CH19 : Calcul différentiel - aspects théoriques

Table des matières

Dans tout ce chapitre, E et F désigneront deux $\mathbb{R}$-espaces vectoriels normés de dimension finie, avec $\dim(E) \geq 1$, et Ω un ouvert non vide de E.
On notera en général $n = \dim(E)$, $m = \dim(F)$. Lorsqu'on munit E et F de bases, elles seront notées respectivement $\mathcal{B}_E = (e_1, \cdots, e_n)$, $\mathcal{B}_F = (e'_1, \cdots, e'_m)$.
On aura la plupart du temps en contexte une application $f : \Omega \subset E \to F$.

I Applications différentiables

1) Dérivée directionnelle

Si $f : I \to F$, où I est un intervalle de $\mathbb{R}$ d'intérieur non vide, alors f est dérivable en a ssi $\lim\limits_{x\to a} \dfrac{f(x) - f(a)}{x - a}$ existe et est finie. Cela revient à dire qu'il existe un unique vecteur de F, noté $f'(a)$, tel que :

$$\forall x \in \Omega, \qquad f(x) = f(a) + (x-a)f'(a) + (x-a)\varepsilon(x), \text{ où } \varepsilon(x) \underset{x\to a}{\longrightarrow} 0_F.$$

(voir CH.15 sur les fonctions vectorielles).

Si on envisage maintenant le cas d'une fonction **à variable vectorielle**, c'est-à-dire $f : \Omega \to F$ avec $\Omega \subset E$, alors on ne peut reprendre telle quelle ce DL_1 en a pour donner un sens à la "dérivabilité" de f en a, car l'expression $(x - a)f'(a)$ n'a plus de sens (on ne peut pas faire le "produit" de $x - a \in E$ par $f'(a) \in F$).
Pour contrer ce problème, on peut avoir envie d'essayer de se restreindre à des fonctions scalaires, pour se ramener à un cadre bien connu. D'où l'idée :

Définition 1 (Dérivée directionnelle suivant un vecteur)
Soit $f : \Omega \to F$, soit $a \in \Omega$ et $v \in E$.
On dit que f est **dérivable en a selon le vecteur v** *lorsque la fonction d'une variable réelle $t \mapsto f(a + tv)$ est dérivable en 0. On note alors*

$$D_v f(a) = \left[\frac{d}{dt}(t \mapsto f(a+tv))\right]_{t=0} = \lim_{t\to 0} \frac{1}{t}\big(f(a+tv) - f(a)\big) \in F,$$

et on dit que $D_v f(a)$ est la **dérivée directionnelle de f en a selon v**.
Lorsqu'elle est bien définie, l'application $D_v f : \left\{\begin{array}{ccc} \Omega & \longrightarrow & F \\ a & \longmapsto & D_v f(a) \end{array}\right.$ est appelée **dérivée directionnelle de f selon v**.

Remarque

- *Si $v = 0_E$, alors la dérivée directionnelle de f en tout point a est bien définie, mais elle est nulle, donc ne sert pas à grand chose.*
- *Lorsque $v \neq 0_E$, l'application $t \mapsto a + tv$ est un paramétrage de la droite affine passant par a et dirigée par v (la droite $\Delta_{a,v} = a + Vect(v)$).*
  L'application $t \mapsto f(a + tv)$ est donc un paramétrage de la restriction de f à $\Delta_{a,v}$.
- *Comme Ω est ouvert, il existe un voisinage $]-\delta, \delta[$ de 0 dans $\mathbb{R}$ tel que*

$$|t| \leq \delta \implies a + tv \in \Omega \implies f(a+tv) \text{ existe.}$$

  En effet, puisque $a \in \Omega$ et Ω est un ouvert de E, il existe $r > 0$ tel que $\overline{B(a,r)} \subset \Omega$, donc $a + tv \in \overline{B(a,r)} \subset \Omega$ dès que $|t| \times \|v\| \leq r$.
  Ainsi, l'application $t \mapsto f(a + tv)$ est bien définie au voisinage de $t = 0$.

Exemple
Soit $f : \left\{\begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & x - y^2 \end{array}\right.$.
Déterminer les dérivées directionnelles de f en $(0,0)$.

On a, pour tout $v = (a, b) \in \mathbb{R}^2$, $f\big((0,0) + tv\big) = f(ta, tb) = ta - t^2b^2$, qui admet comme dérivée a en 0. Ainsi $D_v f(0,0) = a$.

ATTENTION !
Une fonction peut admettre en un point des dérivées directionnelles selon tout vecteur sans même être continue !

Exemple
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ telle que $f(x, y) = \dfrac{y^3}{x}$ lorsque $x \neq 0$ et $f(0, y) = 0$.
Montrer que f est dérivable en $(0,0)$ selon n'importe quel vecteur et montrer que f n'est pas continue en $(0,0)$.

Si $v = (0, b)$, alors $f\big((0,0) + tv\big) = f(0, tb) = 0$ donc $D_v f(0,0) = 0$.
Si $v = (a, b)$ avec $a \neq 0$, alors $f\big((0,0) + tv\big) = f(ta, tb) = t^2\frac{b^3}{a}$ qui admet pour dérivée 0 en 0. Donc $D_v f(0,0) = 0$.
Ainsi f admet des dérivées directionnelles selon toutes les directions en 0, et elles sont nulles.
Mais f n'est pas continue en 0 ! En effet, $f\left(\frac{1}{n^3}, \frac{1}{n}\right) = 1$ et donc $f\left(\frac{1}{n^3}, \frac{1}{n}\right)$ ne tend pas vers $f(0,0) = 0$ lorsque $n \to +\infty$, ce qui contredit le critère séquentiel de continuité.

Ceci est problématique car on souhaite, comme pour les fonctions scalaires, qu'une fonction dérivable soit continue. Ainsi les dérivées directionnelles ne forment pas un bon outil pour définir la dérivabilité d'une fonction à variable vectorielle. C'est pourquoi on va introduire une notion plus complexe, celle de différentielle en un point.

2) Différentiabilité en un point

Revenons au cas des fonctions à variable réelle : si $f : I \to F$ est dérivable en a (avec I intervalle ouvert de $\mathbb{R}$), alors pour tout $h \in \mathbb{R}$ tel que $a + h \in I$,

$$f(a+h) = f(a) + hf'(a) + h\varepsilon(h), \text{ où } \varepsilon(h) \underset{h\to 0}{\longrightarrow} 0_F,$$

et on remarque que l'application $\left\{\begin{array}{ccc} \mathbb{R} & \longrightarrow & F \\ h & \longmapsto & hf'(a) \end{array}\right.$ est linéaire, donc continue puisqu'on est en dimension finie (ce qui assure $f(a+h) \underset{h\to 0}{\longrightarrow} f(a)$).

On va donc exploiter ce point de vue :

$$\boxed{\text{lorsque } h \text{ est petit,} \qquad f(a+h) = \text{valeur } f(a) + \text{application linéaire en } h + \text{erreur en } o(h)},$$

c'est la grande idée du calcul différentiel, qui permet de généraliser la notion de dérivée.

Définition 2 (Application différentiable en un point)
Soit $f : \Omega \to F$ et soit $a \in \Omega$.
On dit que f est **différentiable en a** *lorsqu'il existe une application linéaire $L : E \to F$ et une application $\varepsilon : E \to F$ telle que $\varepsilon(h) \underset{h\to 0_E}{\longrightarrow} 0_F$ et pour tout $h \in E$ tel que $a + h \in \Omega$, on a*

$$f(a+h) = f(a) + L(h) + \|h\|\varepsilon(h).$$

Remarque

- *Là encore, le fait que Ω soit ouvert assure que dès que $\|h\|$ est suffisamment petit, on a $a+h \in \Omega$.*
- *L'ensemble de définition de la fonction ε importe peu, il suffit qu'elle soit définie au voisinage de 0_E.*
- *f est différentiable en a ssi il existe $L : E \to F$ linéaire et de $\tilde{\varepsilon} : \Omega \to F$ de limite nulle en a telles que :*
$$\forall x \in \Omega, \qquad f(x) = f(a) + L(x-a) + (x-a)\tilde{\varepsilon}(x).$$

- *On peut bien sûr utiliser les petits-o, et écrire* $f(a+h) = f(a) + L(h) + o(h)$ *ou encore* $f(x) = f(a) + L(x-a) + o(x-a)$.
  Le "$o(h)$*" peut aussi se noter* $o(\|h\|)$ *puisque la notion de négligeabilité s'énonce avec des normes (tout comme la notion de limite).*
- *Comme* E *est de dimension finie,* L *est automatiquement continue vu que linéaire. Pour votre culture, lorsqu'il s'agit de définir le concept de différentiabilité avec* E *non nécessairement de dimension finie, on prend la même définition en imposant* L *linéaire* ***et continue****.*
- *Attention, au contraire de* f *et de* ε*, l'application* L *est nécessairement définie sur* E *entier (et pas que sur* Ω *par exemple) car linéaire.*

Propriété 3 (Différentiable implique continue)
Si f *est différentiable en* a*, alors* f *est continue en* a*.*

Preuve
L'application L est linéaire sur E de dimension finie, donc continue. Ainsi $L(h) \underset{h\to 0_E}{\longrightarrow} 0_F$, ce qui entraîne $f(a+h) \underset{h\to 0_E}{\longrightarrow} f(a)$ (puisque $\|h\|\varepsilon(h) \to 0_F$), et donc f est continue en a.

Théorème 4 (Unicité de la différentielle et lien avec les dérivées directionnelles)
Si f *est différentiable en* a*, alors :*

(i) *pour tout* $v \in E$*,* $D_v f(a)$ *existe et* $L(v) = D_v f(a)$ *;*

(ii) *l'application linéaire* L *est unique. On dit que* L *est la* **différentielle de f en a** *(ou l'***application linéaire tangente à f en a***). On note* $L = df(a)$ *(ou* df_a*), d'où la relation :*

$$\forall v \in E, \qquad df(a)(v) = D_v f(a).$$

Preuve
Soient $v \in E$ et $t \in \mathbb{R}$ tel que $a+tv \in \Omega$. Alors en posant $h = tv$ dans la définition de la différentiabilité :

$$f(a+tv) = f(a) + L(tv) + |t|\|v\|\varepsilon(tv)$$

donc, par linéarité de L :

$$t \neq 0 \implies \frac{1}{t}\big(f(a+tv) - f(a)\big) = L(v) + \frac{|t|}{t}\|v\|\varepsilon(tv) = L(v) \pm \|v\|\varepsilon(tv),$$

avec $\varepsilon(tv) \underset{t\to 0}{\longrightarrow} 0_F$ par composition de limites. Donc $\frac{1}{t}\big(f(a+tv) - f(a)\big) \underset{t\to 0}{\longrightarrow} L(v)$, ce qui prouve que f admet une dérivée directionnelle en a selon v, et que $D_v f(a) = L(v)$. Le reste s'en déduit.

