fonction  $f$  est de classe  $\mathcal{C}^1$  sur  $U$  si et seulement si ses dérivées partielles dans une base existent en tout point de  $U$  et sont continues sur  $U$ .

*Démonstration.* Pour  $i \in \llbracket 1; n \rrbracket$ , on pose  $dx_i \in \mathcal{L}(E, \mathbb{R})$  définie par  $dx_i(e_j) = \delta_{i,j}$  pour tout  $(i, j) \in \llbracket 1; n \rrbracket^2$ . La famille  $(dx_i)_{i \in \llbracket 1; n \rrbracket}$  est une base de  $\mathcal{L}(E, \mathbb{R})$  (appelée *base duale*) car libre et génératrice. Par conséquent, la famille  $(\varepsilon_j dx_i)_{(i,j) \in \llbracket 1; n \rrbracket \times \llbracket 1; m \rrbracket}$  est une base de  $\mathcal{L}(E, F)$  car libre et de cardinal égal à  $\dim E \times \dim F$ . Pour  $\varphi \in \mathcal{L}(E, F)$ , on a la décomposition

$$\varphi = \sum_{i=1}^n \varphi(e_i) dx_i = \sum_{(i,j) \in \llbracket 1; n \rrbracket \times \llbracket 1; m \rrbracket} \varphi_j(e_i) dx_i \varepsilon_j$$

Supposons  $f$  différentiable. On a

$$df(a) = \sum_{(i,j) \in \llbracket 1; n \rrbracket \times \llbracket 1; m \rrbracket} (df_j(a) \cdot e_i) dx_i \varepsilon_j = \sum_{(i,j) \in \llbracket 1; n \rrbracket \times \llbracket 1; m \rrbracket} \partial_i f_j(a) dx_i \varepsilon_j$$

d'où

$$df = \sum_{(i,j) \in \llbracket 1; n \rrbracket \times \llbracket 1; m \rrbracket} (\partial_i f_j) dx_i \varepsilon_j = \sum_{i=1}^n \partial_i f dx_i$$

Or, on sait

$$\begin{aligned} df \in \mathcal{C}^0(U, \mathcal{L}(E, F)) &\iff \forall (i, j) \in \llbracket 1; n \rrbracket \times \llbracket 1; m \rrbracket \quad \partial_i f_j \in \mathcal{C}^0(U, \mathbb{R}) \\ &\iff \forall i \in \llbracket 1; n \rrbracket \quad \partial_i f \in \mathcal{C}^0(U, F) \end{aligned}$$

Le sens direct en résulte. Démontrons la réciproque dans le cas où  $\dim E = 2$ . On munit  $E$  de la norme infinie relativement à la base  $\mathcal{B}$ . Soit  $a \in U$ . Comme  $U$  est ouvert, il existe  $\eta > 0$  tel que  $B(a, \eta) \subset U$ . Soit  $h \in E$  tel que  $\|h\|_\infty = \max(|h_1|, |h_2|) < \eta$ . On a

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1 + h_1, a_2) - f(a_1, a_2) \\ &= \int_0^{h_2} \partial_2 f(a_1 + h_1, a_2 + t) dt + \int_0^{h_1} \partial_1 f(a_1 + t, a_2) dt \end{aligned}$$

Par suite

$$\begin{aligned} f(a+h) - f(a) - h_1 \partial_1 f(a) - h_2 \partial_2 f(a) &= \\ &= \int_0^{h_2} [\partial_2 f(a_1 + h_1, a_2 + t) - \partial_2 f(a_1, a_2)] dt + \int_0^{h_1} [\partial_1 f(a_1 + t, a_2) - \partial_1 f(a_1, a_2)] dt \end{aligned}$$

Pour  $\varepsilon > 0$ , comme  $\partial_1 f$  et  $\partial_2 f$  sont continues en  $a$ , il existe  $\delta > 0$  tel que pour  $\|h\|_\infty < \delta$

$$\|\partial_2 f(a_1 + h_1, a_2 + t) - \partial_2 f(a_1, a_2)\| \leq \varepsilon \quad \text{et} \quad \|\partial_1 f(a_1 + t, a_2) - \partial_1 f(a_1, a_2)\| \leq \varepsilon$$

Ainsi

$$\|f(a+h) - f(a) - h_1 \partial_1 f(a) - h_2 \partial_2 f(a)\| \leq \varepsilon(|h_1| + |h_2|) \leq 2\varepsilon \|h\|_\infty$$

D'où

$$f(a+h) - f(a) - h_1 \partial_1 f(a) - h_2 \partial_2 f(a) = o(h)$$

ce qui prouve que l'application  $f$  est différentiable sur  $U$  avec

$$\forall a \in U \quad df(a) \cdot h = h_1 \partial_1 f(a) + h_2 \partial_2 f(a)$$

La continuité des dérivées partielles équivaut à la continuité des dérivées partielles des applications coordonnées qui équivaut à la continuité de  $df$ .  $\square$

