

CALCUL DIFFÉRENTIEL

I. Fonctions de classe $\mathcal{C}^1$.

Dans tout ce paragraphe, U est un ouvert de $\mathbb{R}^p$, et f est une application de U vers $\mathbb{R}$. L'espace vectoriel $\mathbb{R}^p$ sera toujours muni de sa norme euclidienne canonique notée $\| \cdot \|$.

1. Dérivée selon un vecteur.

Soit a un point de U , soit v un vecteur de $\mathbb{R}^p$. Comme U est ouvert, il existe un réel $\delta > 0$ tel que $\forall t \in [-\delta, \delta] \quad a + tv \in U$.

En effet, le point a est intérieur à U donc il existe $r > 0$ tel que la boule fermée $\overline{B}(a, r) = \{x \in \mathbb{R}^p \mid \|x - a\| \leq r\}$ soit incluse dans U . On a donc $a + tv \in U$ dès que $|t| \leq \frac{r}{\|v\|}$ (si v est non nul, et s'il est nul c'est évident!).

L'application $\varphi_v : t \mapsto f(a + tv)$ est donc définie au moins sur l'intervalle $[-\delta, \delta]$.

Définition. Si l'application $\varphi_v : t \mapsto f(a + tv)$ est dérivable en 0, on dit que f admet une **dérivée au point a selon le vecteur v** , et on pose

$$D_v f(a) = \varphi'_v(0) .$$

Ainsi, sous réserve d'existence,

$$D_v f(a) = \varphi'_v(0) = \left[\frac{d}{dt} (f(a + tv)) \right]_{t=0} = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} .$$

Interprétation. Lorsque t décrit $\mathbb{R}$, le point $m = a + tv$ décrit, dans $\mathbb{R}^p$, la droite affine passant par a et de vecteur directeur v (si ce vecteur est non nul). La dérivée de f en a selon v donne une idée de comment varie $f(m)$ lorsque le point m s'éloigne du point a dans la direction du vecteur v . Si cette dérivée directionnelle existe, on a en effet le développement limité à l'ordre un:

$$\varphi_v(t) = f(a + tv) = f(a) + D_v f(a) \cdot t + o(t) .$$

2. Dérivées partielles.

Définition 1. Soit $a = (a_1, \dots, a_p)$ un point de U , soit $i \in \llbracket 1, p \rrbracket$. L'application

$$\varphi_i : x \mapsto \varphi_i(x) = f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_p)$$

est la i -ième **application partielle** de f au point a .

Commentaire. Cette application est définie sur une partie ouverte (on l'admettra) de $\mathbb{R}$, à savoir l'ensemble des réels x tels que le p -uplet $(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_p)$ appartienne à U . On a, en particulier, $\varphi_i(a_i) = f(a)$.

Définition 2. Avec les notations introduites ci-dessus, si cette application partielle φ_i est dérivable au point a_i , on dit que f admet au point a une **dérivée partielle d'ordre 1** par rapport à la i -ième variable. On note alors

$$\partial_i f(a) = \varphi'_i(a_i) , \quad \text{ou encore} \quad \frac{\partial f}{\partial x_i}(a) = \varphi'_i(a_i) .$$

Commentaire. Les dérivées partielles sont donc les dérivées des applications partielles (qui sont, elles, des fonctions "ordinaires" d'une variable réelle à valeurs réelles).

Commentaire. En revenant à la définition de la dérivée d'une fonction d'une variable, et en faisant éventuellement une translation de la variable, on a ainsi, sous réserve d'existence,

$$\begin{aligned} \partial_i f(a) &= \left[\frac{d}{dx} f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_p) \right]_{x=a_i} \\ &= \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_p) - f(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_p)}{t} \end{aligned}$$

$$= \lim_{t \rightarrow 0} \frac{f(a + t e_i) - f(a)}{t} = D_{e_i} f(a),$$

en notant $(e_1, \dots, e_p)$ la base canonique de $\mathbb{R}^p$. Autrement dit, cette dérivée partielle de f en a par rapport à la i -ème variable est aussi la dérivée de f en a selon le i -ème vecteur e_i de la base canonique de $\mathbb{R}^p$.

