Calcul différentiel.

Chap. 14 : cours complet.

1. Fonctions de classe C^1 de $\mathbb{R}^p$ dans $\mathbb{R}$.

Théorème 1.1 et définition 1.1 : dérivées partielles d'une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$ en un point
Définition 1.2 : fonction de classe C^1 sur un ouvert de $\mathbb{R}^p$
Théorème 1.2 : existence, unicité d'un développement limité en un point pour une fonction de classe C^1
Définition 1.3 : différentielle d'une fonction en un point
Théorème 1.3 : classe C^1 et continuité
Théorème 1.4 : opérations sur les fonctions de classe C^1
Théorème 1.5 : règle de la chaîne (dérivée le long d'un arc paramétré)
Théorème 1.6 : caractérisation des fonctions constantes sur un ouvert convexe

2. Gradient.

Définition 2.1 et théorème 2.1 : gradient d'une fonction de classe C^1, expression de la différentielle
Théorème 2.1 : dérivées partielles et changement de variables dans $\mathbb{R}^2$
Exemple 2.1 : changement de coordonnées dans $\mathbb{R}^2$, exemple du changement polaires-cartésiennes
Exemple 2.2 : expression du gradient en coordonnées polaires

3. Dérivées d'ordre 2, étude d'extrema de fonctions à valeurs réelles.

Définition 3.1 : dérivées partielles d'ordre 2 d'une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$ en un point
Définition 3.2 : fonction de classe C^2 de $\mathbb{R}^p$ dans $\mathbb{R}$
Théorème 3.1 : de Schwarz
Définition 3.3 : extremum d'une fonction à valeurs réelles
Définition 3.4 : point critique d'une fonction de classe C^1 à valeurs réelles
Théorème 3.2 : condition nécessaire d'extremum pour une fonction de classe C^1 sur un ouvert

Calcul différentiel.

Dans tout le chapitre, on considèrera des fonctions de $\mathbb{R}^p$ dans $\mathbb{R}$, et en pratique, p vaudra 2 ou 3.
Au besoin, on notera : $\mathscr{B} = (e_1, \dots, e_p)$, la base canonique de $\mathbb{R}^p$.
Enfin, on utilisera la norme $\|\,.\,\|_\infty$ dans $\mathbb{R}^p$ et on admettra au besoin que toutes les notions développées dans le chapitre ne dépendent pas de la norme choisie.

1. Fonctions de classe C^1 de $\mathbb{R}^p$ dans $\mathbb{R}$.

Théorème 1.1 et définition 1.1 : dérivées partielles d'une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$ en un point

Soit f une fonction définie d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$, et soit : $a \in U$.
Alors il existe une boule ouverte centrée en a sur laquelle f est définie.
Pour : $1 \leq i \leq p$, la fonction de $\mathbb{R}$ dans $\mathbb{R}$ donnée par : $t \mapsto f(a+t.e_i)$, est alors définie sur un intervalle ouvert autour de 0.

On dit alors que f admet une $i^{\text{ème}}$ dérivée partielle en a si et seulement si $\lim\limits_{t\to 0}\frac{1}{t}[f(a+t.e_i)-f(a)]$ existe.

On note cette limite : $\lim\limits_{t\to 0}\frac{1}{t}[f(a+t.e_i)-f(a)] = \frac{\partial f}{\partial x_i}(a) = \partial_i f(a)$.

En écrivant : $a = (a_1, \dots, a_p)$, cette limite quand elle existe correspond à la dérivée a_i de la fonction de $\mathbb{R}$ dans $\mathbb{R}$ (appelée partielle $i^{\text{ème}}$ fonction partielle en a) définie par : $t \mapsto f(a_1,\dots,a_{i-1},t,a_{i+1},\dots,a_p)$.

Démonstration :

- Puisque U est ouvert, pour la norme $\|\,.\,\|_\infty$ dans $\mathbb{R}^p$, il existe : $\delta > 0$, $B(a,\delta) \subset U$.

On a alors : $\forall\ 1 \leq i \leq p$, $\forall\ t \in\ ]-\delta,+\delta[$, $\|(a+t.e_i)-a\|_\infty = |t|.\|e_i\|_\infty < \delta.1 = \delta$, et : $(a+t.e_i) \in U$.