## Caractérisation des fonctions constantes

**Corollaire 10.** Soit  $U$  ouvert connexe par arcs de  $E$  et  $f \in \mathcal{C}^1(U, F)$ . On a

$$f \text{ constante} \iff df = 0 \iff \forall i \in \llbracket 1; n \rrbracket \quad \partial_i f = 0$$

Pour la réciproque de la première équivalence dans le cas général :

*Démonstration.* Soit  $(a, b) \in U^2$ . Par connexité par arcs, il existe  $\varphi \in \mathcal{C}^0([0; 1], U)$  tel que  $\varphi(0) = a$ ,  $\varphi(1) = b$ . L'ensemble  $\varphi([0; 1])$  est compact comme image directe d'un compact par une application continue. L'application  $d(\cdot, E \setminus U)$  est continue car 1-lipschitzienne. Par conséquent, elle atteint son minimum sur le compact  $\varphi([0; 1])$  en un point  $\varphi(t_0)$  hors du fermé  $E \setminus U$  avec  $t_0 \in [0; 1]$  (il s'agit en fait de  $d(\varphi([0; 1]), E \setminus U)$ ). Ainsi

$$\forall t \in [0; 1] \quad d(\varphi(t), E \setminus U) \geq \varepsilon = d(\varphi(t_0), E \setminus U) > 0$$

Ainsi, pour  $t \in [0; 1]$  et  $x \in E \setminus U$ , on a

$$d(x, \varphi(t)) \geq d(\varphi(t), E \setminus U) \geq d(\varphi(t_0), E \setminus U) \geq \varepsilon$$

d'où par contraposée  $x \in B(\varphi(t), \varepsilon) \implies x \in U$

D'après le théorème de Heine, l'application  $\varphi$  est uniformément continue sur le compact  $[0; 1]$ .

On dispose alors de  $\eta > 0$  tel que

$$\forall (s, t) \in [0; 1]^2 \quad |s - t| < \eta \implies \|\varphi(s) - \varphi(t)\| < \varepsilon$$

On choisit une subdivision  $(t_i)_{0 \leq i \leq n}$  assez fine pour avoir  $|t_{i+1} - t_i| < \eta$  pour tout  $i \in \llbracket 0; n-1 \rrbracket$ . Par construction, on a

$$\forall i \in \llbracket 0; n-1 \rrbracket \quad [\varphi(t_i); \varphi(t_{i+1})] \subset U$$

En effet, soit  $i \in \llbracket 0; n-1 \rrbracket$  et  $\lambda \in [0; 1]$ . On a

$$\|\lambda\varphi(t_i) + (1 - \lambda)\varphi(t_{i+1}) - \varphi(t_i)\| = (1 - \lambda)\|\varphi(t_{i+1}) - \varphi(t_i)\| < \varepsilon$$

d'où  $\lambda\varphi(t_i) + (1 - \lambda)\varphi(t_{i+1}) \in B(\varphi(t_i), \varepsilon) \subset U$

D'après le corollaire précédent, on

$$\forall i \in \llbracket 0; n-1 \rrbracket \quad f(a_i) = f(a_{i+1})$$

ce qui prouve que la suite  $(f(a_i))_{0 \leq i \leq n}$  est constante et par conséquent

$$f(a) = f(a_0) = f(a_n) = f(b)$$

d'où le résultat. □

## Théorème de Schwarz

**Théorème 18 (Théorème de Schwarz).** Soit  $U$  ouvert de  $E$  et  $f \in \mathcal{C}^2(U, F)$ . On a

$$\forall (i, j) \in \llbracket 1; n \rrbracket^2 \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

*Démonstration.* Seuls les cas avec  $i \neq j$  méritent l'attention. On suppose  $(i, j) = (1, 2)$  pour alléger la rédaction. Comme on va figer toutes les coordonnées sauf les deux premières, on note abusivement  $f(x) = f(x_1, x_2)$  pour  $x \in U$  au lieu de  $f(x_1, \dots, x_n)$ . Soit  $a \in U$ . Il existe un voisinage  $\mathcal{V}$  de 0 dans  $\mathbb{R}^2$  tel que  $a + (h_1, h_2, 0, \dots, 0) \in U$  pour  $h = (h_1, h_2) \in \mathcal{V}$ . On pose

$$\forall (h_1, h_2) \in \mathcal{V} \quad \Delta(h_1, h_2) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1, a_2)$$

Fixons  $h \in \mathcal{V}$ . D'après le théorème des accroissements finis appliqué à

$$\varphi : t \mapsto u(t, h_2) = f(a_1 + t, a_2 + h_2) - f(a_1 + t, a_2)$$

il existe  $c_1$  entre 0 et  $h_1$  tel que

$$\Delta(h) = \varphi(h_1) - \varphi(0) = h_1 \varphi'(c_1) = h_1 (\partial_1 f(a_1 + c_1, a_2 + h_2) - \partial_1 f(a_1 + c_1, a_2))$$