Remarque. L'existence de dérivées partielles n'entraîne pas la continuité. Soit en effet la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = \frac{xy}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$.

Cette fonction est discontinue à l'origine puisque, pour x non nul, $f(x, x) = \frac{1}{2} \neq f(0, 0) = 0$. Elle admet pourtant des dérivées partielles en tout point du plan: en effet, par opérations algébriques (produit, quotient avec dénominateur non nul), on obtient, pour $(x, y) \neq (0, 0)$, $\frac{\partial f}{\partial x}(x, y) = \frac{y(y^2 - x^2)}{(x^2 + y^2)^2}$ et $\frac{\partial f}{\partial y}(x, y) = \frac{x(x^2 - y^2)}{(x^2 + y^2)^2}$. Par ailleurs, les applications partielles $x \mapsto f(x, 0)$ et $y \mapsto f(0, y)$ sont nulles sur $\mathbb{R}$ tout entier, d'où l'existence de dérivées partielles à l'origine avec $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$.

3. Fonctions de classe $\mathcal{C}^1$.

Définition. Soit U un ouvert de $\mathbb{R}^p$. Une fonction $f : U \rightarrow \mathbb{R}$ est dite **de classe $\mathcal{C}^1$** sur U si elle admet, en tout point de U , des dérivées partielles d'ordre 1, et si les p applications dérivées partielles $\partial_1 f, \dots, \partial_p f$ sont continues sur U .

Opérations. On montre sans difficulté que, si $f : U \rightarrow \mathbb{R}$ et $g : U \rightarrow \mathbb{R}$ sont de classe $\mathcal{C}^1$ sur U , alors il en est de même de $\alpha f + g$, où α est un réel. L'ensemble $\mathcal{C}^1(U, \mathbb{R})$ des fonctions de classe $\mathcal{C}^1$ sur U est alors muni d'une structure de $\mathbb{R}$ -espace vectoriel. On a, de plus, pour tout $i \in \llbracket 1, p \rrbracket$, la relation

$$\partial_i(\alpha f + g) = \alpha \partial_i f + \partial_i g.$$

Avec les mêmes hypothèses, fg est de classe $\mathcal{C}^1$ sur U avec $\partial_i(fg) = f \cdot \partial_i g + g \cdot \partial_i f$, ou encore avec d'autres notations, $\frac{\partial(fg)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) \cdot g(a) + f(a) \cdot \frac{\partial g}{\partial x_i}(a)$.

4. Développements limités d'ordre 1.

Définition. Soit $f : U \rightarrow \mathbb{R}$ avec U un ouvert de $\mathbb{R}^p$, soit $a \in U$. On dit que f admet un **développement limité d'ordre 1** au point a (en abrégé un $DL_1(a)$) s'il existe une forme linéaire λ sur $\mathbb{R}^p$ telle que

$$f(a + h) = f(a) + \lambda(h) + o(\|h\|)$$

lorsque $h \in \mathbb{R}^p$ tend vers 0.

On peut écrire cela sous la forme

$$f(a + h) = f(a) + \lambda(h) + h \varepsilon(h), \quad \text{avec} \quad \lim_{h \rightarrow 0} \varepsilon(h) = 0.$$

L'existence d'un $DL_1(a)$ correspond à la notion de **différentiabilité** (*terme hors programme*) de la fonction f au point a .

Un exemple élémentaire. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = xy^2 - y$ (c'est une fonction "polynomiale" à deux variables), soit $a = (1, 3) \in \mathbb{R}^2$. Par un simple calcul

algébrique, on peut obtenir un $DL_1(a)$ de la fonction f . Soit un vecteur “petit déplacement” $h = (u, v) \in \mathbb{R}^2$, alors

$$\begin{aligned} f(a+h) &= f(1+u, 3+v) \\ &= (1+u)(3+v)^2 - (3+v) \\ &= 6+9u+5v+6uv+v^2+uv^2 \\ &= 6+(9u+5v)+o(\|h\|) . \end{aligned}$$