- D'autre part on a : $\forall\ 1 \leq i \leq p$, $\forall\ t \in\ ]-\delta,+\delta[$, $\frac{1}{t}.[f(a+t.e_i)-f(a)] = \frac{f_i(h)-f_i(a_i)}{h-a_i}$, où on a noté :

$h = t + a_i$, et : $f_i(h) = f(a_1,\dots,a_{i-1},h,a_{i+1},\dots,a_p)$,

et la quantité précédente admet bien une limite en 0 si et seulement si f_i est dérivable en a_i (et dans ce cas, la limite en 0 correspond bien à la dérivée en a_i de f_i).

Définition 1.2 : fonction de classe C^1 sur un ouvert de $\mathbb{R}^p$

Soit f une fonction définie d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$.
On dit que f est de classe C^1 sur U si et seulement si f admet des dérivées partielles par rapport à toutes ses variables en tout point de U et si ces p fonctions sont toutes continues sur U.

Théorème 1.2 : existence et unicité d'un développement limité en un point pour une fonction de classe C^1

Soit f une fonction définie d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$.
Si f est de classe C^1 sur U, alors la fonction : $h \mapsto f(a+h)$, est définie sur une boule ouverte B_a autour de 0 et :

$$\forall\ a \in U,\ \forall\ h \in B_a,\ f(a+h) = f(a) + \sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a) + \|h\|_\infty.\varepsilon(h) = f(a) + \sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a) + o(h),$$

où on a noté : $h = \sum\limits_{i=1}^{p} h_i.e_i = (h_1,\dots,h_p)$, et ou : $\lim\limits_{h\to 0}\varepsilon(h) = 0$.

Ce développement limité à l'ordre 1 est unique, autrement dit :

si : $\exists\ (\alpha_1, \dots, \alpha_p) \in \mathbb{R}^p$, tel que : $\forall\ a \in U$, $\forall\ h \in B_a$, $f(a+h) = f(a) + \sum\limits_{i=1}^{p}\alpha_i.h_i + o(h)$,

alors : $\forall\ 1 \leq i \leq p$, $\alpha_i = \frac{\partial f}{\partial x_i}(a)$.

Démonstration : (admise)

• On sait qu'il existe : δ > 0, tel que : ∀ h ∈ $\mathbb{R}^p$, ($\|h\|_\infty < \delta$) ⇒ ((a + h) ∈ U), comme on l'a vu plus haut et donc la fonction : h ↦ $f(a+h)$, est définie au moins que la boule ouverte : $B_a = B_\infty(0,\delta)$.

Soit : $h = \sum_{i=1}^{p} h_i.e_i = (h_1,...,h_p)$, tel que : $0 < \|h\|_\infty < \delta$.

On pose alors : $\varepsilon(h) = \frac{1}{\|h\|}.\left[f(a+h) - f(a) - \sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a) \right]$.

On va passer de a à (a+h) en ajoutant une par une toutes les composantes supplémentaires de h, soit :

$$\varepsilon(h) = \frac{1}{\|h\|}.\left[\sum_{i=1}^{p} [\, f(a+h_1.e_1+...+h_i.e_i) - f(a+h_1.e_1+...+h_{i-1}.e_{i-1})] - \sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a) \right],$$

autrement dit en utilisant une somme télescopique.

On applique alors le théorème des accroissements finis aux p fonctions d'une seule variable :

∀ 1 ≤ i ≤ p, φ_i : t ↦ $f(a+h_1.e_1+...+t.e_i)$, sur les segments $[0,h_i]$,

et on peut écrire :

∀ 1 ≤ i ≤ n, ∃ θ_i ∈]0,1[, $\varphi_i(h_i) - \varphi_i(0) = h_i.\varphi_i'(\theta_i.h_i) = h_i.\frac{\partial f}{\partial x_i}(a+h_1.e_1+...+h_{i-1}.e_{i-1}+\theta_i.h_i.e_i)$.

Donc : ∃ $(\theta_1, \ldots, \theta_p)$ ∈ $]0,1[^p$, $\varepsilon(h) = \sum_{i=1}^{p} \frac{h_i}{\|h\|}.\left[\frac{\partial f}{\partial x_i}(a+h_1.e_1+...+h_{i-1}.e_{i-1}+\theta_i.h_i.e_i) - \frac{\partial f}{\partial x_i}(a) \right]$.