Puis, avec le théorème des accroissements finis appliqué à  $\psi : t \mapsto \partial_1 f(a_1 + c_1, a_2 + t)$ , il existe  $c_2$  entre 0 et  $h_2$  tel que

$$\Delta(h) = h_1 [\psi(h_2) - \psi(0)] = h_1 h_2 \psi'(c_2) = h_1 h_2 \partial_2 \partial_1 f(a_1 + c_1, a_2 + c_2)$$

Par continuité de  $\partial_2 \partial_1 f$ , il vient

$$\frac{\Delta(h)}{h_1 h_2} \xrightarrow{h \rightarrow (0,0)} \partial_2 \partial_1 f(a)$$

En permutant les deux et troisièmes termes dans l'écriture de  $\Delta(h)$  et en suivant une démarche identique à précédemment, on obtient

$$\frac{\Delta(h)}{h_1 h_2} \xrightarrow{h \rightarrow (0,0)} \partial_1 \partial_2 f(a)$$

Le résultat suit. □

## Le théorème de Taylor-Young à l'ordre deux

**Théorème 20 (Théorème de Taylor-Young à l'ordre 2).** Soit  $U$  ouvert de  $\mathbb{R}^n$ ,  $a \in U$  et  $f \in \mathcal{C}^2(U, \mathbb{R})$ . On a, en confondant  $H_f(a)$  et l'application linéaire qui lui est canoniquement associée

$$f(a + h) \underset{h \rightarrow 0_{\mathbb{R}^n}}{=} f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle H_f(a) \cdot h, h \rangle + o(\|h\|^2)$$

qu'on peut aussi écrire, en confondant  $h$  avec un vecteur de  $\mathcal{M}_{n,1}(\mathbb{R})$

$$f(a + h) \underset{h \rightarrow 0_{\mathbb{R}^n}}{=} f(a) + \nabla f(a)^\top h + \frac{1}{2} h^\top H_f(a) h + o(\|h\|^2)$$

*Démonstration.* Soit  $h \in E$  et  $I$  un intervalle ouvert contenant  $[0; 1]$  tel  $a + th \in U$  pour  $t \in [0; 1]$ . On pose

$$\forall t \in I \quad \varphi(t) = f(a + th)$$

La fonction  $\varphi = f \circ (t \mapsto a + th)$  est de classe  $\mathcal{C}^2$  sur  $I$  par composition. Par Taylor reste intégral et linéarité de l'intégrale sur le segment  $[0; 1]$  :

$$\begin{aligned} \varphi(1) &= \varphi(0) + \varphi'(0) + \int_0^1 \varphi''(t)(1-t) dt \\ &= \varphi(0) + \varphi'(0) + \int_0^1 \varphi''(0)(1-t) dt + \int_0^1 [\varphi''(t) - \varphi''(0)](1-t) dt \end{aligned}$$

Par dérivation, on trouve pour  $t \in I$

$$\varphi'(t) = df(a + th) \cdot h = \sum_{j=1}^n h_j \partial_j f(a + th)$$

puis

$$\begin{aligned} \varphi''(t) &= \sum_{j=1}^n h_j \frac{d}{dt} [\partial_j f \circ (t \mapsto a + th)] \\ &= \sum_{j=1}^n h_j d(\partial_j f)(a + th) \cdot h = \sum_{j=1}^n h_j \left( \sum_{i=1}^n h_i \partial_i \partial_j f(a + th) \right) \end{aligned}$$



autrement dit 
$$\varphi''(t) = \sum_{1 \leq i, j \leq n} h_i h_j \partial_i \partial_j f(a + th) = \langle H_f(a + th) \cdot h, h \rangle$$

Par linéarité de l'intégrale sur le segment  $[0; 1]$ , il vient

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2} \varphi''(0) + \sum_{1 \leq i, j \leq n} h_i h_j \int_0^1 [\partial_i \partial_j f(a + th) - \partial_i \partial_j f(a)] (1 - t) dt$$

Soit  $(i, j) \in \llbracket 1; n \rrbracket^2$ . La fonction  $\partial_i \partial_j f$  est continue sur  $U$ . Ainsi, pour  $\varepsilon > 0$ , il existe  $\delta > 0$  tel que, pour  $\|h\| \leq \delta$ , on a

$$\forall t \in [0; 1] \quad \|\partial_i \partial_j f(a + th) - \partial_i \partial_j f(a)\| \leq \varepsilon$$

Par inégalité triangulaire, on obtient pour  $\|h\| \leq \delta$

$$\left| \sum_{1 \leq i, j \leq n} h_i h_j \int_0^1 [\partial_i \partial_j f(a + th) - \partial_i \partial_j f(a)] (1 - t) dt \right| \leq \sum_{1 \leq i, j \leq n} |h_i| |h_j| \int_0^1 \varepsilon (1 - t) dt \leq \varepsilon \|h\|_1^2$$