En effet, on vérifie facilement que, lorsque $h = (u, v) \rightarrow (0, 0)$, on a $uv = o(\|h\|)$ et $uv^2 = o(\|h\|)$, par exemple en choisissant $\|h\| = \|h\|_\infty = \max\{|u|, |v|\}$ (mais peu importe puisqu’en dimension finie, les normes sont toutes équivalentes). L’application $\lambda : h = (u, v) \mapsto 9u + 5v$ est une forme linéaire sur $\mathbb{R}^2$, canoniquement représentée par la matrice-ligne $L = (9 \ 5)$. Enfin, $f(a) = 6$. On a donc obtenu un développement limité à l’ordre un de f au voisinage du point $a = (1, 3)$. On peut remarquer que les coefficients de ce développement limité sont les dérivées partielles premières de f au point a , en effet $\frac{\partial f}{\partial x}(a) = 9$ et $\frac{\partial f}{\partial y}(a) = 5$.

Propriété d’unicité. Si f admet un développement limité d’ordre 1 en un point a de U , la forme linéaire λ apparaissant dans l’écriture de ce développement est unique.

Preuve. Supposons que $f(a+h) = f(a) + \lambda(h) + o(\|h\|)$ et $f(a+h) = f(a) + \mu(h) + o(\|h\|)$ lorsque h tend vers 0 dans $\mathbb{R}^p$, où λ et μ sont deux formes linéaires sur $\mathbb{R}^p$. Par différence, en posant $\alpha = \lambda - \mu$ (c’est une nouvelle forme linéaire sur $\mathbb{R}^p$), on a $\alpha(h) = o(\|h\|)$

lorsque $h \rightarrow 0$, ce qui signifie que $\lim_{h \rightarrow 0} \frac{|\alpha(h)|}{\|h\|} = 0$. Si x est un vecteur non nul de $\mathbb{R}^p$,

alors en posant $u_n = \frac{x}{n}$ pour $n \in \mathbb{N}^*$, on a $\lim_{n \rightarrow +\infty} u_n = 0$ donc $\lim_{n \rightarrow +\infty} \frac{|\alpha(u_n)|}{\|u_n\|} = 0$ par

caractérisation séquentielle de la limite. Mais $\frac{|\alpha(u_n)|}{\|u_n\|}$ ne dépend pas de l’entier n et vaut

$\frac{|\alpha(x)|}{\|x\|}$, donc $|\alpha(x)| = 0$ pour tout x , donc $\alpha = 0$ et $\lambda = \mu$.

Une forme linéaire λ sur $\mathbb{R}^p$ est canoniquement représentée par une matrice-ligne $L = (\alpha_1 \ \cdots \ \alpha_p) \in \mathcal{M}_{1,p}(\mathbb{R})$, c’est alors l’application de $\mathbb{R}^p$ vers $\mathbb{R}$ telle que

$$\forall h = (h_1, \dots, h_p) \in \mathbb{R}^p \quad \lambda(h) = \alpha_1 h_1 + \cdots + \alpha_p h_p = \sum_{i=1}^p \alpha_i h_i .$$

L’existence d’un développement limité d’ordre 1 au point $a = (a_1, \dots, a_p) \in U$ pour une fonction $f : U \rightarrow \mathbb{R}$ équivaut donc à l’existence de p réels $\alpha_1, \dots, \alpha_p$ tels que

$$f(a+h) = f(a_1+h_1, \dots, a_p+h_p) = f(a) + \sum_{i=1}^p \alpha_i h_i + o(\|h\|)$$

lorsque $h = (h_1, \dots, h_p) \rightarrow (0, \dots, 0)$ dans $\mathbb{R}^p$. Les réels α_i (alors uniques) sont les **coefficients** de ce développement limité.

5. Différentielle en un point d'une fonction de classe $\mathcal{C}^1$.

On admettra le théorème suivant:

Théorème. Toute fonction de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^p$ admet, en tout point a de U , un développement limité d'ordre 1, dont les coefficients sont les dérivées partielles d'ordre 1 de f au point a , i.e.

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

lorsque $h = (h_1, \dots, h_p) \rightarrow (0, \dots, 0)$ dans $\mathbb{R}^p$.