Il suffit alors de dire que toutes les fonctions dérivées partielles sont continues en a, et donc que :

∀ α > 0, ∃ 0 < δ' < δ, ∀ k ∈ $\mathbb{R}^p$, ($\|h\|_\infty \le \delta'$) ⇒ (∀ 1 ≤ i ≤ p, $\left\| \frac{\partial f}{\partial x_i}(a+k) - \frac{\partial f}{\partial x_i}(a) \right\|_\infty \le \frac{\alpha}{p}$).

Par conséquent avec les notations précédentes, pour : h ∈ $\mathbb{R}^p$ tel que : $\|h\|_\infty \le \delta'$, on a :

∀ 1 ≤ i ≤ p, $\|h_1.e_1+...+h_{i-1}.e_{i-1}+\theta_i.h_i.e_i\|_\infty = \max(|h_1|,...,|h_{i-1}|,|\theta_i.h_i|) \le \max(|h_1|,...,|h_{i-1}|,|h_i|) \le \|h\|_\infty \le \delta'$,

et donc : ∀ 1 ≤ i ≤ p, $\left\| \frac{\partial f}{\partial x_i}(a+h_1.e_1+...+h_{i-1}.e_{i-1}+\theta_i.h_i.e_i) - \frac{\partial f}{\partial x_i}(a) \right\|_\infty \le \frac{\alpha}{p}$.

Pour h tel que : $0 < \|h\|_\infty \le \delta'$, on remarque que : ∀ 1 ≤ i ≤ p, $\frac{|h_i|}{\|h\|_\infty} \le 1$, et on en déduit que :

$$\|\varepsilon(h)\|_\infty \le \sum_{i=1}^{p} \frac{|h_i|}{\|h\|_\infty}.\left\| \frac{\partial f}{\partial x_i}(a+h_1.e_1+...+h_{i-1}.e_{i-1}+\theta_i.h_i.e_i) - \frac{\partial f}{\partial x_i}(a) \right\|_\infty \le p.\frac{\alpha}{p} = \alpha.$$

Autrement dit on a montré que : ∀ α > 0, ∃ δ' > 0, ∀ h ∈ $\mathbb{R}^p$, ($0 < \|h\|_\infty \le \delta'$) ⇒ ($\|\varepsilon(h)\|_\infty \le \alpha$), et la fonction ε tend vers 0 quand h tend vers 0.

On a donc bien : $f(a+h) = f(a) + \sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a) + \|h\|_\infty.\varepsilon(h) = f(a) + \sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a) + o(h)$.

• L'unicité enfin vient du fait que si : ∀ h ∈ B_a, $f(a) + \sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a) + o(h) = f(a) + \sum_{i=1}^{p} h_i.\alpha_i + o(h)$,

alors : ∀ h ∈ B_a, $\sum_{i=1}^{p} h_i.\left(\frac{\partial f}{\partial x_i}(a) - \alpha_i \right) = o(h) = \|h\|_\infty.\varepsilon(h)$.

Si on prend alors le cas particulier : ∀ 1 ≤ i ≤ p, $h_i = t.e_i = (0,...,0,t,0,...,0)$, avec : t > 0, on obtient :

∀ 1 ≤ i ≤ p, $t.\left(\frac{\partial f}{\partial x_i}(a) - \alpha_i \right) = t.\varepsilon(t.e_i)$, soit : $\left(\frac{\partial f}{\partial x_i}(a) - \alpha_i \right) = \varepsilon(t.e_i)$,

et en faisant tendre t vers 0, on aboutit à : $0 = \lim_{t \to 0} \varepsilon(t.e_i) = \left(\frac{\partial f}{\partial x_i}(a) - \alpha_i \right)$, soit bien : $\frac{\partial f}{\partial x_i}(a) = \alpha_i$.

Définition 1.3 : différentielle d'une fonction en un point

Soit f une fonction définie d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$, de classe C^1.
On appelle différentielle de f en a, la forme linéaire notée $df(a)$, définie par :

$$\forall\ h = \sum_{i=1}^{p} h_i.e_i \in \mathbb{R}^p,\ \ df(a)(h) = \sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a) = \sum_{i=1}^{p} h_i.\partial_i f(a).$$

On a donc alors pour h au voisinage de a : $f(a+h) = f(a) + df(a)(h) + o(h)$.

Théorème 1.3 : classe C^1 et continuité

Si f est une fonction de classe C^1 d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$, alors f est continue sur U.

Démonstration :

Si f est de classe C^1 sur U, alors pour : a ∈ U, et : h ∈ $\mathbb{R}^p$, tel que : (a + h) ∈ U, on a :
$f(a+h) = f(a) + df(a)(h) + o(h)$.