Les normes sur  $\mathbb{R}^n$  étant équivalentes, on a donc prouvé

$$\sum_{1 \leq i, j \leq n} h_i h_j \int_0^1 [\partial_i \partial_j f(a + th) - \partial_i \partial_j f(a)] (1 - t) dt \underset{h \rightarrow 0_{\mathbb{R}^n}}{=} o(\|h\|^2)$$

Ainsi 
$$f(a + h) \underset{h \rightarrow 0_{\mathbb{R}^n}}{=} f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle H_f(a) \cdot h, h \rangle + o(\|h\|^2)$$

□

## Équations aux dérivées partielles d'ordre 1

**Proposition 24.** Soient  $I, J$  des intervalles ouverts non vides de  $\mathbb{R}$ . Pour  $f \in \mathcal{C}^1(I \times J, \mathbb{R})$ , on a

$$\frac{\partial f}{\partial x} = 0 \iff \forall (x, y) \in I \times J \quad f(x, y) = B(y) \quad \text{avec} \quad B \in \mathcal{C}^1(J, \mathbb{R})$$

*Démonstration.* La fonction  $x \mapsto f(x, y)$  est de dérivée nulle sur un intervalle donc constante vis-à-vis de  $x$  mais dépend éventuellement de  $y$  d'où  $f(x, y) = B(y)$  pour  $(x, y) \in I \times J$ . Pour  $x_0 \in I$ , on a  $y \mapsto B(y) = f \circ (y \mapsto (x_0, y))$  de classe  $\mathcal{C}^1$  par composition d'où  $B \in \mathcal{C}^1(J, \mathbb{R})$ . La réciproque est immédiate. □

**Remarque :** On dispose d'un résultat identique pour  $\frac{\partial f}{\partial y}$  par symétrie des rôles.

**Exemple :** On a

$$\frac{\partial f}{\partial x}(x, y) = y \iff f : (x, y) \mapsto xy + B(y) \quad \text{avec} \quad B \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$$

**Définition 20.** Une application  $\varphi$  est un changement de variable de classe  $\mathcal{C}^1$  d'un ouvert  $U$  de  $\mathbb{R}^n$  sur  $V = \varphi(U) \subset \mathbb{R}^n$  si  $\varphi$  est bijective de  $U$  sur  $V$ , de classe  $\mathcal{C}^1$  avec  $\varphi^{-1}$  de classe  $\mathcal{C}^1$ . On dit aussi que  $\varphi$  est un  $\mathcal{C}^1$ -difféomorphisme.

**Exemple :** Passage en coordonnées polaires

L'application 
$$\begin{cases} ]0; +\infty[ \times ]-\pi; \pi[ \rightarrow \mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\}) \\ (r, \theta) \mapsto (r \cos(\theta), r \sin(\theta)) \end{cases}$$

est un changement de variables.

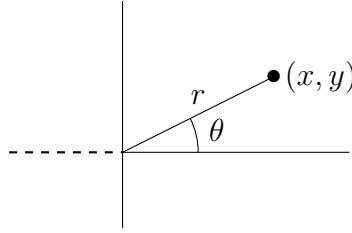


FIGURE 12 – Coordonnées polaires

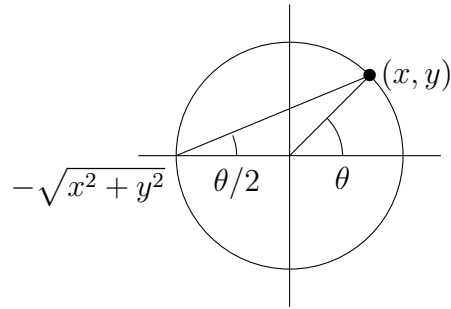
Par trigonométrie, on a

$$\tan\left(\frac{\theta}{2}\right) = \frac{y}{x + \sqrt{x^2 + y^2}}$$

d'où

$$\theta = 2 \operatorname{Arctan}\left(\frac{y}{x + \sqrt{x^2 + y^2}}\right)$$

et  $r = \sqrt{x^2 + y^2}$  d'où le caractère  $\mathcal{C}^1$  de l'application réciproque.



**Méthode avec changement de variables :** Deux configurations sont possibles.

• On a le changement de variables  $(u, v) = \varphi(x, y)$  avec  $\varphi : U \rightarrow V$  et  $U, V$  des ouverts de  $\mathbb{R}^2$ . Soit  $f \in \mathcal{C}^1(U, \mathbb{R})$ . On a le diagramme commutatif

$$\begin{array}{ccc} U & \xrightarrow{f} & \mathbb{R} \\ \varphi \downarrow & \nearrow \tilde{f} & \\ V & & \end{array}$$

encore  $\tilde{f} = f \circ \varphi^{-1}$  qui est donc de classe  $\mathcal{C}^1$ . D'après la règle de la chaîne, il vient

$$\begin{cases} \frac{\partial f}{\partial x} = \frac{\partial \tilde{f}}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial \tilde{f}}{\partial v} \frac{\partial v}{\partial x} \\ \frac{\partial f}{\partial y} = \frac{\partial \tilde{f}}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial \tilde{f}}{\partial v} \frac{\partial v}{\partial y} \end{cases}$$