Définition. La “partie linéaire” de ce développement limité, i.e. la forme linéaire sur $\mathbb{R}^p$ définie par

$$h = (h_1, \dots, h_p) \mapsto \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) h_i = \sum_{i=1}^p h_i \partial_i f(a)$$

est appelée **différentielle** de f en a , et notée $df(a)$.

Notation. Si f est de classe $\mathcal{C}^1$, si $a \in U$, si $h = (h_1, \dots, h_p) \in \mathbb{R}^p$, le réel $\sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) h_i$, qu'il serait logique de noter $df(a)(h)$, est couramment noté $df(a) \cdot h$. Ainsi, le $DL_1(a)$ d'une fonction f de classe $\mathcal{C}^1$ se mettra sous la forme

$$(*) \quad f(a+h) = f(a) + df(a) \cdot h + o(\|h\|).$$

Remarque. Le nombre $df(a) \cdot h = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) h_i$ est aussi la dérivée de f au point a selon

le vecteur h , que nous avons notée précédemment $D_h f(a)$. En effet, il résulte de $(*)$ que, si le vecteur h est fixé, l'application de variable réelle $\varphi_h : t \mapsto f(a+th)$ admet en 0 un développement limité d'ordre un, qui est $\varphi_h(t) = f(a) + t df(a) \cdot h + o(t)$, ce qui montre que φ_h est dérivable en 0, avec par définition $D_h f(a) = \varphi'_h(0) = df(a) \cdot h$.

Conséquence. Toute fonction de classe $\mathcal{C}^1$ sur U est continue sur U .

Preuve. En effet, du $DL_1(a)$ ci-dessus, il résulte immédiatement (notamment en utilisant la continuité des applications linéaires en dimension finie) que $\lim_{h \rightarrow 0} f(a+h) = f(a)$.

Remarque. Approximation du premier ordre. Si f est de classe $\mathcal{C}^1$ sur U , le vecteur $h = (h_1, \dots, h_p)$ intervenant dans l'écriture du $DL_1(a)$ de f est souvent interprété comme un vecteur “petit déplacement”, on le note parfois $da = (dx_1, \dots, dx_p)$. Le $DL_1(a)$ s'écrit alors

$$f(a+da) = f(a) + \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) dx_i + o(\|da\|).$$

On dit parfois que, “dans une approximation du premier ordre”, on a

$$f(a+da) \simeq f(a) + \sum_{i=1}^p \frac{\partial f}{\partial x_i}(a) dx_i.$$

II. Règle de la chaîne.

1. Dérivée le long d'un arc.

Soit U un ouvert de $\mathbb{R}^p$, soit $f : U \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$. On peut interpréter f comme un **champ de scalaires** sur U .

Soit d'autre part I un intervalle de $\mathbb{R}$ et $x_1, \dots, x_p$ des fonctions de classe $\mathcal{C}^1$ de I vers $\mathbb{R}$ telles que

$$\forall t \in I \quad (x_1(t), \dots, x_p(t)) \in U.$$

Si, pour $t \in I$, on pose $\varphi(t) = (x_1(t), \dots, x_p(t))$, alors la fonction vectorielle $\varphi : I \rightarrow U \subset \mathbb{R}^p$ peut être interprétée comme un **arc paramétré** de classe $\mathcal{C}^1$, à support dans U .

Pour tout $t \in I$, posons enfin $g(t) = f(\varphi(t)) = f(x_1(t), \dots, x_p(t))$. On a ainsi $g = f \circ \varphi$, qui est une application de I vers $\mathbb{R}$.