Or $df(a)$, comme application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}$, espaces vectoriels de dimension finie, est continue.

Donc : $\lim_{h\to 0} df(a)(h) = df(a)(0) = 0$, de même que : $\lim_{h\to 0} o(h) = 0$,

d'où on déduit la continuité de f en a et donc sur U.

Théorème 1.4 : opérations sur les fonctions de classe C^1

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

Si f et g sont des fonctions de classe C^1 de U dans $\mathbb{R}$, et ψ une fonction définie au moins de : f(U) ⊂ $\mathbb{R}$, dans $\mathbb{R}$, et de classe C^1, alors :

- ∀ (λ,μ) ∈ $\mathbb{R}^2$, $(\lambda.f + \mu.g)$ est de classe C^1 sur U, et : ∀ a ∈ U, $d(\lambda.f + \mu.g)(a) = \lambda.df(a) + \mu.dg(a)$,
- $(f.g)$ est de classe C^1 sur U, et : ∀ a ∈ U, ∀ 1 ≤ i ≤ p, $\dfrac{\partial(f.g)}{\partial x_i}(a) = \dfrac{\partial f}{\partial x_i}(a).g(a) + f(a).\dfrac{\partial g}{\partial x_i}(a)$,
- si de plus f ne s'annule pas sur U, $\dfrac{1}{f}$ est de classe C^1 sur U, et :

  ∀ a ∈ U, $d\left(\dfrac{1}{f}\right)(a) = -\dfrac{1}{(f(a))^2}.df(a)$, ou : ∀ 1 ≤ i ≤ p, $\dfrac{\partial}{\partial x_i}\left(\dfrac{1}{f}\right)(a) = -\dfrac{1}{(f(a))^2}.\dfrac{\partial f}{\partial x_i}(a)$,
- si g ne s'annule pas sur U, alors $\dfrac{f}{g}$ est de classe C^1 sur U, et :

  ∀ a ∈ U, $d\left(\dfrac{f}{g}\right)(a) = \dfrac{g(a).df(a) - f(a).dg(a)}{(g(a))^2}$, de même avec les dérivées partielles.
- $\psi o f$ est de classe C^1 sur U et : ∀ a ∈ U, ∀ 1 ≤ i ≤ p, $\dfrac{\partial(\psi o f)}{\partial x_i}(a) = \psi'(f(a)).\dfrac{\partial f}{\partial x_i}(a)$.

Démonstration :

- pour : a ∈ U, et : 1 ≤ i ≤ p, $\lim_{t\to 0}\frac{1}{t}[f(a+t.e_i) - f(a)]$ et $\lim_{t\to 0}\frac{1}{t}[g(a+t.e_i) - g(a)]$.

Donc $\lim_{t\to 0}\frac{1}{t}[(\lambda.f+\mu.g)(a+t.e_i) - (\lambda.f+\mu.g)(a)]$ existe aussi, et :

$$\lim_{t\to 0}\frac{1}{t}[(\lambda.f+\mu.g)(a+t.e_i) - (\lambda.f+\mu.g)(a)] = \lambda.\lim_{t\to 0}\frac{1}{t}[f(a+t.e_i) - f(a)] + \mu.\lim_{t\to 0}\frac{1}{t}[f(a+t.e_i) - f(a)],$$

donc : ∀ a ∈ U, ∀ 1 ≤ i ≤ p, $\dfrac{\partial}{\partial x_i}(\lambda.f+\mu.g)(a) = \lambda.\dfrac{\partial f}{\partial x_i}(a).+\mu.\dfrac{\partial g}{\partial x_i}(a)$, et les dérivées partielles obtenues sont clairement continues sur U, autrement dit $(\lambda.f+\mu.g)$ est de classe C^1 sur U.

De plus : ∀ a ∈ U, $d(\lambda.f+\mu.g)(a) = \lambda.df(a) + \mu.dg(a)$.

- pour : a ∈ U, et pour : 1 ≤ i ≤ p, la fonction φ_i : t ↦ $(f.g)(a+t.e_i) = f(a+t.e_i).g(a+t.e_i)$, est une fonction qui admet une dérivée en 0, comme produit de deux fonctions d'un intervalle de $\mathbb{R}$ dans $\mathbb{R}$ ayant cette propriété.