En pratique, on utilise cette écriture « à la physicienne » pour résoudre simplement les EDP (équations aux dérivées partielles). On peut aussi obtenir cette égalité par le produit des matrices jacobiniennes correspondant à l'égalité  $df = d\tilde{f}(\varphi) \circ d\varphi$  :

$$\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial \tilde{f}}{\partial u} & \frac{\partial \tilde{f}}{\partial v} \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$$

• L'autre configuration est celle du changement de variables  $(x, y) = \varphi(u, v)$ . On a le diagramme commutatif

$$\begin{array}{ccc} V & \xrightarrow{\tilde{f}} & \mathbb{R} \\ \varphi \downarrow & \nearrow f & \\ U & & \end{array}$$

de la chaîne, il vient

$$\begin{cases} \frac{\partial \tilde{f}}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} \\ \frac{\partial \tilde{f}}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v} \end{cases}$$

**Exemples :** 1. Résoudre l'équation aux dérivées partielles

$$2 \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} = x^2 \quad (\text{E}_1)$$

à l'aide du changement de variables

$$\begin{cases} u = x \\ v = x + 2y \end{cases} \quad \text{avec} \quad (x, y) \in \mathbb{R}^2$$

On cherche  $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$  solution de l'équation aux dérivées partielles. Notons  $(u, v) = \varphi(x, y) = (x, x + 2y)$  le changement de variables. On a le diagramme

$$\begin{array}{ccc} \mathbb{U} & \xrightarrow{f} & \mathbb{R} \\ \varphi \downarrow & \nearrow \tilde{f} & \\ \mathbb{V} & & \end{array} \quad \text{donc}$$

$$f = \tilde{f} \circ \varphi.$$

On trouve

$$\begin{cases} \frac{\partial f}{\partial x} = \frac{\partial \tilde{f}}{\partial u} + \frac{\partial \tilde{f}}{\partial v} \\ \frac{\partial f}{\partial y} = 2 \frac{\partial \tilde{f}}{\partial v} \end{cases}$$

Ainsi  $f \in S_{E_1} \iff 2 \frac{\partial \tilde{f}}{\partial u} = u^2 \iff \tilde{f}(u, v) = \frac{u^3}{6} + A(v)$

avec  $A \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ . En revenant aux coordonnées d'origine, on conclut que les solutions de (E<sub>1</sub>) sont les fonctions de la forme

$$\boxed{\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = \frac{x^3}{6} + A(x + 2y) \quad \text{avec} \quad A \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})}$$

2. Résoudre l'équation aux dérivées partielles

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = 0 \quad (\text{E}_2)$$

avec les coordonnées polaires  $(x, y) = \varphi(r, \theta) = (r \cos(\theta), r \sin(\theta))$  où  $(r, \theta) \in ]0; +\infty[ \times ]-\pi; \pi[$  et  $(x, y) \in \Omega = \mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\})$ . On est dans la seconde configuration de changement de variables avec  $\tilde{f} = f \circ \varphi$ . D'après la règle de la chaîne, il vient

$$\begin{cases} \frac{\partial \tilde{f}}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \cos(\theta) \frac{\partial f}{\partial x} + \sin(\theta) \frac{\partial f}{\partial y} \\ \frac{\partial \tilde{f}}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = -r \sin(\theta) \frac{\partial f}{\partial x} + r \cos(\theta) \frac{\partial f}{\partial y} \end{cases}$$

En particulier

$$r \frac{\partial \tilde{f}}{\partial r} = x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y}$$

D'où  $f \in S_{E_2} \iff r \frac{\partial \tilde{f}}{\partial r} = 0 \iff \frac{\partial \tilde{f}}{\partial r} = 0 \iff \tilde{f}(r, \theta) = A(\theta)$

avec  $A \in \mathcal{C}^1(]-\pi; \pi[, \mathbb{R})$ . En revenant aux coordonnées d'origine, on conclut que les solutions de  $(E_2)$  sont les fonctions de la forme

$$\forall (x, y) \in \Omega \quad f(x, y) = A \left[ 2 \operatorname{Arctan} \left( \frac{x}{x + \sqrt{x^2 + y^2}} \right) \right] \quad \text{avec} \quad A \in \mathcal{C}^1(]-\pi; \pi[, \mathbb{R})$$

## Équations aux dérivées partielles d'ordre 2

**Proposition 25.** Soient  $I, J$  des intervalles ouverts non vides de  $\mathbb{R}$ . Pour  $f \in \mathcal{C}^2(I \times J, \mathbb{R})$ , on a

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y} = 0 &\iff f(x, y) = A(x) + B(y) \\ \frac{\partial^2 f}{\partial x^2} = 0 &\iff f(x, y) = xA(y) + B(y) \\ \frac{\partial^2 f}{\partial y^2} = 0 &\iff f(x, y) = yA(x) + B(x) \end{aligned}$$

avec  $A, B$  de classe  $\mathcal{C}^2$ .