Théorème. Alors g est de classe $\mathcal{C}^1$ sur I et on a

$$\forall t \in I \quad g'(t) = \sum_{i=1}^p x'_i(t) \partial_i f(x_1(t), \dots, x_p(t)) = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(\varphi(t)) x'_i(t).$$

On peut réécrire cela sous la forme

$$\frac{d}{dt} \left(f(x_1(t), \dots, x_p(t)) \right) = \sum_{i=1}^p x'_i(t) \frac{\partial f}{\partial x_i}(x_1(t), \dots, x_p(t)),$$

ou encore

$$\frac{d}{dt} \left(f(\varphi(t)) \right) = df(\varphi(t)) \cdot \varphi'(t).$$

Preuve. Pour $t \in I$ et h réel “petit” (en tous cas tel que $t+h \in I$), on a

$$g(t+h) = f(\varphi(t+h)) = f(\varphi(t) + h \varphi'(t) + h \varepsilon(h)) \quad \text{avec} \quad \lim_{h \rightarrow 0} \varepsilon(h) = 0,$$

en utilisant le développement limité d'ordre 1 de la fonction vectorielle φ au point t . Mais f admet aussi un DL_1 au point $\varphi(t)$, qui s'écrit $f(\varphi(t)+k) = f(\varphi(t)) + df(\varphi(t)) \cdot k + o(\|k\|)$, où la lettre k représente ici un vecteur “petit déplacement” dans $\mathbb{R}^p$. Comme $\lim_{h \rightarrow 0} (h \varphi'(t) + h \varepsilon(h)) = 0$, il est légitime de remplacer k par cette expression dans ce dernier DL , on a donc

$$\begin{aligned} g(t+h) &= f(\varphi(t) + h \varphi'(t) + h \varepsilon(h)) \\ &= f(\varphi(t)) + df(\varphi(t)) \cdot (h \varphi'(t) + h \varepsilon(h)) + o(\|k\|) \\ &= g(t) + h df(\varphi(t)) \cdot \varphi'(t) + \text{des cacahuètes}. \end{aligned}$$

On a utilisé la linéarité de la différentielle $df(\varphi(t))$ pour scinder le terme du milieu en deux termes. Le lecteur méticuleux s'assurera que les termes restants non explicités et relégués au rang de “cacahuètes” sont bien “négligeables dans une approximation du premier ordre”, autrement dit qu'ils peuvent s'écrire sous la forme $o(h)$.

On vient de prouver que la fonction $g : I \rightarrow \mathbb{R}$ admet en tout point de l'intervalle I un développement limité à l'ordre 1, elle est donc dérivable sur I . Sa dérivée au point

$t \in I$ est le coefficient du terme du premier ordre, ce qui donne bien la formule attendue:

$$\forall t \in I \quad g'(t) = \mathrm{d}f(\varphi(t)) \cdot \varphi'(t) = \sum_{i=1}^p x'_i(t) \frac{\partial f}{\partial x_i}(\varphi(t)).$$

Enfin, comme f est de classe $\mathcal{C}^1$, ses dérivées partielles $\frac{\partial f}{\partial x_i}$ sont continues, on voit alors que g' est continue comme composée, produit et somme de fonctions continues. Donc g est de classe $\mathcal{C}^1$ sur I .

2. Application aux dérivées partielles d'une fonction composée.

Soient U un ouvert de $\mathbb{R}^n$, Ω un ouvert de $\mathbb{R}^p$, soit $X : U \rightarrow \Omega$ une application (un "changement de variables") de classe $\mathcal{C}^1$: pour tout $u = (u_1, \dots, u_n) \in U$, posons

$$X(u) = (x_1(u_1, \dots, u_n), \dots, x_p(u_1, \dots, u_n)) ,$$

dire que l'application X est de classe $\mathcal{C}^1$ signifie que ses p fonctions coordonnées $x_1, \dots, x_p$ sont de classe $\mathcal{C}^1$ sur U . Soit d'autre part $f : \Omega \rightarrow \mathbb{R}$ une application de classe $\mathcal{C}^1$. Soit enfin l'application

$$g = f \circ X : \begin{cases} U & \rightarrow \mathbb{R} \\ u = (u_1, \dots, u_n) & \mapsto g(u) = f(x_1(u_1, \dots, u_n), \dots, x_p(u_1, \dots, u_n)) \end{cases} .$$

Alors g est de classe $\mathcal{C}^1$ sur U , et ses dérivées partielles se calculent par les formules

$$\boxed{\forall i \in \llbracket 1, n \rrbracket \quad \frac{\partial g}{\partial u_i} = \sum_{j=1}^p \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial u_i} .}$$

La preuve est identique à celle du paragraphe précédent.