De plus : $\forall$ 1 ≤ i ≤ p, $\varphi_i'(0) = \dfrac{\partial (f.g)}{\partial x_i}(a) = f(a).\dfrac{\partial g}{\partial x_i}(a) + \dfrac{\partial f}{\partial x_i}(a).g(a)$.

On vient donc d'établir que $(f.g)$ admet des dérivées partielles en tout point de U, continues sur U.
Donc $(f.g)$ est bien de classe C^1 sur U autrement dit $(f.g)$ est de classe C^1 sur U.
De plus : : $\forall$ a ∈ U, $\forall$ h ∈ U, tel que : (a+h) ∈ U, on a :

$$d(f.g)(a)(h) = \sum_{i=1}^{p}\left[f(a).\frac{\partial g}{\partial x_i}(a) + \frac{\partial f}{\partial x_i}(a).g(a)\right].h_i = f(a).\sum_{i=1}^{p} h_i.\frac{\partial g}{\partial x_i}(a) + g(a).\sum_{i=1}^{p} h_i.\frac{\partial f}{\partial x_i}(a),$$

ce qui s'écrit encore : $d(f.g)(a)(h) = f(a).dg(a)(h) + g(a).df(a)(h) = [f(a).dg(a) + g(a).df(a)](h)$.

• de même, si f ne s'annule pas sur U, on a : $\forall$ a ∈ U, $\forall$ 1 ≤ i ≤ p, $\psi_i : t \mapsto \dfrac{1}{f(a+t.e_i)}$, admet une dérivée en 0 comme inverse d'une fonction d'un intervalle de ℝ dans ℝ qui en admet une, et :

$\forall$ 1 ≤ i ≤ p, $\psi_i'(0) = \dfrac{\partial}{\partial x_i}\left[\dfrac{1}{f}\right](a) = -\dfrac{\partial f}{\partial x_i}(a).\dfrac{1}{(f(a))^2}$, qui est bien continue sur U.

De plus, on obtient de la même façon que précédemment la forme annoncée de la différentielle.
• le point suivant combine simplement les deux points qui le précédent.
• pour : a ∈ U, et pour : 1 ≤ i ≤ p, la fonction $\varphi_i : t \mapsto (\psi o f)(a+t.e_i) = \psi(f(a+t.e_i))$, est une fonction qui admet une dérivée en 0, comme composée de deux fonctions entre intervalles de ℝ et à valeurs dans ℝ ayant cette propriété.

De plus : $\forall$ 1 ≤ i ≤ p, $(\psi o f)'(0) = \dfrac{\partial(\psi o f)}{\partial x_i}(a) = \psi'(f(a)).\dfrac{\partial f}{\partial x_i}(a)$.

A nouveau, on constate pour finir que ces dérivées partielles sont bien continues sur U.

Théorème 1.5 : règle de la chaîne (dérivée le long d'un arc paramétré)

Soit f une fonction d'un ouvert U de $\mathbb{R}^p$ dans ℝ de classe C^1.
Soit φ une fonction de classe C^1 d'un intervalle I de ℝ dans $\mathbb{R}^p$ de telle sorte que : φ(I) ⊂ U, pour laquelle on note : $\forall$ t ∈ I, $\varphi(t) = (x_1(t),\ldots,x_p(t))$, avec : $\forall$ 1 ≤ i ≤ p, x_i de classe C^1 de I dans ℝ.
Alors $(f o \varphi)$ est de classe C^1 de I dans ℝ, et :

$$\forall t \in I,\ (f o\varphi)'(t) = df(\varphi(t))(\varphi'(t)) = \sum_{i=1}^{p}\frac{\partial f}{\partial x_i}(\varphi(t)).x_i'(t).$$

Démonstration :

Pour : t ∈ I, chaque fonction x_i admet un développement limité à l'ordre 1 en t :
$\forall$ h ∈ ℝ, avec : (t + h) ∈ I, on a : $x_i(t+h) = x_i(t) + h.x_i'(t) + o(h)$.