ATTENTION !
La réciproque du point (i) est fausse ! Une application qui possède des dérivées directionnelles en a n'est pas nécessairement différentiable en a (elle peut ne même pas être continue, voir les exemples du début).
La formule du point (ii) permet donc seulement de déterminer le candidat pour $df(a)$, pas de montrer que f est différentiable en a. Le seul moyen pour cela est de vérifier la définition.

Notation
Ainsi, lorsque f est différentiable en a, on peut écrire

$$f(a+h) = f(a) + df(a)(h) + o(h),$$

avec $df(a) \in \mathcal{L}(E,F)$.
Pour tout vecteur $h \in E$, on note en général $df(a) \cdot h$ plutôt que $df(a)(h)$ (comme on travaille en dimension finie, on peut assimiler cela à un produit matriciel). Ainsi, la formule du point (ii) se réécrit :

$$\forall v \in E, \qquad df(a) \cdot v = D_v f(a).$$

Remarque (IMPORTANT)
$df(a)$ est l'application linéaire $E \to F$ qui approche le mieux $f(a+h) - f(a)$ pour h voisin de 0_E.

Exemple

- Toute fonction f constante est différentiable en tout $a \in \Omega$, avec $df(a) = 0_{\mathcal{L}(E,F)}$.
- Si $f : E \to F$ est linéaire, alors f est différentiable en tout $a \in E$ et $df(a) = f$.
- *Soit* $f : \left\{ \begin{array}{ccc} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ M & \longmapsto & M^2 \end{array} \right.$ *et* $A \in \mathcal{M}_n(\mathbb{R})$.
  Montrer que f est différentiable en A et expliciter $df(A)$.
  En effet, pour $H \in \mathcal{M}_n(\mathbb{R})$, $f(A+H) = (A+H)^2 = A^2 + AH + HA + H^2$.
  Or, $H \mapsto AH + HA$ est linéaire.
  Et, munissant $\mathcal{M}_n(\mathbb{R})$ d'une norme sous-multiplicative, on a $\|H^2\| \leq \|H\|^2$ donc
  $$\|H^2\| = o(\|H\|) = o(H).$$
  On en conclut que f est différentiable en A avec $df(A) : \left\{ \begin{array}{ccc} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ H & \longmapsto & AH + HA \end{array} \right.$.

Méthode
Par contraposition, on a donc une méthode pour montrer qu'une fonction n'est pas différentiable en un point : trouver au moins un vecteur selon lequel elle n'admet pas de dérivée directionnelle en ce point !

Exemple
Montrer que l'application $\|.\| : E \to \mathbb{R}$ n'est pas différentiable en 0_E.
Soit $u \in E \backslash \{0_E\}$ et $t \in \mathbb{R}$. On a $\|tu\| = |t|\|u\|$ donc $\frac{\|tu\| - \|0_E\|}{t} \underset{t \to 0_+}{\longrightarrow} \|u\|$ et $\frac{\|tu\| - \|0_E\|}{t} \underset{t \to 0_-}{\longrightarrow} -\|u\| \neq \|u\|$, ce qui montre que $\|.\|$ n'admet pas de dérivée directionnelle en 0_E selon tous les vecteurs, donc qu'elle n'est pas différentiable en 0_E.

Enfin, le résultat qui suit fait le lien explicite entre les notions de différentiabilité et de dérivabilité pour une fonction de la variable réelle :

Propriété 5 (Différentiabilité d'une fonction d'une variable réelle)
Soient I un intervalle ouvert de $\mathbb{R}$, soit $f : I \to F$ et $a \in I$.
Alors f est dérivable en a ssi f est différentiable en a et, lorsque c'est le cas, on a
$$f'(a) = df(a) \cdot 1,$$
et donc,
$$\forall h \in \mathbb{R}, \qquad df(a) \cdot h = hf'(a).$$

Preuve

$\boxed{\Longrightarrow}$ Supposons f dérivable en a. Alors on sait qu'on a, pour $h \to 0$, $f(a+h) = f(a) + hf'(a) + o(h)$, avec $h \to hf'(a)$ qui est linéaire. Donc f est différentiable en a et $df(a) \cdot h = hf'(a)$.

$\boxed{\Longleftarrow}$ Supposons f différentiable en a.
Pour $h \neq 0$, on a $\frac{1}{h}\big(f(a+h) - f(a)\big) = \frac{1}{h}\big(df(a) \cdot h + o(h)\big) = df(a) \cdot 1 + o(1)$ par linéarité de $df(a)$.
Donc $\frac{1}{h}\big(f(a+h) - f(a)\big) \underset{h \to 0}{\longrightarrow} df(a) \cdot 1$, ce qui prouve que f est dérivable en a et que $f'(a) = df(a) \cdot 1$.

3) Différentiabilité sur un ouvert

Définition 6 (Application différentiable sur un ouvert)
On dit que $f : \Omega \to F$ est différentiable sur Ω lorsque f est différentiable en tout point de Ω.

Définition 7 (Application différentielle)
Si f est différentiable sur Ω, alors on peut définir l'application $df : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & \mathcal{L}(E,F) \\ a & \longmapsto & df(a) \end{array} \right.$.
On l'appelle **différentielle de f**.

ATTENTION !
La différentielle df n'est pas une application linéaire, c'est la différentielle $df(a)$ **en un point $a \in E$ fixé** qui est linéaire.

Exemple

- Si f est constante, alors df est nulle.
- Si f est linéaire, alors df est constante égale à f.
- On notera, grâce à la linéarité, que les applications coordonnées $\mathbb{R}^n \to \mathbb{R}$, ou encore la partie réelle et la partie imaginaire, sont des applications différentiables, de différentielles constantes et égales à elle-même.

Définition 8 (Fonctions coordonnées de f dans une base de F)
Etant donnée une base $\mathcal{B}_F = (e'_1, \dots, e'_m)$ de F et une application $f : \Omega \to F$, on appelle **fonctions coordonnées de f dans $\mathcal{B}_F$** *les applications $f_1 : \Omega \to \mathbb{R}, \cdots, f_m : \Omega \to \mathbb{R}$ définies par :*

$$\forall x \in \Omega, \qquad f(x) = \sum_{j=1}^{m} f_j(x) e'_j.$$

On pourra noter en abrégé $[f]_{\mathcal{B}_F} = (f_1, \cdots, f_m)$.

Vocabulaire
Les fonctions coordonnées $(f_1, \cdots, f_m)$ sont aussi appelées les **composantes** de f dans la base de F choisie.

Comme on l'avait déjà aperçu pour la dérivabilité des fonctions vectorielles, la différentiabilité peut être caractérisée sur les fonctions coordonnées de f.

Propriété 9 (Différentiabilité coordonnée par coordonnée)
Soit $f : \Omega \to F$ avec $m = \dim(F) \geq 1$. On note $(f_1, \dots, f_m)$ les fonctions coordonnées de f dans une base quelconque $\mathcal{B}_F$.
Alors f est différentiable ssi $f_1, \dots, f_m$ sont différentiables et dans ce cas :

$$\forall a \in \Omega, \qquad [df(a)]_{\mathcal{B}_F} = (df_1(a), \cdots, df_m(a)).$$

Preuve
Pour toutes applications $L : E \to F$ et $\varepsilon : E \to F$, notons $(L_1, \cdots, L_m)$ et $(\varepsilon_1, \cdots, \varepsilon_m)$ leurs coordonnées respectives dans $\mathcal{B}_F$. On a alors :

$$f(a+h) = f(a) + L(h) + \|h\|\varepsilon(h) \iff \forall j, \quad f_j(a+h) = f_j(a) + L_j(h) + \|h\|\varepsilon_j(h),$$

ainsi que les équivalences :

$$L \in \mathcal{L}(E,F) \iff \forall j, \quad L_j \in \mathcal{L}(E,\mathbb{R}),$$
$$\varepsilon(h) \underset{h \to 0_E}{\longrightarrow} 0_F \iff \forall j, \quad \varepsilon_j(h) \underset{h \to 0_E}{\longrightarrow} 0,$$

donc f est diff en a ssi $\forall j$, f_j est diff en a et on a dans ce cas $df(a) = L$ et $df_j(a) = L_j$ pour tout j, donc

$$[df(a)]_{\mathcal{B}_F} = [L]_{\mathcal{B}_F} = (L_1, \cdots, L_m) = (df_1(a), \cdots, df_m(a)).$$

Exemple
La fonction $f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x,y) & \longmapsto & (x+2y,1) \end{array} \right.$ est différentiable car ses fonctions coordonnées $(x,y) \mapsto x+2y$ et $(x,y) \mapsto 1$ le sont en tant qu'application linéaire pour la première et en tant qu'application constante pour la deuxième.

4) Dérivées partielles dans une base

Soit $\mathcal{B}_E = (e_1,\dots,e_n)$ une base de E. On décompose un vecteur h quelconque dans cette base : $h = \sum\limits_{i=1}^{n} h_i e_i$. On obtient alors, si f est différentiable en un certain $a \in \Omega$:

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

par linéarité de $df(a)$. Et donc,

$$df(a)\cdot h = \sum_{i=1}^{n} h_i D_{e_i} f(a),$$

d'après la formule qui relie différentielle en a et dérivées directionnelles en a.

Il suffit donc de connaître les n dérivées directionnelles $D_{e_i} f(a)$ pour connaître l'application $df(a)$.

Ceci justifie l'introduction de la notion suivante :

Définition 10 (Dérivées partielles en un point par rapport à une base)
Soit $\mathcal{B}_E = (e_1,\dots,e_n)$ une base de E.

(i) *Lorsqu'elles existent, on note $\dfrac{\partial f}{\partial x_1}(a),\dots,\dfrac{\partial f}{\partial x_n}(a)$ les dérivées directionnelles de f en a selon les vecteurs $e_1,\dots,e_n$. Pour tout $i \in [\![1,n]\!]$, on appelle $\dfrac{\partial f}{\partial x_i}(a)$ la **i-ème dérivée partielle** de f en a relativement à la base $\mathcal{B}_E$.*
Par définition, on a donc :

$$\forall i \in [\![1,n]\!], \qquad \frac{\partial f}{\partial x_i}(a) = D_{e_i} f(a) = \left[\frac{d}{dt}(t \mapsto f(a+te_i))\right]_{t=0} = \lim_{t\to 0}\frac{1}{t}\big(f(a+te_i)-f(a)\big)$$

lorsque cette limite existe.