*Démonstration.* On a

$$\frac{\partial^2 f}{\partial x \partial y} = 0 \iff \frac{\partial f}{\partial y} = b(y) \iff f(x, y) = A(x) + B(y)$$

D'après la proposition 24, la fonction  $b$  est de classe  $\mathcal{C}^1$  et une de ses primitives  $B$  est de classe  $\mathcal{C}^2$ . Ensuite, pour  $y_0 \in J$ , on a  $A : x \mapsto f \circ (x \mapsto (x, y_0)) - B(y_0)$  de classe  $\mathcal{C}^2$  par opérations sur de telles fonctions.

Puis 
$$\frac{\partial^2 f}{\partial x^2} = 0 \iff \frac{\partial f}{\partial x} = A(y) \iff f(x, y) = xA(y) + B(y)$$

Soit  $(x_1, x_2) \in I^2$  avec  $x_1 \neq x_2$ . On pose  $\varphi_1 : y \mapsto x_1 A(y) + B(y)$  et  $\varphi_2 : y \mapsto x_2 A(y) + B(y)$ . On trouve

$$\forall y \in J \quad A(y) = \frac{\varphi_1(y) - \varphi_2(y)}{x_1 - x_2} \quad \text{et} \quad B(y) = \frac{x_2 \varphi_1(y) - x_1 \varphi_2(y)}{x_2 - x_1}$$

Les fonctions  $\varphi_1, \varphi_2$  sont de classe  $\mathcal{C}^2$  par composition puisqu'on a  $\varphi_1 = f \circ (y \mapsto (x_1, y))$  et  $\varphi_2 = f \circ (y \mapsto (x_2, y))$  et le résultat suit pour les fonctions  $A$  et  $B$ . Les réciproques sont immédiates.  $\square$

**Définition 21.** Une application  $\varphi$  est un changement de variable de classe  $\mathcal{C}^k$  d'un ouvert  $U$  de  $\mathbb{R}^n$  sur  $V = \varphi(U) \subset \mathbb{R}^n$  si  $\varphi$  est bijective de  $U$  sur  $V$ , de classe  $\mathcal{C}^k$  avec  $\varphi^{-1}$  de classe  $\mathcal{C}^k$ . On dit aussi que  $\varphi$  est un  $\mathcal{C}^k$ -difféomorphisme.

**Remarque :** Il existe un théorème très efficace pour caractériser un  $\mathcal{C}^k$ -difféomorphisme (théorème d'inversion globale) mais celui-ci est hors-programme. Caractériser une telle application ne fait donc pas partie des attendus des candidats aux concours.

### Exemples importants : 1. L'équation des ondes

Soit  $c > 0$ . Résolvons 
$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0$$

où  $f : (x, t) \mapsto f(x, t)$  avec  $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$  à l'aide du changement de variable

$$\begin{cases} u = x + ct \\ v = x - ct \end{cases}$$

On a  $f = \tilde{f} \circ \varphi$ . Avec la règle de la chaîne, il vient

$$\begin{cases} \frac{\partial f}{\partial x} = \frac{\partial \tilde{f}}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial \tilde{f}}{\partial v} \frac{\partial v}{\partial x} = \frac{\partial \tilde{f}}{\partial u} + \frac{\partial \tilde{f}}{\partial v} \\ \frac{\partial f}{\partial t} = \frac{\partial \tilde{f}}{\partial u} \frac{\partial u}{\partial t} + \frac{\partial \tilde{f}}{\partial v} \frac{\partial v}{\partial t} = c \frac{\partial \tilde{f}}{\partial u} - c \frac{\partial \tilde{f}}{\partial v} \end{cases}$$

puis, en regroupant les dérivées croisées d'après le théorème de Schwarz

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x} \left( \frac{\partial \tilde{f}}{\partial u} + \frac{\partial \tilde{f}}{\partial v} \right) \\ &= \frac{\partial^2 \tilde{f}}{\partial u^2} \frac{\partial u}{\partial x} + \frac{\partial^2 \tilde{f}}{\partial v \partial u} \frac{\partial v}{\partial x} + \frac{\partial^2 \tilde{f}}{\partial u \partial v} \frac{\partial u}{\partial x} + \frac{\partial^2 \tilde{f}}{\partial v^2} \frac{\partial v}{\partial x} = \frac{\partial^2 \tilde{f}}{\partial u^2} + 2 \frac{\partial^2 \tilde{f}}{\partial u \partial v} + \frac{\partial^2 \tilde{f}}{\partial v^2} \end{aligned}$$