Donc : $\forall$ h ∈ ℝ, avec : (t + h) ∈ I, on peut écrire :

$$(f o\varphi)(t+h) = f(x_1(t) + h.x_1'(t) + o_1(h),\ldots,x_p(t) + h.x_p'(t) + o_p(h)) = f(\varphi(t) + k),$$

où on a noté : $k = (h.x_1'(t) + o_1(h),\ldots,h.x_p'(t) + o_p(h))$, et donc :

$$(f o\varphi)(t+h) = \sum_{i=1}^{p}\frac{\partial f}{\partial x_i}(\varphi(t)).(h.x_i'(t) + o_i(h)) = h.\sum_{i=1}^{p}\frac{\partial f}{\partial x_i}(\varphi(t)).x_i'(t) + \sum_{i=1}^{p}\frac{\partial f}{\partial x_i}(\varphi(t)).h.\varepsilon_i(h).$$

De plus : $\displaystyle\sum_{i=1}^{p}\frac{\partial f}{\partial x_i}(\varphi(t)).h.\varepsilon_i(h) = h.\sum_{i=1}^{p}\frac{\partial f}{\partial x_i}(\varphi(t)).\varepsilon_i(h) = h.\varepsilon(h)$, où la fonction ε, comme combinaison linéaire de fonctions qui tendent vers 0, tend encore vers 0 quand h tend vers 0.

Enfin, on a : $\displaystyle\lim_{h\to 0}\frac{(f o\varphi)(t+h) - (f o\varphi)(t)}{h} = \sum_{i=1}^{p}\frac{\partial f}{\partial x_i}(\varphi(t)).x_i'(t) + \lim_{h\to 0}\varepsilon(h) = \sum_{i=1}^{p}\frac{\partial f}{\partial x_i}(\varphi(t)).x_i'(t)$,

donc $(f o\varphi)$ est dérivable sur I, et sa dérivée est bien celle annoncée.
Et pour terminer cette dérivée est une somme de produits de composées de fonctions continues et est donc continue sur I, ce qui prouve que $(f o\varphi)$ est bien de classe C^1 sur I.

Théorème 1.6 : caractérisation des fonctions constantes sur un ouvert convexe

Soit f une fonction f de classe C^1 définie sur un ouvert convexe U de $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}$.
Alors f est constante sur U si et seulement si : $\forall$ a $\in$ U, $df(a) = 0$.

Démonstration :

• Si f est constante sur U, alors ses fonctions partielles (qu'on avait notées f_i) le sont aussi, donc elles ont toutes une dérivée nulle et : $\forall$ a $\in$ U, $df(a) = 0$.

• Supposons maintenant que la différentielle de f soit nulle pour tout a dans U, $df(a) = 0$.

Soit : (a,b) $\in$ U, et la fonction φ : t $\mapsto$ $(1-t).a + t.b$, décrivant le segment qui relie a à b.

Cette fonction est clairement de classe C^1 sur [0,1], et : $\forall$ t $\in$ [0,1], $\varphi'(t) = b - a = (b_1 - a_1, ..., b_p - a_p)$.

La fonction $(fo\varphi)$ est alors définie et de classe C^1 sur [0,1], et :

$\forall$ t $\in$ [0,1], $(fo\varphi)'(t) = df(\varphi(t))(b-a) = 0$.

Comme fonction de [0,1] dans $\mathbb{R}$, $(fo\varphi)$ est donc constante et : $f(a) = (fo\varphi)(0) = (fo\varphi)(1) = f(b)$.

f est donc bien constante sur U.

2. Gradient.

Définition 2.1 et théorème 2.1 : gradient d'une fonction de classe C^1, expression de la différentielle

Soit f une fonction de classe C^1 d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$.

Pour : a $\in$ U, on appelle gradient de f en a le vecteur défini par : $grad(f)(a) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a).e_i$.

Si on munit $\mathbb{R}^p$ de sa structure euclidienne canonique, alors :

$\forall$ a $\in$ U, $\forall$ h $\in$ $\mathbb{R}^p$, on a : $df(a)(h) = (grad(f)(a)|h)$.

Si on note le produit scalaire par •, on peut encore écrire cette égalité avec : $df(a)(h) = grad(f)(a) \bullet h$, et si on note le gradient à l'aide de l'opérateur ∇ (nabla), on a alors :

$\forall$ a $\in$ U, $\forall$ h $\in$ $\mathbb{R}^p$, $df(a)(h) = \nabla(f)(a) \bullet h$.

Il arrive parfois qu'on écrive alors aussi (abusivement) : $df(a)(h) = df(a) \bullet h$, en confondant la forme linéaire et le vecteur qui lui est associé.

Démonstration :

Pour la structure euclidienne canonique de $\mathbb{R}^p$ la base canonique est orthonormale et donc on a bien :

$\forall$ a $\in$ U, $\forall$ h $\in$ $\mathbb{R}^p$, $df(a)(h) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a).h_i = (grad(f)(a)|h)$.