(ii) *Sous réserve d'existence, l'application $\dfrac{\partial f}{\partial x_i} =: \left\{ \begin{array}{ccc} \Omega & \longrightarrow & F \\ a & \longmapsto & \dfrac{\partial f}{\partial x_i}(a) \end{array} \right.$ est appelée **i-ème fonction dérivée partielle** de f (relativement à la base $\mathcal{B}_E$), pour tout $i \in [\![1,n]\!]$.*

Notation
On note aussi $\partial_i f(a) = \dfrac{\partial f}{\partial x_i}(a) = \dfrac{\partial f}{\partial e_i}(a)$ la i^e dérivée partielle de f en a.
La notation $\dfrac{\partial f}{\partial e_i}(a)$ est la plus claire car elle indique le vecteur suivant lequel on dérive f.
Mais la notation la plus utilisée est $\dfrac{\partial f}{\partial x_i}(a)$, il faut bien comprendre que le "x_i" est symbolique, il indique juste que l'on dérive par rapport à la i^e coordonnée, et sous-entend que les coordonnées d'un vecteur $x \in E$ dans $\mathcal{B}_E$ seront notées $(x_1,\cdots,x_n)$.

Exemple
Soit $f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & xy^3 \end{array} \right.$.
Calculer les dérivées partielles de f en tout point $a=(x,y)$ dans la base canonique.

On obtient facilement :

$$\forall (x,y) \in \mathbb{R}^2, \qquad \frac{\partial f}{\partial x}(x,y) = y^3, \qquad \frac{\partial f}{\partial y}(x,y) = 3xy^2.$$

Pour la première dérivée partielle, on fixe $y \in \mathbb{R}$ et on dérive la fonction $x \mapsto f(x,y)$, et pour la seconde c'est l'inverse (ou alors on repasse par les taux d'accroissement).

Remarque

- *En pratique, on se rend compte que* $\dfrac{\partial f}{\partial x_i}(a) = \lim\limits_{t \to 0} \dfrac{1}{t}\big(f(a+te_i) - f(a)\big)$ *n'est rien d'autre que la dérivée en* 0 *de la fonction de la variable réelle* $f_i : t \mapsto f(a+te_i)$*, fonction qu'on appelle souvent* ***i*-ième application partielle** *au point* a*.*
- *Lorsqu'elles sont bien définies, les dérivées partielles* $\dfrac{\partial f}{\partial x_i}$ *sont comme* f *des fonctions à variable et valeurs vectorielles, et il y en a* n *(autant que* $\dim(E)$*).*

Exemple
Soit $E = \mathbb{R}_2[X]$ *muni de sa base canonique* $\mathcal{C} = (1, X, X^2)$ *et*

$$f : \left\{ \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ P = a_0 + a_1 X + a_2 X^2 & \longmapsto & f(P) = \sin(a_2) + \cos(a_1) + \cos(a_0) a_2^3 \end{array} \right. .$$

Calculer la troisième dérivée partielle de f *en* $Q = 1 + X^2$ *relativement à* $\mathcal{C}$*.*

Première méthode :
On introduit la troisième application partielle au point Q :

$$f_3 : t \mapsto f(Q + tX^2) = f\big(1 + (1+t)X^2\big) = \sin(1+t) + \cos(1)(1+t)^3.$$

Cette fonction est polynomiale donc dérivable sur $\mathbb{R}$, et on a

$$\forall t \in \mathbb{R}, \qquad f_3'(t) = \cos(1+t) + 3\cos(1)(1+t)^2,$$

donc f admet une troisième dérivée partielle en Q et

$$\partial_3 f(Q) = \frac{\partial f}{\partial X^2}(Q) = \frac{\partial f}{\partial a_2}(Q) = f_3'(0) = 4\cos(1).$$

Deuxième méthode :
On calcule $\dfrac{\partial f}{\partial a_2}$ en un polynôme quelconque $P = a_0 + a_1 X + a_2 X^2$, puis on évalue en $Q = 1 + X^2$.
Pour cela, on fixe a_0, a_1 et on dérive l'expression de $f(P)$ par rapport à a_2 (elle est bien dérivable comme somme et composée de fonctions dérivables sur $\mathbb{R}$). Cela donne

$$\forall P = a_0 + a_1 X + a_2 X^2 \in \mathbb{R}_2[X], \qquad \frac{\partial f}{\partial a_2}(P) = \cos(a_2) + 3\cos(a_0) a_2^2,$$

d'où

$$\frac{\partial f}{\partial a_2}(Q) = \cos(1) + 3\cos(1) \times 1^2 = 4\cos(1).$$

Propriété 11 (Lien entre différentielle et dérivées partielles)
Si f *est différentiable en* a *alors les dérivées partielles de* f *en* a *relativement à toute base* $\mathcal{B}_E$ *existent et vérifient :* $\forall i \in [\![1, n]\!]$, $\quad \dfrac{\partial f}{\partial x_i}(a) = df(a) \cdot e_i$*. De plus, on a l'expression :*

$$\forall h \in E, \qquad df(a) \cdot h = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a),$$

où $[h]_{\mathcal{B}_E} = (h_1, \cdots, h_n)$*.*

Preuve
La première formule est directe d'après la formule du th. 4, puisque les dérivées partielles sont des dérivées directionnelles. La seconde formule s'en déduit par linéarité de $df(a)$.

<u>ATTENTION !</u>
Une fonction peut admettre toutes ses dérivées partielles sans être différentiable !
Cf. l'exemple donné plus haut de $f : \mathbb{R}^2 \to \mathbb{R}$ telle que $f(x, y) = \frac{y^3}{x}$ lorsque $x \neq 0$ et $f(0, y) = 0$.
Ce n'est que lorsqu'une fonction est différentiable que l'on est sûr de pouvoir exprimer sa différentielle à l'aide des dérivées partielles.

Corollaire 12 (Formule de Taylor-Young à l'ordre 1)
Si f est différentiable en a, alors on a le $DL_1(a)$:

$$f(a + h) = f(a) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + o(h),$$

où les $\dfrac{\partial f}{\partial x_i}(a)$ désignent les dérivées partielles en a dans une quelconque base de E.

Preuve
Simple réécriture de la définition de la différentiabilité en a, en exprimant le terme linéaire $df(a) \cdot h$ en fonction des dérivées partielles $\frac{\partial f}{\partial x_i}(a)$ grâce à la prop. précédente.

Remarque
Si f possède des dérivées partielles en a, alors on a la réciproque :

$$f(a + h) = f(a) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + o(h) \implies f \text{ est différentiable en } a.$$

En effet, si cette écriture est vraie, alors l'application $L : h \mapsto \sum_{i=1}^{n} h_i \dfrac{\partial f}{\partial x_i}(a)$ étant linéaire, la définition de la différentiabilité en a est vérifiée.
Cela donne une méthode pour montrer que f est différentiable en a, en connaissant simplement ses dérivées partielles en a : il suffit de vérifier que

$$\left\| f(a + h) - f(a) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) \right\| \underset{h \to 0_E}{=} o(\|h\|)$$

(on teste la seule différentielle possible : l'application $h \mapsto \sum_{i=1}^{n} h_i \dfrac{\partial f}{\partial x_i}(a)$).

5) Classe $\mathcal{C}^1$

On a la généralisation suivante du concept de classe $\mathcal{C}^1$ déjà vu pour les fonctions à variable réelle :

Définition 13 (Fonction de classe $\mathcal{C}^1$)
Une fonction $f : \Omega \to F$ est dite **de classe $\mathcal{C}^1$** *lorsqu'elle est différentiable et lorsque l'application $df : \Omega \to \mathcal{L}(E, F)$ est continue.*

Remarque
La continuité de df est compliquée à traduire, car df est un objet abstrait (ce n'est d'ailleurs pas une application linéaire !). Si on munit $\mathcal{L}(E, F)$ de la norme subordonnée $\|\!| \cdot |\!\|$, définie par

$$\forall u \in \mathcal{L}(E, F), \qquad |\!|\!| u |\!|\!| = \sup_{x \neq 0_E} \frac{\|u(x)\|_F}{\|x\|_E},$$

alors la continuité de df s'écrit :

$$\forall a \in \Omega,\ \forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in \Omega, \quad \|x-a\| \leq \delta \implies \|df(x) - df(a)\| \leq \varepsilon,$$

Heureusement, le théorème suivant donne une caractérisation simple de la classe $\mathcal{C}^1$.

Théorème 14 (Caractérisation des fonctions de classe $\mathcal{C}^1$)
Une fonction $f : \Omega \to F$ est de classe $\mathcal{C}^1$ ssi ses dérivées partielles (relativement à une base quelconque de E) existent en tout point de Ω et sont continues sur Ω.

Remarque
Du moment qu'il existe une base sur laquelle cette condition est satisfaite, la fonction est de classe $\mathcal{C}^1$ et ses dérivées partielles relativement à n'importe quelle base existent et sont continues.
En pratique, on choisit donc la base qui nous arrange !

Preuve (non traitée en classe)

$\boxed{\Longrightarrow}$: Supposons f de classe $\mathcal{C}^1$, c'est-à-dire que f est différentiable sur Ω et df est continue.
Soit $\mathcal{B}_E = (e_1, \ldots, e_n)$ une base de E. Comme f est différentiable, ses dérivées partielles dans la base $\mathcal{B}_E$ existent et vérifient :

$$\forall i \in \{1, \ldots, n\},\ \forall a \in \Omega,\ \partial_i f(a) = df(a) \cdot e_i = df(a)(e_i).$$

Or $df : \Omega \to \mathcal{L}(E,F)$ est continue et, e_i étant fixé, l'application $\varepsilon_i : \left\{ \begin{array}{ccc} \mathcal{L}(E,F) & \longrightarrow & F \\ u & \longmapsto & u(e_i) \end{array} \right.$
("évaluation en e_i") est linéaire sur un espace de dimension finie donc continue.
Par composition, on en déduit que pour tout $i \in [\![1,n]\!]$, la dérivée partielle $\partial_i f = \varepsilon_i \circ df$ est continue.

$\boxed{\Longleftarrow}$: La réciproque est non exigible. On se contente d'en voir l'idée dans un cadre simplifié, celui de fonctions à deux variables (cas $E = \mathbb{R}^2$ avec $\mathcal{B}_E$ la base canonique).
On suppose donc que f admet des dérivées partielles $\partial_1 f$ et $\partial_2 f$ continues.