et

$$\begin{aligned} \frac{\partial^2 f}{\partial t^2} &= \frac{\partial}{\partial t} \left( c \frac{\partial \tilde{f}}{\partial u} - c \frac{\partial \tilde{f}}{\partial v} \right) \\ &= c \frac{\partial^2 \tilde{f}}{\partial u^2} \frac{\partial u}{\partial t} + c \frac{\partial^2 \tilde{f}}{\partial v \partial u} \frac{\partial v}{\partial t} - c \frac{\partial^2 \tilde{f}}{\partial u \partial v} \frac{\partial u}{\partial t} - c \frac{\partial^2 \tilde{f}}{\partial v^2} \frac{\partial v}{\partial t} = c^2 \frac{\partial^2 \tilde{f}}{\partial u^2} - 2c^2 \frac{\partial^2 \tilde{f}}{\partial u \partial v} + c^2 \frac{\partial^2 \tilde{f}}{\partial v^2} \end{aligned}$$

Ainsi

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0 \iff \frac{\partial^2 \tilde{f}}{\partial u \partial v} = 0$$

et d'après le résultat de la proposition 25, on a

$$\frac{\partial^2 \tilde{f}}{\partial u \partial v} = 0 \iff \forall (u, v) \in \mathbb{R}^2 \quad \tilde{f}(u, v) = A(u) + B(v)$$

avec A, B fonctions de  $\mathcal{C}^2(\mathbb{R}, \mathbb{R})$ . En revenant aux coordonnées d'origine, on conclut que les solutions de l'équation des ondes sont les fonctions de la forme

$$\boxed{\forall (x, t) \in \mathbb{R}^2 \quad f(x, t) = A(x + ct) + B(x - ct) \quad \text{avec} \quad A, B \text{ dans } \mathcal{C}^2(\mathbb{R}, \mathbb{R})}$$

## 2. Laplacien en coordonnées polaires

**Définition 22.** Soit  $U$  ouvert de  $\mathbb{R}^n$  et  $f \in \mathcal{C}^2(U, \mathbb{R})$ . On appelle Laplacien de  $f$  la fonction notée  $\Delta f$  définie par

$$\Delta f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} = \sum_{i=1}^n \partial_i^2 f$$

Soit  $f \in \mathcal{C}^2(U, \mathbb{R})$  avec  $U \subset \mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\})$ . Exprimons le Laplacien de  $f$  en fonction des coordonnées polaires  $(r, \theta)$  liées à  $(x, y)$  par le changement de variables  $(x, y) = (r \cos(\theta), r \sin(\theta))$ . On a établi précédemment

$$\begin{cases} \frac{\partial \tilde{f}}{\partial r} = \cos(\theta) \frac{\partial f}{\partial x} + \sin(\theta) \frac{\partial f}{\partial y} \\ \frac{\partial \tilde{f}}{\partial \theta} = -r \sin(\theta) \frac{\partial f}{\partial x} + r \cos(\theta) \frac{\partial f}{\partial y} \end{cases}$$

Toujours avec la règle de la chaîne et en regroupant les dérivées croisées d'après le théorème de Schwarz, il vient

$$\begin{aligned}
\frac{\partial^2 \tilde{f}}{\partial r^2} &= \frac{\partial}{\partial r} \left( \cos(\theta) \frac{\partial f}{\partial x} + \sin(\theta) \frac{\partial f}{\partial y} \right) \\
&= \cos(\theta) \frac{\partial}{\partial r} \left( \frac{\partial f}{\partial x} \right) + \sin(\theta) \frac{\partial}{\partial r} \left( \frac{\partial f}{\partial y} \right) \\
&= \cos(\theta) \left( \frac{\partial^2 f}{\partial x^2} \frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y \partial x} \frac{\partial y}{\partial r} \right) + \sin(\theta) \left( \frac{\partial^2 f}{\partial x \partial y} \frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y^2} \frac{\partial y}{\partial r} \right) \\
&= \cos(\theta) \left( \cos(\theta) \frac{\partial^2 f}{\partial x^2} + \sin(\theta) \frac{\partial^2 f}{\partial x \partial y} \right) + \sin(\theta) \left( \cos(\theta) \frac{\partial^2 f}{\partial x \partial y} + \sin(\theta) \frac{\partial^2 f}{\partial y^2} \right) \\
\frac{\partial^2 \tilde{f}}{\partial r^2} &= \cos(\theta)^2 \frac{\partial^2 f}{\partial x^2} + 2 \sin(\theta) \cos(\theta) \frac{\partial^2 f}{\partial x \partial y} + \sin(\theta)^2 \frac{\partial^2 f}{\partial y^2}
\end{aligned}$$