Théorème 2.1 : dérivées partielles et changement de variables dans $\mathbb{R}^2$

Soit f une fonction d'un ouvert U de $\mathbb{R}^2$ dans $\mathbb{R}$ de classe C^1.
Soit Φ : (u,v) $\mapsto$ $(x(u,v), y(u,v))$, un changement de variables de $\mathbb{R}^2$ dans $\mathbb{R}^2$, tel que x et y sont des fonctions de classe C^1 de $\mathbb{R}^2$ dans $\mathbb{R}$.
La fonction composée $fo\Phi$ admet des dérivées partielles par rapport à u et v en tout point : (a,b) de $\mathbb{R}^2$, tel que : Φ(a,b) $\in$ U, et on a :

$\forall$ (a,b) $\in$ $\mathbb{R}^2$, tel que : (x,y) = Φ(a,b) $\in$ U,

• $\frac{\partial(fo\Phi)}{\partial u}(a,b) = \frac{\partial f}{\partial x}(\Phi(a,b)).\frac{\partial x}{\partial u}(a,b) + \frac{\partial f}{\partial y}(\Phi(a,b)).\frac{\partial y}{\partial u}(a,b)$, et :

• $\frac{\partial(fo\Phi)}{\partial v}(a,b) = \frac{\partial f}{\partial x}(\Phi(a,b)).\frac{\partial x}{\partial v}(a,b) + \frac{\partial f}{\partial y}(\Phi(a,b)).\frac{\partial y}{\partial v}(a,b)$.

Démonstration :

Soit : (a,b) $\in$ $\mathbb{R}^2$, tel que : Φ(a,b) $\in$ U.
Considérons les deux fonctions partielles définies autour de a, et à valeurs dans $\mathbb{R}$, données par :

• h $\mapsto$ $x_1(h) = x(h,b)$,

• h $\mapsto$ $x_2(h) = y(h,b)$,

et la fonction φ définie autour de a et à valeurs dans $\mathbb{R}^2$, donnée par : h $\mapsto$ $(x_1(h), x_2(h))$.

La fonction φ est de classe C^1 d'un intervalle contenant a dans $\mathbb{R}$ et le théorème précédent montre alors que $(fo\varphi)$ est de classe C[1] sur cet intervalle et que en particulier :

$$(fo\varphi_1)'(a) = \frac{\partial f}{\partial x}(\varphi_1(a)).x_1'(a) + \frac{\partial f}{\partial y}(\varphi_1(a)).x_2'(a) = \frac{\partial f}{\partial x}(\Phi(a,b)).\frac{\partial x}{\partial u}(a,b) + \frac{\partial f}{\partial y}(\Phi(a,b)).\frac{\partial y}{\partial u}(a,b).$$

Or ce qu'on vient de calculer, c'est la dérivée partielle de : (u,v) $\mapsto (fo\Phi)(u,v)$, par rapport à u en (a,b),

soit : $\frac{\partial(fo\Phi)}{\partial u}(a,b) = \frac{\partial f}{\partial x}(\Phi(a,b)).\frac{\partial x}{\partial u}(a,b) + \frac{\partial f}{\partial y}(\Phi(a,b)).\frac{\partial y}{\partial u}(a,b)$.

Le même résultat s'obtient pour la dérivée partielle de $(fo\Phi)$ par rapport à v en (a,b).

Exemple 2.1 : changement de coordonnées dans $\mathbb{R}^2$, exemple du changement polaires-cartésiennes

- Soit f une fonction de classe C^1 d'un ouvert U de $\mathbb{R}^2$ dans $\mathbb{R}$, Φ une fonction de classe C^1 d'un ouvert V de $\mathbb{R}^2$ dans U, et : $F = fo\Phi$.

On note alors : $\forall$ (u,v) $\in$ V, $\Phi(u,v) = (x(u,v), y(u,v))$.

On a dans ce cas : $\forall$ (u,v) $\in$ V, $F(u,v) = f(x(u,v), y(u,v))$.

Il arrive souvent qu'on désigne par u et v les variables de F et (abusivement) par x et y les variables de f, x et y étant à la fois vues comme coordonnées dans U et comme fonctions des « anciennes » coordonnées u et v.