* Soit $(a,b) \in \Omega$. Montrons que f est différentiable en (a,b).
  Pour cela, il suffit de montrer que

  $$f(a+h, b+k) - f(a,b) - h\partial_1 f(a,b) - k\partial_2 f(a,b) = o(\|(h,k)\|)$$

  au voisinage de $(k,h) = (0,0)$ (on teste la seule différentielle possible).
  On note $r > 0$ tel que $B_f\big((a,b), r\big) \subset \Omega$.
  Pour tout $(h,k) \in E$ tel que $\|(h,k)\| \leq r$, on introduit l'accroissement
  $\Delta(h,k) = f(a+h,b+k) - f(a,b) = f(a+h,b+k) - f(a,b+k) + f(a,b+k) - f(a,b)$.
  Par hypothèse, les applications $x \mapsto f(x, b+k)$ et $y \mapsto f(a,y)$ sont dérivables de dérivées respectives $x \mapsto \partial_1 f(x, b+k)$ et $y \mapsto \partial_2 f(a,y)$ continues. On a donc, par le théorème fondamental de l'analyse,

  $$\Delta(h,k) = \int_a^{a+h} \partial_1 f(x, b+k)dx + \int_b^{b+k} \partial_2 f(a,y)dy = h\int_0^1 \partial_1 f(a+th, b+k)dt + k\int_0^1 \partial_2 f(a, b+tk)dt.$$

  On introduit maintenant le reste $r(h,k) = f(a+h,b+k) - f(a,b) - h\partial_1 f(a,b) - k\partial_2 f(a,b)$.
  Par ce qui précède, on a :

  $$r(h,k) = h\int_0^1 \big(\partial_1 f(a+th, b+k) - \partial_1 f(a,b)\big)dt + k\int_0^1 \big(\partial_2 f(a, b+tk) - \partial_2 f(a,b)\big)dt.$$

  Soit $\varepsilon > 0$. La continuité de $\partial_1 f$ et $\partial_2 f$ nous dit :

  $$\exists \eta_1 > 0,\ \forall (s,t) \in \mathbb{R}^2,\ \|(s,t)\| \leq \eta_1 \implies \|\partial_1 f(a+s, b+t) - \partial_1 f(a,b)\| \leq \frac{\varepsilon}{2}.$$

  $$\exists \eta_2 > 0,\ \forall t \in \mathbb{R},\ |t| \leq \eta_2 \implies \|\partial_2 f(a, b+t) - \partial_2 f(a,b)\| \leq \frac{\varepsilon}{2}.$$

On note $\eta = \min(\eta_1, \eta_2)$. On a donc, pour tout (h,k) tel que $\|(h,k)\| \leq \eta$,

$$|r(h,k)| \leq |h| \int_0^1 \|\partial_1 f(a+th, b+k) - \partial_1 f(a,b)\| \, dt + |k| \int_0^1 \|\partial_2 f(a, b+tk) - \partial_2 f(a,b)\| \, dt,$$

$$|r(h,k)| \leq |h|\frac{\varepsilon}{2} + |k|\frac{\varepsilon}{2} \leq \varepsilon \|(h,k)\|_\infty.$$

Ainsi $r(h,k)$ est négligeable devant $\|(h,k)\|_\infty$, d'où :

$$f(a+h, b+k) = f(a,b) + h\partial_1 f(a,b) + k\partial_2 f(a,b) + o(h,k),$$

avec $(h,k) \mapsto h\partial_1 f(a,b) + k\partial_2 f(a,b)$ linéaire, ce qui prouve que f est différentiable en (a,b).

Remarque
Cette preuve n'est au fond qu'une généralisation de celle du théorème fondamental de l'analyse.

* On sait alors qu'on a, pour tout $(a,b) \in \Omega$ et tout $(h,k) \in \mathbb{R}^2$,

$$df(a,b) \cdot (h,k) = \partial_1 f(a,b)h + \partial_2 f(a,b)k,$$

donc $df : (a,b) \mapsto \partial_1 f(a,b)p_1 + \partial_2 f(a,b)p_2$, où p_1 et p_2 sont les deux projections canoniques de $\mathbb{R}^2$. Comme $\partial_1 f$ et $\partial_2 f$ sont continues par hypothèse, on en déduit que df est continue.

Exemple

- Les fonctions constantes sont de classe $\mathcal{C}^1$ car de dérivées partielles nulles, donc continues.
- Les applications linéaires sont de classe $\mathcal{C}^1$ car leurs dérivées partielles sont constantes donc continues.
  En effet, pour tout $a \in \Omega$, $\partial_i L(a) = dL(a) \cdot e_i = L(e_i)$.
- Il s'ensuit notamment que pour toute base $\mathcal{B}_E = (e_i)_{1 \leq i \leq n}$ de E, les applications coordonnées

$$e_i^* : \left\{ \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x_i \end{array} \right. , \qquad 1 \leq i \leq n,$$

(où $(x_1, \cdots, x_n) = [x]_{\mathcal{B}_E}$) sont de classe $\mathcal{C}^1$.

Finissons par une évidence :

Propriété 15 ($\mathcal{C}^1$ implique $\mathcal{C}^0$)
Les fonctions de classe $\mathcal{C}^1$ sont continues.

Preuve
Car différentiables.

En résumé, pour $f : \Omega \to \mathbb{R}$:

$$\boxed{f \text{ de classe } \mathcal{C}^1 \implies f \text{ différentiable} \left\{ \begin{array}{ll} \implies & f \text{ continue} \\ \implies & f \text{ possède des dérivées partielles en tout point de } \Omega \end{array} \right. ,}$$

(et les implications réciproques sont fausses) et

$$\boxed{f \text{ de classe } \mathcal{C}^1 \iff f \text{ possède des dérivées partielles } \textbf{continues} \text{ en tout point de } \Omega.}$$

6) Matrice jacobienne

En tant qu'application linéaire, la différentielle $df(a) \in \mathcal{L}(E,F)$ peut se représenter par sa matrice A dans un couple de bases. En notant $(e_1,\cdots,e_n)$ une base de E et $(e'_1,\cdots,e'_m)$ une base de F, les colonnes de cette matrice A vont être définies par

$$\forall j \in [\![1,n]\!], \qquad C_j(A) = [df(a)\cdot e_j]_{\mathcal{B}_F} = \left[\frac{\partial f}{\partial x_j}(a)\right]_{\mathcal{B}_F} = \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(a) & \cdots & \frac{\partial f_m}{\partial x_j}(a)\end{pmatrix}^\top$$

(en notant $(f_1,\cdots,f_m)$ les coordonnées de f dans la base $\mathcal{B}_F$ et en dérivant "coordonnée par coordonnée").

Définition 16 (Matrice jacobienne dans un couple de bases)
Soit $\mathcal{B}_E = (e_1,\cdots,e_n)$ une base de E et $\mathcal{B}_F = (e'_1,\cdots,e'_m)$ une base de F.
Soit $f : \Omega \to F$ une application différentiable en $a \in \Omega$.
On appelle **matrice jacobienne de f dans les bases $(\mathcal{B}_E,\mathcal{B}_F)$** *la matrice de l'application linéaire $df(a) \in \mathcal{L}(E,F)$ dans les bases $(\mathcal{B}_E,\mathcal{B}_F)$. On la note :*

$$Jac\, f(a) = Mat_{\mathcal{B}_E,\mathcal{B}_F}\big(df(a)\big) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{i,j} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a)\end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R}),$$

où $(f_1,\cdots,f_m)$ sont les coordonnées de f dans la base $\mathcal{B}_F$ et $(\frac{\partial f_i}{\partial x_1},\cdots,\frac{\partial f_i}{\partial x_n})$ désignent les dérivées partielles des $f_i : \Omega \to \mathbb{R}$ relativement à la base $\mathcal{B}_E$.

Notation
On pourra noter plus simplement $J_f(a)$ la matrice jacobienne de f en a.

Remarque

- *Bien retenir que les lignes de la jacobienne correspondent aux composantes f_i que l'on dérive, alors que les colonnes correspondent aux variables x_j par rapport auxquelles on dérive.*
- *En utilisant les notations allégées pour les dérivées partielles, on a donc :*

$$Jac\, f(a) = \big(\partial_j f_i(a)\big)_{i,j} = \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \vdots & & \vdots \\ \partial_1 f_m(a) & \dots & \partial_n f_m(a)\end{pmatrix}.$$

- *Le $DL_1(a)$ de f se réécrit donc en coordonnées :*

$$[f(a+h)]_{\mathcal{B}_F} = [f(a)]_{\mathcal{B}_F} + Jac\, f(a) \times \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + o(h),$$

ou encore

$$\begin{pmatrix} f_1(a+h) \\ \vdots \\ f_m(a+h)\end{pmatrix} = \begin{pmatrix} f_1(a) \\ \vdots \\ f_m(a)\end{pmatrix} + Jac\, f(a) \times \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + o(h),$$

en notant $(h_1,\dots,h_n)$ les coordonnées de h dans $\mathcal{B}_E$ et $(f_1,\cdots,f_m)$ les composantes de f dans $\mathcal{B}_F$.

- *En général, pour une fonction $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, on représente la matrice jacobienne dans les bases canoniques. On ne le précise d'ailleurs souvent même pas.*

Exemple

- Soit $f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ (x,y,z) & \longmapsto & (2x^2+y-z^3, xyz) \end{array}\right.$. On a $Jac\, f(x,y,z) = \begin{pmatrix} 4x & 1 & -3z^2 \\ yz & xz & xy \end{pmatrix}$.

- Soit $\varphi : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (r,\theta) & \longmapsto & (r\cos\theta, r\sin\theta) \end{array} \right.$. On a $Jac\,\varphi(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$.
Notons que par lecture de cette matrice, on en déduit le DL_1 en (r,θ) de φ :
$$\varphi(r+h,\theta+k) = \varphi(r,\theta) + (h\cos\theta - kr\sin\theta, h\sin\theta + kr\cos\theta) + o(h,k).$$

7) Gradient d'une fonction numérique

Dans ce paragraphe, les fonctions sont à valeurs dans $\mathbb{R}$ (d'où le terme "fonction numérique"), et définies sur un ouvert Ω de E, où E est un **espace euclidien** ($\mathbb{R}$-e.v. de dimension finie muni d'un produit scalaire $(\cdot|\cdot)$).

Théorème 17 (Gradient en un point)
Soit $f : \Omega \to \mathbb{R}$ différentiable en $a \in \Omega$. Il existe un unique vecteur de E, noté $\nabla f(a)$, tel que
$$\forall h \in E, \qquad df(a) \cdot h = (\nabla f(a)|h).$$
Ce vecteur $\nabla f(a)$ est appelé **gradient de f en a**.

Preuve
$df(a) : E \to \mathbb{R}$ est une forme linéaire, donc le résultat est direct d'après le théorème de représentation des formes linéaires sur un espace euclidien.