puis, avec les règles usuelles de dérivation d'un produit de fonctions

$$\begin{aligned}
\frac{\partial^2 \tilde{f}}{\partial \theta^2} &= \frac{\partial}{\partial \theta} \left( -r \sin(\theta) \frac{\partial f}{\partial x} + r \cos(\theta) \frac{\partial f}{\partial y} \right) \\
&= -r \cos(\theta) \frac{\partial f}{\partial x} - r \sin(\theta) \frac{\partial}{\partial \theta} \left( \frac{\partial f}{\partial x} \right) - r \sin(\theta) \frac{\partial f}{\partial y} + r \cos(\theta) \frac{\partial}{\partial \theta} \left( \frac{\partial f}{\partial y} \right) \\
&= -r \frac{\partial \tilde{f}}{\partial r} - r \sin(\theta) \left( \frac{\partial^2 f}{\partial x^2} \frac{\partial x}{\partial \theta} + \frac{\partial^2 f}{\partial y \partial x} \frac{\partial y}{\partial \theta} \right) + r \cos(\theta) \left( \frac{\partial^2 f}{\partial x \partial y} \frac{\partial x}{\partial \theta} + \frac{\partial^2 f}{\partial y^2} \frac{\partial y}{\partial \theta} \right) \\
&= -r \frac{\partial \tilde{f}}{\partial r} - r \sin(\theta) \left( -r \sin(\theta) \frac{\partial^2 f}{\partial x^2} + r \cos(\theta) \frac{\partial^2 f}{\partial x \partial y} \right) + r \cos(\theta) \left( -r \sin(\theta) \frac{\partial^2 f}{\partial x \partial y} + r \cos(\theta) \frac{\partial^2 f}{\partial y^2} \right) \\
\frac{\partial^2 \tilde{f}}{\partial \theta^2} &= -r \frac{\partial \tilde{f}}{\partial r} + r^2 \left( \sin(\theta)^2 \frac{\partial^2 f}{\partial x^2} - 2 \sin(\theta) \cos(\theta) \frac{\partial^2 f}{\partial x \partial y} + \cos(\theta)^2 \frac{\partial^2 f}{\partial y^2} \right)
\end{aligned}$$

Finalement, on obtient

$$\Delta f = \frac{\partial^2 \tilde{f}}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 \tilde{f}}{\partial \theta^2} + \frac{1}{r} \frac{\partial \tilde{f}}{\partial r}$$

## L'entropie d'un gaz parfait

Dans le cas d'un gaz parfait, l'entropie  $S : ]0; +\infty[^2 \rightarrow \mathbb{R}, (t, v) \mapsto S(t, v)$  vérifie la relation suivante :

$$dS = nc \frac{dT}{T} + nR \frac{dV}{V}$$

Les notations  $(dT, dV)$  désignent la base duale de la base canonique de  $\mathbb{R}^2$ , c'est-à-dire

$$\forall (t, v) \in \mathbb{R}^2 \quad dT(t, v) = t \quad \text{et} \quad dV(t, v) = v$$

$$\text{On a} \quad \forall (T, V, t, v) \in ]0; +\infty[^2 \times \mathbb{R}^2 \quad dS(T, V) \cdot (t, v) = nc \frac{t}{T} + nR \frac{v}{V}$$

Ainsi, par intégration le long d'un chemin, il vient pour  $(T_0, V_0)$  et  $(T, V)$  dans  $]0; +\infty[^2$

$$\begin{aligned}
S(T, V) - S(T_0, V_0) &= \int_0^1 dS((T_0, V_0) + u(T - T_0, V - V_0)) \cdot (T - T_0, V - V_0) du \\
&= \int_0^1 \left( nc \frac{T - T_0}{T_0 + u(T - T_0)} + nR \frac{V - V_0}{V_0 + u(V - V_0)} \right) du \\
S(T, V) - S(T_0, V_0) &= [nc \ln(T_0 + u(T - T_0)) + nR \ln(V_0 + u(V - V_0))]_0^1
\end{aligned}$$

On conclut

$$S(T, V) - S(T_0, V_0) = nc \ln \left( \frac{T}{T_0} \right) + nR \ln \left( \frac{V}{V_0} \right)$$

**Remarque :** Le calcul présenté en physique est cohérent avec le formalisme mathématique puisqu'on observe par linéarité de l'intégrale puis changement de variables ( $t = T_0 + u(T - T_0)$  pour la première intégrale et  $v = V_0 + u(V - V_0)$  pour la deuxième) :

$$\begin{aligned} S(T, V) - S(T_0, V_0) &= \int_0^1 nc \frac{T - T_0}{T_0 + u(T - T_0)} du + \int_0^1 nR \frac{V - V_0}{V_0 + u(V - V_0)} du \\ &= nc \int_{T_0}^T \frac{dt}{t} + nR \int_{V_0}^V \frac{dv}{v} \end{aligned}$$

Plus généralement, pour une entropie vérifiant

$$dS = f(T)dT + g(V)dV$$

avec  $f$  et  $g$  continues, on a

$$S(T, V) - S(T_0, V_0) = \int_{T_0}^T f(t) dt + \int_{V_0}^V g(v) dv$$

Ceci s'explique par le caractère « séparé » des variables dans l'écriture de la différentielle  $dS$ .