Notons que la relation encadrée est un peu abrégée: en effet, elle n'indique pas en quel(s) point(s) les différentes dérivées partielles sont évaluées. Cette formule, complètement explicitée, s'écrirait

$$\forall i \in \llbracket 1, n \rrbracket \quad \forall u \in U \quad \frac{\partial g}{\partial u_i}(u) = \sum_{j=1}^p \frac{\partial f}{\partial x_j}(X(u)) \frac{\partial x_j}{\partial u_i}(u) .$$

Un cas particulier est à connaître, c'est celui du "**passage en coordonnées polaires**":

soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$. Posons d'autre part $x(r, \theta) = r \cos \theta$ et $y(r, \theta) = r \sin \theta$ pour tout couple $(r, \theta) \in \mathbb{R}^2$. Soit enfin $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$\forall (r, \theta) \in \mathbb{R}^2 \quad g(r, \theta) = f(r \cos \theta, r \sin \theta) = f(x(r, \theta), y(r, \theta)) .$$

Alors g est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et on a les relations

$$\begin{aligned} \frac{\partial g}{\partial r} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = (\cos \theta) \frac{\partial f}{\partial x} + (\sin \theta) \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial \theta} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = (-r \sin \theta) \frac{\partial f}{\partial x} + (r \cos \theta) \frac{\partial f}{\partial y} . \end{aligned}$$

3. Caractérisation des fonctions constantes.

Proposition. Soit U un ouvert convexe de $\mathbb{R}^p$, soit $f : U \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$. Alors f est constante sur U si et seulement si $\forall a \in U \quad df(a) = 0$.

La condition énoncée est que, en tout point a de U , la différentielle de f en a est la forme linéaire nulle sur $\mathbb{R}^p$. Comme cette différentielle s'exprime à l'aide des dérivées partielles d'ordre 1 de f , un énoncé équivalent est de dire (toujours si f est de classe $\mathcal{C}^1$ sur un ouvert convexe U) que f est constante sur U si et seulement si

$$\forall a \in U \quad \forall i \in \llbracket 1, p \rrbracket \quad \frac{\partial f}{\partial x_i}(a) = 0.$$

Preuve. Le sens direct est évident: si f est constante sur U , toutes ses dérivées partielles sont nulles en tout point de U .

Réciproquement, si $df = 0$ sur U , soient $a \in U$ et $b \in U$, montrons que $f(a) = f(b)$. Comme U est convexe, le segment $[a, b]$ est inclus dans U , i.e. $\forall t \in [0, 1] \quad (1-t)a + tb \in U$. Posons $x(t) = (1-t)a + tb$ pour tout $t \in [0, 1]$, alors la fonction vectorielle x est de classe $\mathcal{C}^1$ sur $[0, 1]$, à valeurs dans U , avec $x'(t) = b - a$ pour tout $t \in [0, 1]$. Soit maintenant la fonction composée $g = f \circ x$, à savoir $g : [0, 1] \rightarrow \mathbb{R}$ telle que $g(t) = f((1-t)a + tb)$ pour tout $t \in [0, 1]$. D'après la règle de la chaîne (dérivée le long d'un "arc" qui est ici un segment de droite), cette fonction g est de classe $\mathcal{C}^1$ sur $[0, 1]$ avec $\forall t \in [0, 1] \quad g'(t) = df(x(t)) \cdot x'(t) = 0$. Cette fonction g d'une variable, dérivable et de dérivée nulle, est donc constante sur $[0, 1]$, d'où $g(0) = g(1)$, i.e. $f(a) = f(b)$.

III. Gradient.

Définition. Soit $f : U \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$, où U est un ouvert de $\mathbb{R}^p$. On munit $\mathbb{R}^p$ de sa structure euclidienne canonique. Pour tout $a \in U$, la différentielle de f en a , notée $df(a)$, est une forme linéaire sur $\mathbb{R}^p$. D'après le théorème de représentation de Riesz, il existe alors un unique vecteur v de $\mathbb{R}^p$ tel que

$$\forall h \in \mathbb{R}^p \quad df(a) \cdot h = (v|h).$$

Ce vecteur v est appelé **gradient** de f en a , et noté $\nabla f(a)$.