Remarque
On dispose donc d'une définition **intrinsèque** *du gradient, c'est-à-dire qui ne dépend pas d'un choix de base.*

Propriété 18 (Expression du gradient en base orthonormée)
Soit $\mathcal{B}_E = (e_1, \cdots, e_n)$ une base orthonormée de E, et $f : \Omega \to \mathbb{R}$ différentiable en $a \in \Omega$. Alors
$$[\nabla f(a)]_{\mathcal{B}_E} = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a) \right).$$

Preuve
Notons $(y_1, \cdots, y_n)$ les coordonnées de $\nabla f(a) \in E$ dans la base $\mathcal{B}_E = (e_1, \cdots, e_n)$. Puisque la base est orthonormée, on a
$$\forall i \in [\![1,n]\!], \qquad y_i = (\nabla f(a)|e_i) = df(a) \cdot e_i = \frac{\partial f}{\partial x_i}(a).$$

Corollaire 19 (Fonction gradient d'une fonction de classe $\mathcal{C}^1$)
Si $f : \Omega \to \mathbb{R}$ est de classe $\mathcal{C}^1$, alors la fonction gradient $\nabla f : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & E \\ a & \longmapsto & \nabla f(a) \end{array} \right.$ est continue.

Vocabulaire
On dit que le gradient ∇f est un **champ de vecteurs** de E (nom donné aux applications $\Omega \subset E \to E$).

Preuve
Si on fixe une base orthonormée de E, alors les composantes de l'application ∇f (qui est bien définie sur Ω car f est différentiable) sont les dérivées partielles de f (d'après la prop. précédente), qui sont continues (puisque f est de classe $\mathcal{C}^1$). Donc ∇f est continue, puisque ses composantes dans une base le sont.

Exemple

- Si $E = \mathbb{R}^n$ muni de sa structure euclidienne canonique, alors pour $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ différentiable, on a
$$\forall a \in \Omega, \qquad \nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a)\right) \in \mathbb{R}^n$$
(dans $\mathbb{R}^n$, on confond vecteur et coordonnées dans la base canonique, comme d'habitude).
- Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^m$ munis de leurs structures euclidiennes canonique, alors pour $f : \left\{ \begin{array}{ccc} \Omega \subset \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ x & \longmapsto & (f_1(x), \cdots, f_m(x)) \end{array} \right.$ différentiable, on a
$$\forall a \in \Omega, \qquad J_f(a) = \begin{pmatrix} -\nabla f_1(a)- \\ \vdots \\ -\nabla f_m(a)- \end{pmatrix}.$$
("on empile en ligne les gradients des composantes pour former la jacobienne")

Remarque (Le gradient pointe la direction de plus grande pente)
Si $\nabla f(a) \neq 0$, alors $\nabla f(a)$ pointe la direction de plus grande pente de f, c'est-à-dire que $\nabla f(a)$ est positivement colinéaire au vecteur v unitaire tel que le nombre réel $D_v f(a)$ est maximal.
En effet, pour tout vecteur $v \in E$ unitaire, on a
$$D_v f(a) = df(a) \cdot v = (\nabla f(a)|v) \leq \|\nabla f(a)\| \times \|v\| = \|\nabla f(a)\|,$$
(inégalité de Cauchy-Schwarz), avec égalité si et seulement si il existe $\lambda > 0$ tel que $v = \lambda \nabla f(a)$, c'est-à-dire ssi $v = \frac{\nabla f(a)}{\|\nabla f(a)\|}$.
Cette "direction de plus grande pente" est donc la droite affine passant par a suivant laquelle les valeurs de f croissent le plus rapidement au voisinage de a.

Exemple (Fonction définie par une intégrale)
Soit $\varphi : \mathbb{R} \to \mathbb{R}$ continue et $f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \displaystyle\int_x^y \varphi(t)dt \end{array} \right.$.
Montrer que f est de classe $\mathcal{C}^1$ et déterminer son gradient.
Soit Φ une primitive de φ sur $\mathbb{R}$. On a $f(x,y) = \Phi(y) - \Phi(x)$, avec Φ dérivable, donc f admet en tout point une première et une deuxième dérivées partielles :
$$\frac{\partial f}{\partial x}(x,y) = -\varphi(x) \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = \varphi(y).$$
La continuité de φ implique celle de $\dfrac{\partial f}{\partial x}$ et de $\dfrac{\partial f}{\partial y}$. Ainsi f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$, et pour tout $(x,y) \in \mathbb{R}^2$:
$$\nabla f(x,y) = (-\varphi(x), \varphi(y)).$$

Exemple (Gradient de la norme euclidienne)
Soit E un espace euclidien et $f = \|\cdot\| : E \to \mathbb{R}$ la norme euclidienne sur E.

1. *Montrer que f est différentiable en tout point $a \neq 0_E$ et déterminer $df(a)$.*
2. *En déduire $\nabla f(a)$ pour tout $a \neq 0_E$.*

Remarque
L'exemple précédent montre que l'approche des DL pour déterminer la différentielle est rapidement limitée lorsque les calculs se compliquent. C'est pourquoi on va établir des formules générales pour calculer des différentielles, notamment pour calculer la différentielle d'une fonction composée.

II Opérations sur les applications différentiables

On revient au cadre général des applications $f : \Omega \subset E \to F$, où E et F sont des $\mathbb{R}$-evn de dimension finie et Ω est un ouvert de E.

1) Combinaisons linéaires

Théorème 20 (Différentielle d'une combinaison linéaire en un point)
Soient $f : \Omega \to F$, $g : \Omega \to F$ et $(\lambda, \mu) \in \mathbb{R}^2$.
Si f et g sont différentiables en $a \in \Omega$, alors $\lambda f + \mu g$ aussi et

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

Preuve
Soit $h \in E$ tel que $a + h \in \Omega$. On a :

$$(\lambda f+\mu g)(a+h) = \lambda f(a+h)+\mu g(a+h) = \lambda\big(f(a)+df(a)\cdot h+\|h\|\varepsilon(h)\big)+\mu\big(g(a)+dg(a)\cdot h+\|h\|\tilde{\varepsilon}(h)\big)$$

$$= (\lambda f + \mu g)(a) + \big(\lambda df(a) + \mu dg(a)\big) \cdot h + \|h\|\big(\lambda\varepsilon(h) + \mu\tilde{\varepsilon}(h)\big),$$

avec $h \mapsto \big(\lambda df(a) + \mu dg(a)\big) \cdot h$ linéaire et $\lambda\varepsilon(h) + \mu\tilde{\varepsilon}(h) \underset{h\to 0_E}{\longrightarrow} 0_F$, d'où $\lambda f + \mu g$ est différentiable en a et $d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a)$.

Corollaire 21 (Différentielle d'une combinaison linéaire)
Si f et g sont différentiables sur Ω alors, pour tout $(\lambda, \mu) \in \mathbb{R}^2$, $\lambda f + \mu g$ est différentiable sur Ω, avec

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

Il s'ensuit que l'ensemble des fonctions différentiables de Ω vers F est un sous-espace vectoriel de F^Ω (les fonctions de Ω dans F).

Preuve
On applique la proposition précédente en tout point $a \in \Omega$.

2) Composition

G désigne ici un autre $\mathbb{R}$-evn de dimension finie.

Théorème 22 (Différentielle d'une composée en un point)
Soient Ω un ouvert de E et Ω' un ouvert de F. Soit $f : \Omega \to F$ et $g : \Omega' \to G$ telles que $f(\Omega) \subset \Omega'$.
Si f est différentiable en $a \in \Omega$ et si g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et

$$d(g \circ f)(a) = \big[dg\big(f(a)\big)\big] \circ df(a).$$

Remarque

- *C'est la généralisation de la dérivée d'une composée de deux fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$: si $f : I \to \mathbb{R}$ est dérivable en a et si $g : J \to \mathbb{R}$ est dérivable en $f(a)$ (avec I et J deux intervalles de $\mathbb{R}$ tels que $f(I) \subset J$), alors $g \circ f : I \to \mathbb{R}$ est dérivable en a et*

  $$(g \circ f)'(a) = g'(f(a)) \times f'(a) \in \mathbb{R}.$$

  On peut en effet redémontrer cette formule à l'aide du théorème précédent et de la prop. 5 : f est dérivable en a, donc différentiable en a, et $\forall h \in \mathbb{R},\ df(a) \cdot h = h \times f'(a)$. De même, g est dérivable en $f(a)$, donc différentiable en $f(a)$, et $\forall h \in \mathbb{R},\ dg(f(a)) \cdot h = h \times g'(f(a))$.
  Donc d'après le théorème de composition $g \circ f$ est différentiable en a, ce qui revient à dire que $g \circ f$ est dérivable en a, et

  $$(g\circ f)'(a) = d(g\circ f)(a)\cdot 1 = (dg(f(a))\circ df(a))\cdot 1 = dg(f(a))\cdot(df(a)\cdot 1) = dg(f(a))\cdot f'(a) = g'(f(a))\times f'(a).$$

- *Si on choisit des bases de E, F, G, alors la formule se traduit matriciellement par un produit de jacobiennes :*
$$J_{g\circ f}(a) = J_g(f(a)) \times J_f(a).$$

Preuve
On a
$$f(a+h) = f(a) + df(a)\cdot h + \|h\|\varepsilon(h) = f(a) + h',$$
avec $h' = df(a)\cdot h + \|h\|\varepsilon(h) \underset{h\to 0_E}{\longrightarrow} 0_F$, ainsi que
$$g\big(f(a+h)\big) = g\big(f(a)+h'\big) = g\big(f(a)\big) + dg\big(f(a)\big)\cdot h' + \|h'\|\tilde{\varepsilon}(h'),$$
avec $\tilde{\varepsilon}(h') \underset{h'\to 0_F}{\longrightarrow} 0_G$.
Donc
$$(g\circ f)(a+h) = g\big(f(a)\big) + dg\big(f(a)\big)\cdot\big(df(a)\cdot h\big) + \varphi(h),$$
où $\varphi(h) = \|h\|dg\big(f(a)\big)\cdot\varepsilon(h) + \|h'\|\tilde{\varepsilon}(h')$.
On a $df(a)$ linéaire continue donc $\|df(a)\cdot h\| \le k_f\|h\|$ et donc $\|h'\| \le (k_f + |\varepsilon(h)|)\|h\|$. Ainsi $\varphi(h) = o(h)$.
On a donc écrit
$$(g\circ f)(a+h) = g\big(f(a)\big) + dg\big(f(a)\big)\cdot\big(df(a)\cdot h\big) + o(h),$$
avec $dg\big(f(a)\big)\cdot\big(df(a)\cdot h\big) = \Big(dg\big(f(a)\big)\circ df(a)\Big)\cdot h$ et $dg\big(f(a)\big)\circ df(a)$ linéaire par composition d'applications linéaires, donc $g\circ f$ est différentiable en a et $d(g\circ f)(a) = dg\big(f(a)\big)\circ df(a)$.

Corollaire 23 (Composition extérieure avec une fonction réelle)
Soient I un intervalle réel, $f : \Omega \to \mathbb{R}$ et $\varphi : I \to \mathbb{R}$ telles que $f(\Omega) \subset I$.
Si f est différentiable en $a \in \Omega$ et si φ est dérivable en $f(a)$, alors $\varphi\circ f$ est différentiable en a et :
$$d(\varphi\circ f)(a) = \varphi'(f(a)) \times df(a)$$
(produit du nombre réel $\varphi'(f(a))$ par l'application linéaire $df(a)$).

Preuve
Par composition, on a $d(\varphi\circ f)(a) = d\varphi\big(f(a)\big)\circ df(a)$. Or φ étant à variable réelle, on a pour tout $\tilde{h} \in \mathbb{R}$, $d\varphi\big(f(a)\big)\cdot\tilde{h} = \varphi'\big(f(a)\big)\times\tilde{h}$, donc, pour tout $h \in E$:
$$d(\varphi\circ f)(a)\cdot h = d\varphi\big(f(a)\big)\cdot(df(a)\cdot h) = \varphi'(f(a)) \times (df(a)\cdot h) = (\varphi'(f(a))\times df(a))\cdot h,$$
d'où le résultat.

Corollaire 24 (Dérivée le long d'un arc)
Soient I un intervalle réel, $\gamma : I \to E$ et $f : \Omega \to F$ telles que $\gamma(I) \subset \Omega$.
Si γ est dérivable en $t_0 \in I$ et si f est différentiable en $\gamma(t_0)$, alors $f\circ\gamma$ est dérivable en t_0 et
$$(f\circ\gamma)'(t_0) = df\big(\gamma(t_0)\big)\cdot\gamma'(t_0).$$

Remarque
- *On appelle "arc" une fonction de la variable réelle à valeurs vectorielles $\gamma : I \subset \mathbb{R} \to E$. On peut voir γ comme une "trajectoire" (point de l'espace E qui se déplace en fonction du "temps" $t \in I$).*
- *On rappelle qu'un arc γ est dérivable en t_0 ssi il est différentiable.*
- *Les dérivées directionnelles sont des dérivées selon des arcs particuliers, ceux paramétrant des morceaux de droites (i.e. $\gamma(t) = a + tv$, où $a \in \Omega$ et $v \in E$, et dans ce cas, on a pour tout $t \in I$ $(f\circ\gamma)'(t) = df\big(\gamma(t)\big)\cdot\gamma'(t) = df(a+tv)\cdot v$).*

Preuve
On a $f\circ\gamma$ différentiable (c'est-à-dire dérivable) en t_0 par composition et :
$$(f\circ\gamma)'(t_0) = d(f\circ\gamma)(t_0)\cdot 1 = \Big(df\big(\gamma(t_0)\big)\circ d\gamma(t_0)\Big)\cdot 1 = df\big(\gamma(t_0)\big)\cdot\gamma'(t_0).$$

Corollaire 25 (Dérivée le long d'un arc tracé dans $\mathbb{R}^n$)
On suppose que $E = \mathbb{R}^n$. Soit Ω un ouvert non vide de $\mathbb{R}^n$, soit $f : \Omega \to F$, soient $x_1 : I \to \mathbb{R}, \cdots, x_n : I \to \mathbb{R}$, où I est un intervalle réel.
On suppose que $\big(x_1(t), \dots, x_n(t)\big) \in \Omega$ pour tout $t \in I$.
Soit $g : I \to F$ définie par

$$g(t) = f\big(x_1(t), \dots, x_n(t)\big).$$

Si les x_i sont dérivables en $t_0 \in I$ et si f est différentiable en $\big(x_1(t_0), \dots, x_n(t_0)\big)$, alors g est dérivable en t_0 et

$$\begin{aligned} g'(t_0) &= x_1'(t_0)\frac{\partial f}{\partial x_1}\big(x_1(t_0), \dots, x_n(t_0)\big) + \cdots + x_n'(t_0)\frac{\partial f}{\partial x_n}\big(x_1(t_0), \dots, x_n(t_0)\big) \\ &= \sum_{i=1}^n x_i'(t_0)\frac{\partial f}{\partial x_i}\big(x_1(t_0), \dots, x_n(t_0)\big), \end{aligned}$$

où les $\dfrac{\partial f}{\partial x_i}$ désignent les dérivées partielles de f relativement à la base canonique de $\mathbb{R}^n$.

Remarque
On peut retenir la forme synthétique : $\dfrac{dg}{dt} = \sum\limits_{i=1}^n \dfrac{dx_i}{dt}\dfrac{\partial f}{\partial x_i}$.

Preuve
On note $\gamma : I \to \mathbb{R}^n$ définie par $\gamma(t) = \big(x_1(t), \dots, x_n(t)\big)$. Ainsi $g = f \circ \gamma$.
Par dérivation d'une composée, on a $g'(t_0) = (f \circ \gamma)'(t_0) = df\big(\gamma(t_0)\big) \cdot \gamma'(t_0) = \sum\limits_{i=1}^n x_i'(t_0)\dfrac{\partial f}{\partial x_i}\big(\gamma(t_0)\big)$, vu que $\gamma'(t_0) = \big(x_1'(t_0), \dots, x_n'(t_0)\big)$ (en dérivant "composante par composante", cf. chapitre 15 sur les fonctions vectorielles).

Exemple
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction différentiable et $g : \mathbb{R} \to \mathbb{R}$ définie par $g(t) = f\big(\cos(t), \sin(t)\big)$.
Par composition, la fonction g est dérivable et

$$\forall t \in \mathbb{R}, \qquad g'(t) = -\sin(t)\frac{\partial f}{\partial x}\big(\cos(t), \sin(t)\big) + \cos(t)\frac{\partial f}{\partial y}\big(\cos(t), \sin(t)\big).$$

3) Composition bilinéaire, multilinéaire, produit

On a déjà vu que toute application linéaire $L : E \to F$ est différentiable et

$$\forall a \in E, \qquad dL(a) = L.$$

Pour les applications bilinéaires, la situation est un peu plus compliquée :

Théorème 26 (Différentielle d'une application bilinéaire)
Soient F, G, H trois $\mathbb{R}$-evn de dimension finie, et $B : F \times G \to H$ une application bilinéaire. Alors B est différentiable et

$$\forall (a,b) \in F \times G, \qquad dB(a,b) : \left\{ \begin{array}{ccc} F \times G & \longrightarrow & H \\ (u,v) & \longmapsto & B(u,b) + B(a,v) \end{array} \right. .$$

Symboliquement, on peut écrire

$$\forall (a,b) \in F \times G, \qquad dB(a,b) = B(\cdot, b) + B(a, \cdot).$$

Preuve

On peut généraliser :

Théorème 27 (Différentielle d'une application multilinéaire)
Soient $F_1, \cdots, F_p, G$ des $\mathbb{R}$-evn de dimension finie, et $M : F_1 \times \cdots \times F_p \to G$ une application multilinéaire. Alors M est différentiable et pour tout $(a_1, \cdots, a_p) \in F_1 \times \cdots \times F_p$:

$$dM(a_1, \cdots, a_p) : \left\{ \begin{array}{ccc} F_1 \times \cdots \times F_p & \longrightarrow & G \\ (u_1, \cdots, u_p) & \longmapsto & M(u_1, a_2, \cdots, a_p) + \cdots + M(a_1, \cdots, a_{p-1}, u_p) \end{array} \right. .$$

Preuve

Corollaire 28 (Différentielle d'une composition multilinéaire)
Soient $f_1 : \Omega \to F_1, \cdots, f_p : \Omega \to F_p$ des applications différentiables en $a \in \Omega$ (ouvert de E).
Soit $M : F_1 \times \cdots \times F_p \to G$ une application multilinéaire.
Alors, l'application

$$M(f_1, \cdots, f_p) : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & G \\ x & \longmapsto & M(f_1(x), \cdots, f_p(x)) \end{array} \right. ,$$

est différentiable en a, et

$$d(M(f_1, \cdots, f_p))(a) : h \mapsto M(df_1(a) \cdot h, f_2(a), \cdots, f_p(a)) + \cdots + M(f_1(a), \cdots, f_{p-1}(a), df_p(a) \cdot h).$$

Remarque
Pour une composition bilinéaire, cela s'écrit : si $f : \Omega \to F$, $g : \Omega \to G$ sont différentiables en $a \in \Omega$ et si $B : F \times G \to H$ est bilinéaire, alors $B(f, g) : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & H \\ x & \longmapsto & B(f(x), g(x)) \end{array} \right.$ est différentiable en a et

$$d(B(f, g))(a) : h \mapsto B(df(a) \cdot h, g(a)) + B(f(a), dg(a) \cdot h).$$

Preuve

Corollaire 29 (Différentielle d'un produit)
Soient $f, g : \Omega \to F$. Si F est une $\mathbb{R}$-algèbre de dimension finie et si f et g sont différentiables en $a \in \Omega$, alors $f \times g : \Omega \to F$ est différentiable en a et

$$d(fg)(a) : h \mapsto (df(a) \cdot h) \times g(a) + f(a) \times (dg(a) \cdot h),$$

qu'on peut abréger en

$$d(fg)(a) = df(a) \times g(a) + f(a) \times dg(a).$$

Ainsi le produit de deux fonctions différentiables est différentiables, et l'ensemble des fonctions différentiables de Ω vers F est une sous-algèbre de F^Ω.

Preuve
Car $B : \left\{ \begin{array}{ccc} F \times F & \longrightarrow & F \\ (x, y) & \longmapsto & x \times y \end{array} \right.$ est une application bilinéaire.

Exemple
Les applications polynomiales $\mathbb{R}^n \to \mathbb{R}$ sont différentiables.
La fonction déterminant $\mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ est différentiable (somme et produit de fonctions différentiables).

Exemple
Justifier que l'application $f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x, y, z) & \longmapsto & x^3 + xy^2 + yz^3 \end{array} \right.$ est différentiable et calculer $df(x, y, z)$

en tout $(x,y,z) \in \mathbb{R}^3$. f est polynomiale donc différentiable et, en utilisant les dérivées partielles,

$$\forall (a,b,c) \in \mathbb{R}^3, \qquad df(x,y,z)\cdot(a,b,c) = a(3x^2+y^2) + b(2xy+z^3) + c(3yz^2).$$

Exemple
Les fractions rationnelles sur $\mathbb{R}^p$ sont différentiables sur leur ouvert de définition, car produits d'une fonction polynomiale et d'une composée d'une fonction polynomiale et de la fonction inverse qui est dérivable, donc différentiable, sur $\mathbb{R}^*$.