On a ainsi $\forall h \in \mathbb{R}^p \quad df(a) \cdot h = (\nabla f(a) | h)$.

Comme d'autre part, si $h = (h_1, \dots, h_p) = \sum_{i=1}^p h_i e_i$, on a $df(a) \cdot h = \sum_{i=1}^p \partial_i f(a) h_i$, on déduit que les coordonnées du vecteur gradient $\nabla f(a)$, dans la base canonique $(e_1, \dots, e_p)$ de $\mathbb{R}^p$, sont les dérivées partielles premières de f au point a , i.e.

$$\nabla f(a) = \sum_{i=1}^p \partial_i f(a) e_i = (\partial_1 f(a), \dots, \partial_p f(a)).$$

On peut alors réécrire différents résultats des paragraphes précédents:

- si $f : U \rightarrow \mathbb{R}$ est de classe $\mathcal{C}^1$, alors son développement limité d'ordre 1 en $a \in U$ s'écrit

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

La dérivée de f au point a selon le vecteur h est donc aussi $D_h f(a) = (\nabla f(a) | h)$.

- “dérivée le long d’un arc”: si, de plus, on se donne un arc paramétré de classe $\mathcal{C}^1$ à support dans U , i.e. une application $\varphi : I \rightarrow U$, $t \mapsto m = \varphi(t)$, de classe $\mathcal{C}^1$, et si on pose $g(t) = f(\varphi(t))$ pour tout $t \in I$, alors g est de classe $\mathcal{C}^1$ sur l’intervalle I avec

$$\forall t \in I \quad g'(t) = \left(\nabla f(\varphi(t)) \mid \varphi'(t) \right) = \left(\nabla f(m) \mid \frac{dm}{dt} \right),$$

i.e. $g'(t)$ est le produit scalaire du vecteur gradient de f au point $m = \varphi(t)$ avec le vecteur vitesse instantanée du point mobile m à l’instant t , dans une interprétation cinématique.

- si U est un ouvert convexe de $\mathbb{R}^p$ et si $f : U \rightarrow \mathbb{R}$ est de classe $\mathcal{C}^1$, alors f est constante sur U si et seulement si $\forall a \in U \quad \nabla f(a) = \vec{0}$.

Expression du gradient en coordonnées polaires.

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$. Posons $g(r, \theta) = f(r \cos \theta, r \sin \theta)$ pour $(r, \theta) \in \mathbb{R}^2$. Alors g est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$, et on a obtenu dans le paragraphe **II.2.** les relations

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

En inversant ce système linéaire, dont le déterminant est $\begin{vmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{vmatrix} = r$, non nul si on se place en un point autre que l’origine, on obtient

$$\begin{cases} \frac{\partial f}{\partial x} = \cos \theta \frac{\partial g}{\partial r} - \frac{\sin \theta}{r} \frac{\partial g}{\partial \theta} \\ \frac{\partial f}{\partial y} = \sin \theta \frac{\partial g}{\partial r} + \frac{\cos \theta}{r} \frac{\partial g}{\partial \theta} \end{cases}.$$

On a alors, en notant (e_1, e_2) la base canonique de $\mathbb{R}^2$,

$$\begin{aligned} \nabla f &= \frac{\partial f}{\partial x} e_1 + \frac{\partial f}{\partial y} e_2 \\ &= \frac{\partial g}{\partial r} (\cos \theta e_1 + \sin \theta e_2) + \frac{1}{r} \frac{\partial g}{\partial \theta} (-\sin \theta e_1 + \cos \theta e_2) \\ &= \frac{\partial g}{\partial r} u_r + \frac{1}{r} \frac{\partial g}{\partial \theta} u_\theta, \end{aligned}$$

en notant (u_r, u_θ) la “base mobile” des coordonnées polaires, i.e. la base orthonormale directe obtenue par rotation d’angle θ à partir de la base canonique.