Exemple
Soit $f : \Omega \to \mathbb{R}$ une application différentiable.
Pour tout $n \in \mathbb{N}^*$, $f^n : \Omega \to \mathbb{R}$ est différentiable et $d(f^n) = nf^{n-1}df$, c'est-à-dire
$d(f^n)(a) = \big(nf^{n-1}(a)\big) \times df(a)$, c'est-à-dire

$$d(f^n)(a) : h \mapsto \big(nf^{n-1}(a)\big) \times (df(a)\cdot h)$$

Exemple
Différentiabilité de $f : \left\{ \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \|x\| \end{array} \right.$.
On a déjà vu que f n'est pas différentiable en 0_E.
Introduisons $g : \left\{ \begin{array}{ccc} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \|x\|^2 \end{array} \right.$. On peut écrire $f = \sqrt{g}$, où $\sqrt{\ }$ est dérivable sur $\mathbb{R}_+^*$ et $g(E\backslash\{0_E\}) = \mathbb{R}_+^*$. Comme $g = (Id_E|Id_E)$, avec $(.|.)$ bilinéaire et Id_E linéaire, g est différentiable en tout point par composition, avec $dg(a) = 2\big(Id(a)|dId(a)\big)$, donc $dg(a)\cdot h = 2(a|h)$.
Et donc, pour $a \neq 0_E$,

$$df(a) : h \mapsto \frac{1}{2\sqrt{g(a)}} dg(a)\cdot h = \frac{(a|h)}{\|a\|}.$$

4) Opérations sur les dérivées partielles

Vu qu'une dérivée partielle est une dérivée "classique" (i.e. une dérivée de fonction d'une variable réelle), on dispose directement de formules pour les dérivées partielles d'une somme, d'un produit, etc. (en reprenant les formules de dérivation du CH.15 sur les fonctions vectorielles), et ce même si les fonctions en jeu ne sont pas différentiables.

Théorème 30 (Dérivées partielles d'une combinaison linéaire)
Soient $f, g : \Omega \to F$ *et* $(\lambda,\mu) \in \mathbb{R}^2$.
Si f *et* g *admettent des dérivées partielles en* $a \in \Omega$, *alors* $\lambda f + \mu g$ *aussi et*

$$\forall i \in [\![1,n]\!], \qquad \partial_i(\lambda f + \mu g)(a) = \lambda \partial_i f(a) + \mu \partial_i g(a).$$

Théorème 31 (Dérivées partielles d'une composition bilinéaire)
Soient F, G *et* H *trois* $\mathbb{R}$*-espaces vectoriels de dimension finie,* $f : \Omega \to F$, $g : \Omega \to G$ *et* $B : F \times G \to H$ *une application bilinéaire.*
Si f *et* g *admettent des dérivées partielles en* $a \in \Omega$, *alors* $B(f,g)$ *aussi et*

$$\forall i \in [\![1,n]\!], \qquad \partial_i\big(B(f,g)\big)(a) = B\big(\partial_i f(a), g(a)\big) + B\big(f(a), \partial_i g(a)\big).$$

Corollaire 32 (Dérivées partielles d'un produit)
Soient $f, g : \Omega \to F$. *Si* F *est une* $\mathbb{R}$*-algèbre et si* f *et* g *admettent des dérivées partielles en* a, *alors* $f \times g$ *aussi et*

$$\forall i \in [\![1,n]\!], \qquad \partial_i(fg)(a) = \big(\partial_i f(a)\big) \times g(a) + f(a) \times \big(\partial_i g(a)\big).$$

Lorsqu'on compose deux applications différentiables, on peut exprimer simplement les dérivées partielles de $g \circ f$ en fonction de celles de g et de f. C'est l'objet du résultat suivant :

Théorème 33 (Règle de la chaîne)
Soient Ω un ouvert de E et Ω' un ouvert de F. Soit $f : \Omega \to F$ et $g : \Omega' \to G$ telles que $f(\Omega) \subset \Omega'$.
On munit E d'une base $\mathcal{B}_E = (e_1, \cdots, e_n)$ et F d'une base $\mathcal{B}_F = (e'_1, \cdots, e'_m)$.
Si f est différentiable en $a \in \Omega$ et si g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a, donc admet des dérivées partielles en a (relativement à la base $\mathcal{B}_E$), et :

$$\forall i \in [\![1, n]\!], \qquad \partial_i(g \circ f)(a) = \sum_{k=1}^{m} \partial_i f_k(a) \times \partial_k g\big(f(a)\big),$$

c'est-à-dire

$$\forall i \in [\![1, n]\!], \qquad \frac{\partial(g \circ f)}{\partial e_i}(a) = \sum_{k=1}^{m} \frac{\partial f_k}{\partial e_i}(a) \times \frac{\partial g}{\partial e'_k}(f(a)),$$

où $(f_1, \cdots, f_m)$ désignent les fonctions coordonnées de f dans la base $\mathcal{B}_F$.

Preuve
On note $m = \dim(F)$, on fixe une base quelconque $(e'_1, \dots, e'_m)$ de F, de sorte qu'on puisse écrire $f = \sum\limits_{k=1}^{m} f_k e'_k$, à comprendre au sens où, pour tout $x \in \Omega$, $f(x) = \sum\limits_{k=1}^{m} f_k(x) e'_k$.
Comme vu précédemment, on a $d(g \circ f)(a) = \big[dg\big(f(a)\big)\big] \circ df(a)$.
Or $\partial_i(g \circ f)(a) = d(g \circ f)(a) \cdot e_i$ donc $\partial_i(g \circ f)(a) = \big[dg\big(f(a)\big)\big] \circ df(a) \cdot e_i = \big[dg\big(f(a)\big)\big] \cdot \partial_i f(a)$.
Par linéarité de ∂_i, on a $\partial_i f(a) = \sum\limits_{k=1}^{m} \partial_i f_k(a) \cdot e'_k$ et, par linéarité de $dg\big(f(a)\big)$, on obtient : $\partial_i(g \circ f)(a) = \sum\limits_{k=1}^{m} \partial_i f_k(a) \cdot [dg\big(f(a)\big) \cdot e'_k] = \sum\limits_{k=1}^{m} \partial_i f_k(a) \partial_k g\big(f(a)\big)$.

Corollaire 34 (Règle de la chaîne pour une composée $\mathbb{R}^n \to \mathbb{R}^m \to \mathbb{R}$)
Soient $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, $g : \Omega' \subset \mathbb{R}^m \to \mathbb{R}$ avec $f(\Omega) \subset \Omega'$.
Notons $(f_1, \cdots, f_m)$ les composantes de f.
Si f est différentiable en $a \in \Omega$, si g est différentiable en $f(a)$, alors les dérivées partielles de

$$g \circ f : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ (x_1, \cdots, x_n) & \longmapsto & g\left(\underbrace{f_1(x_1, \cdots, x_n)}_{=y_1}, \cdots, \underbrace{f_m(x_1, \cdots, x_n)}_{=y_m} \right) \end{array} \right.$$

en $a = (a_1, \cdots, a_n)$ valent :

$$\forall i \in [\![1, n]\!], \qquad \frac{\partial(g \circ f)}{\partial x_i}(a) = \sum_{k=1}^{m} \frac{\partial f_k}{\partial x_i}(a) \times \frac{\partial g}{\partial y_k}(f(a)).$$

Preuve
Cas particulier du théorème précédent (les coordonnées sont exprimées dans les bases canoniques).

Exemple (Cas de fonctions à deux variables)
Soient $g : \mathbb{R}^2 \to \mathbb{R}$ et $f : \mathbb{R}^2 \to \mathbb{R}^2$ différentiables.
Introduisons les fonctions coordonnées u et v de f, de sorte qu'on écrive $f(x, y) = \big(u(x, y), v(x, y)\big)$.
La règle de la chaîne donne alors :

$$\frac{\partial(g \circ f)}{\partial x}(x, y) = \frac{\partial u}{\partial x}(x, y)\frac{\partial g}{\partial u}\big(u(x, y), v(x, y)\big) + \frac{\partial v}{\partial x}(x, y)\frac{\partial g}{\partial v}\big(u(x, y), v(x, y)\big),$$

$$\frac{\partial(g \circ f)}{\partial y}(x, y) = \frac{\partial u}{\partial y}(x, y)\frac{\partial g}{\partial u}\big(u(x, y), v(x, y)\big) + \frac{\partial v}{\partial y}(x, y)\frac{\partial g}{\partial v}\big(u(x, y), v(x, y)\big).$$

Exemple (Passage des coordonnées cartésiennes aux coordonnées polaires)

Soit $f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(x,y) \end{array} \right.$ une fonction de classe $\mathcal{C}^1$.
On pose, pour tout $(r,\theta) \in \mathbb{R}^2$,

$$F(r,\theta) = f(r\cos\theta, r\sin\theta)$$

(on dit que F est *l'expression de f en coordonnées polaires*). Alors $F \in \mathcal{C}^1(\mathbb{R}^2,\mathbb{R})$ et :

$$\frac{\partial F}{\partial r}(r,\theta) = \cos\theta \, \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta \, \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta),$$

$$\frac{\partial F}{\partial \theta}(r,\theta) = -r\sin\theta \, \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta \, \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta).$$

5) Formule d'intégration

Théorème 35 (Intégration le long d'un arc)
Soit $f : \Omega \to F$ de classe $\mathcal{C}^1$ et $\gamma : [0,1] \to \Omega$ un arc paramétré ("chemin") de classe $\mathcal{C}^1$, d'extrémités $\gamma(0) = a$ et $\gamma(1) = b$. Alors $f(b) - f(a) = \displaystyle\int_0^1 df\big(\gamma(t)\big) \cdot \gamma'(t)dt$.

Preuve
Soit $\varphi : \left\{ \begin{array}{ccc} [0,1] & \longrightarrow & F \\ t & \longmapsto & f\big(\gamma(t)\big) \end{array} \right.$.
Par composition, φ est dérivable, avec $\varphi'(t) = df\big(\gamma(t)\big) \cdot \gamma'(t) = \sum\limits_{i=1}^{n} x_i'(t) \dfrac{\partial f}{\partial x_i}\big(\gamma(t)\big)$, où on a choisi une base $(e_1, \ldots, e_n)$ quelconque de E, dans laquelle $\gamma = \sum\limits_{i=1}^{n} x_i e_i$.
Les x_i' sont continues ainsi que les $\dfrac{\partial f}{\partial x_i}$ vu que f est $\mathcal{C}^1$, donc φ' est bien continue.
Ainsi : $\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt$, c'est-à-dire $f(b) - f(a) = \displaystyle\int_0^1 df\big(\gamma(t)\big) \cdot \gamma'(t)dt$.

Exemple
Si le segment $[a,b]$ est inclus dans Ω, alors on peut le paramétrer par l'arc $\mathcal{C}^1$ donné par : $\gamma(t) = a + t(b-a)$, et on a $f(b) - f(a) = \displaystyle\int_0^1 df\big(a + t(b-a)\big) \cdot (b-a)dt$.

Corollaire 36 (Caractérisation des fonctions constantes)
Si Ω est un ouvert connexe par arcs et si $f : \Omega \to F$ est de classe $\mathcal{C}^1$, alors f est constante si et seulement si $df : \Omega \to \mathcal{L}(E,F)$ est identiquement nulle.

Preuve
$\boxed{\Longrightarrow}$ Si f est constante, alors sa différentielle en tout point est nulle, et donc df est identiquement nulle.

$\boxed{\Longleftarrow}$ Preuve uniquement dans le cas où Ω est convexe.

Soit $(a,b) \in \Omega^2$. Par convexité, $[a,b] \subset \Omega$ donc $f(b) - f(a) = \displaystyle\int_0^1 df\big(a + t(b-a)\big) \cdot (b-a)dt = \displaystyle\int_0^1 0_F dt = 0_F$, donc $f(b) = f(a)$. Ainsi f est constante.
(ça fonctionne aussi si Ω est étoilé, en fixant un point a qui "voit" tout point b en ligne droite).

Remarque
Attention, le cas Ω connexe par arcs n'est pas trivial ! Certes on peut relier deux points a et b de Ω par un chemin continu γ, mais on n'est pas sûr que γ soit $\mathcal{C}^1$, ce qui fait qu'on ne peut pas a priori appliquer la formule $f(b) - f(a) = \displaystyle\int_0^1 df\big(a + t(b-a)\big) \cdot (b-a)dt$.
On notera cependant que, Ω étant ouvert, il existe au voisinage de chaque point $a \in \Omega$ une boule $B_o(a,r) \subset \Omega$. Une telle boule étant convexe, f est constante en restriction à $B_o(a,r) \subset \Omega$. On dit que f est localement constante.

III Classe d'une fonction

1) Dérivées partielles successives

Définition 37 (Dérivées partielles d'ordre k)
Soit $f : \Omega \to F$ et $\mathcal{B}_E = (e_1, \dots, e_n)$ une base de E.
Par convention, on dit que f est la dérivée partielle d'ordre 0 de f.
*Pour $k \in \mathbb{N}$, lorsqu'elles existent, on appelle **dérivées partielles d'ordre** $k+1$ de f relativement à la base $\mathcal{B}_E$ les dérivées partielles des dérivées partielles d'ordre k de f.*

Notation
Si on note $x_1, \cdots, x_n$ les coordonnées dans la base $\mathcal{B}_E$ de la variable $x \in \Omega$, on note :

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} = \partial_{i_1}(\cdots(\partial_{i_k} f)\cdots).$$

Remarque
Si $\dim(E) = n$, alors si elles existent, les dérivées partielles d'ordre k de $f : \Omega \subset E \to F$ sont au nombre de n^k, et ce sont toutes des fonctions $\Omega \to F$ (comme f).

Exemple
Calculer les 4 dérivées partielles d'ordre 2 de $f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & x^2 y e^{-x} \end{array} \right.$
(relativement à la base canonique de $\mathbb{R}^2$).

2) Classe $\mathcal{C}^k$

Définition 38 (Fonction de classe $\mathcal{C}^k$)
Soit $k \in \mathbb{N}$. On dit que $f : \Omega \to F$ est **de classe $\mathcal{C}^k$** *lorsque ses dérivées partielles d'ordre k relativement à une base quelconque de E existent et sont continues.*
On dit que f est **de classe $\mathcal{C}^\infty$** *lorsqu'elle est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.*

Remarque
- *En réalité, pour $k \geq 2$, la notion de fonction de classe $\mathcal{C}^k$ peut se définir (comme pour la classe $\mathcal{C}^1$), au moyen d'un objet (très) abstrait, appelé différentielle d'ordre k (par exemple, la différentielle d'ordre 2 de $f : \Omega \to F$ en $a \in \Omega$, est, lorsqu'elle existe, $d^2 f(a) = d(df)(a) \in \mathcal{L}(E, \mathcal{L}(E, F))$ et on dit que f est de classe $\mathcal{C}^2$ lorsque $d^2 f : \Omega \to \mathcal{L}(E, \mathcal{L}(E, F))$ est continue), mais cette notion n'est pas au programme.*
- *Lorsque f est de classe $\mathcal{C}^k$, ses dérivées partielles d'ordre k dans* **toute** *base de E existent et sont continues (comme pour la classe $\mathcal{C}^1$, la notion est donc intrinsèque).*
- *Comme d'habitude une fonction est de classe $\mathcal{C}^0$ ssi elle est continue.*

Propriété 39 ($\mathcal{C}^{k+1} \subset \mathcal{C}^k$)
Toute fonction de classe $\mathcal{C}^{k+1}$ est de classe $\mathcal{C}^k$.

Preuve
Si f est de classe $\mathcal{C}^{k+1}$ alors en particulier ses dérivées partielles d'ordre k existent et sont de classe $\mathcal{C}^1$, donc continues.

Exemple
1. Les applications constantes sont de classe $\mathcal{C}^\infty$, et toutes leurs dérivées partielles d'ordre ≥ 1 sont nulles.
2. Les applications linéaires sont de classe $\mathcal{C}^\infty$: leurs dérivées partielles sont constantes donc toutes les dérivées partielles d'ordre ≥ 2 sont nulles.

Par conséquent les applications coordonnées, les parties réelles et imaginaires notamment sont de classe $\mathcal{C}^\infty$.

3) Opérations sur les applications de classe $\mathcal{C}^k$

Dans tout ce paragraphe, $k \in \mathbb{N} \cup \{+\infty\}$.

Théorème 40 (Structure d'espace vectoriel de $\mathcal{C}^k$.)
Soient $f, g : \Omega \to F$ et $(\lambda, \mu) \in \mathbb{R}^2$.
Si f et g sont de classe $\mathcal{C}^k$, alors $\lambda f + \mu g$ aussi.
Donc, l'ensemble $\mathcal{C}^k(\Omega, F)$ des applications de classe $\mathcal{C}^k$ de Ω vers F est un sous-espace vectoriel de F^Ω.

Preuve
Par récurrence, et utilisant

$$\partial_{i_1}(\cdots(\partial_{i_k}(\lambda f+\mu g)\cdots) = \partial_{i_1}(\cdots(\lambda\partial_{i_k} f+\mu\partial_{i_k} g)\cdots) = \cdots = \lambda\partial_{i_1}(\cdots(\partial_{i_k} f)\cdots)+\mu\partial_{i_1}(\cdots(\partial_{i_k} g)\cdots).$$

Théorème 41 (Composition bilinéaire de fonctions $\mathcal{C}^k$)
Soient $f : \Omega \to F$, $g : \Omega \to G$ et $B : F \times G \to H$ bilinéaire.
Si f et g sont de classe $\mathcal{C}^k$, alors $B(f, g)$ aussi.

Preuve
Récurrence encore.

Corollaire 42 (Structure de $\mathbb{R}$-algèbre des fonctions de classe $\mathcal{C}^k$)
Si F est une $\mathbb{R}$-algèbre, alors l'ensemble $\mathcal{C}^k(\Omega, F)$ est une sous-algèbre de F^Ω.

Preuve
Appliquer le théorème précédent à $B : (x, y) \mapsto xy$ qui est une application bilinéaire $F \times F \to F$.

Exemple
Les applications polynomiales $\mathbb{R}^n \to \mathbb{R}$ sont de classe $\mathcal{C}^\infty$.
L'application $\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est de classe $\mathcal{C}^\infty$car polynomiale en les fonctions coordonnées, qui sont elles-mêmes $\mathcal{C}^\infty$.

Théorème 43 (Composée de fonctions $\mathcal{C}^k$)
Soient $f : \Omega \to F$, $g : \Omega' \to G$, où Ω' est un ouvert de F tel que $f(\Omega) \subset \Omega'$.
Si f et g sont de classe $\mathcal{C}^k$, alors $g \circ f$ aussi.

Preuve
Via la formule donnant les dérivées partielles d'une composée.

Et on a enfin le critère classique et bien utile :

Théorème 44 (Classe $\mathcal{C}^k$ par composantes)
Soit $f : \Omega \to F$. Alors f est de classe $\mathcal{C}^k$ ssi les fonctions coordonnées de f dans une base quelconque de F sont de classe $\mathcal{C}^k$.

Preuve
Soit une base $(e'_1, \ldots, e'_m)$ de F. On écrit, pour tout $x \in \Omega$, $f(x) = \sum_{i=1}^{m} f_i(x)e'_i$, ce qui définit les fonctions coordonnées $(f_1, \ldots, f_m)$ de f relativement à la base choisie.
On a donc $f_i = p_i \circ f$, où p_i est la projection selon la i-ème coordonnée, donc est linéaire, donc $\mathcal{C}^\infty$.
Ainsi par composition f de classe $\mathcal{C}^k$ implique f_i de classe $\mathcal{C}^k$ pour tout i.

Et en notant encore e'_i la fonction constante égale à e'_i, qui est donc une fonction $\mathcal{C}^\infty$, l'écriture $f = \sum_{i=1}^{m} f_i e'_i$ montre que si les f_i sont de classe $\mathcal{C}^k$, alors f est $\mathcal{C}^k$ par combinaison linéaire de produits de fonctions $\mathcal{C}^k$.

Exemple

$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x,y) & \longmapsto & (x^2+y^2, xy) \end{array} \right.$ est $\mathcal{C}^\infty$ car ses fonctions coordonnées le sont.

$\varphi : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (r,\theta) & \longmapsto & (r\cos\theta, r\sin\theta) \end{array} \right.$ est $\mathcal{C}^\infty$ car ses fonctions coordonnées le sont.

4) Théorème de Schwarz

Théorème 45 (Théorème de Schwarz)
Si $f : \Omega \to F$ est de classe $\mathcal{C}^2$, alors les dérivées partielles secondes relativement à une quelconque base de E vérifient :

$$\forall (i,j) \in \{1,\ldots,n\}^2, \qquad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Remarque
Ainsi, l'ordre de dérivation suivant les différentes variables importe peu.

Preuve
Non exigible, voir les exercices pour une preuve dans le cas $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$.

Exemple
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^2$ et soit $g : \mathbb{R} \to \mathbb{R}$ définie par $g(t) = f(2t, 1+t^2)$. Calculer g''.

Exemple (Un contre-exemple si f n'est pas $\mathcal{C}^2$)

Soit $f : (x,y) \mapsto f(x,y) = \left\{ \begin{array}{ll} \dfrac{xy^3}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{array} \right.$.

1. *Montrer que f est continue en $(0,0)$.*
2. *Justifier brièvement que f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$, et même de classe $\mathcal{C}^2$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$.*
3. *Montrer que f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.*
4. *Calculer $\dfrac{\partial^2 f}{\partial x \partial y}(0,0)$ et $\dfrac{\partial^2 f}{\partial y \partial x}(0,0)$. Conclure.*