F est alors de classe C^1 de V dans $\mathbb{R}$, et :

$$\forall\ (u_0,v_0) \in V,\ \frac{\partial F}{\partial u}(u_0,v_0) = \frac{\partial f}{\partial x}(x(u_0,v_0), y(u_0,v_0)).\frac{\partial x}{\partial u}(v_0,v_0) + \frac{\partial f}{\partial y}(x(u_0,v_0), y(u_0,v_0)).\frac{\partial y}{\partial u}(u_0,v_0),$$

de même pour la dérivée de F par rapport à sa deuxième variable

Si on considère alors une grandeur physique scalaire E définie sur une portion Π du plan géométrique dans lequel on dispose de deux systèmes de coordonnées, et dont le calcul se fait, soit par une fonction f en référence au premier système de coordonnées (ici (u,v)), soit par F en référence au second système (ici (x,y)), on arrive alors à l'expression « à la physicienne » :

$\forall$ M $\in$ Π, $E(M) = f(x,y) = F(u,v)$, puis :

$\forall$ M $\in$ Π, $\frac{\partial E}{\partial u}(M) = \frac{\partial E}{\partial x}(M).\frac{\partial x}{\partial u} + \frac{\partial E}{\partial y}(M).\frac{\partial y}{\partial u}$, ou plus simplement : $\frac{\partial E}{\partial u} = \frac{\partial E}{\partial x}.\frac{\partial x}{\partial u} + \frac{\partial E}{\partial y}.\frac{\partial y}{\partial x}$,

les dérivées partielles indiquant à chaque fois à quelle façon de calculer E (soit f, soit F) on se réfère.

- En particulier, si on considère le changement de variables polaires-cartésiennes Φ défini par :

$$\forall\ (\rho,\theta) \in\ ]0,+\infty)\times]-\pi,+\pi],\ \begin{cases} x = \rho.\cos(\theta) \\ y = \rho.\sin(\theta) \end{cases},$$

et si on se donne deux fonctions f et g à valeurs scalaires liées par la relation :

$\forall$ $(\rho,\theta) \in\ ]0,+\infty)\times]-\pi,+\pi]$, $F(\rho,\theta) = f(x,y) = f(\rho.\cos(\theta), \rho.\sin(\theta))$,

alors si f est de classe C^1 sur $\mathbb{R}^2$, g est de classe C^1 sur son ensemble de définition et :

$$\begin{cases} \frac{\partial F}{\partial \rho}(\rho,\theta) = \frac{\partial f}{\partial x}(\rho.\cos(\theta), \rho.\sin(\theta)).\frac{\partial x}{\partial \rho}(\rho,\theta) + \frac{\partial f}{\partial y}(\rho.\cos(\theta), \rho.\sin(\theta)).\frac{\partial y}{\partial \rho}(\rho,\theta) \\ \frac{\partial F}{\partial \theta}(\rho,\theta) = \frac{\partial f}{\partial x}(\rho.\cos(\theta), \rho.\sin(\theta)).\frac{\partial x}{\partial \theta}(\rho,\theta) + \frac{\partial f}{\partial y}(\rho.\cos(\theta), \rho.\sin(\theta)).\frac{\partial y}{\partial \theta}(\rho,\theta) \end{cases},$$

ou encore : $\begin{cases} \frac{\partial F}{\partial \rho}(\rho,\theta) = \frac{\partial f}{\partial x}(x,y).\cos(\theta) + \frac{\partial f}{\partial y}(x,y).\sin(\theta) \\ \frac{\partial F}{\partial \theta}(\rho,\theta) = -\frac{\partial f}{\partial x}(x,y).\rho.\sin(\theta) + \frac{\partial f}{\partial y}(x,y).\rho.\cos(\theta) \end{cases}$.

Exemple 2.2 : expression du gradient en coordonnées polaires

Soient U un ouvert de $\mathbb{R}^2$ ne contenant pas (0,0), et f une fonction de classe C^1 de U dans $\mathbb{R}$.

On pose par ailleurs : $F(\rho,\theta) = f(\rho.\cos(\theta), \rho.\sin(\theta)) = f(x,y)$, avec : $x = \rho.\cos(\theta)$, $y = \rho.\sin(\theta)$.

Autrement dit, f et F pourraient représenter la même grandeur physique (attaché au point M de coordonnées cartésiennes (x,y) et polaires (ρ,θ)).

Si on note u_θ et u'_θ les vecteurs : $\overrightarrow{u_\theta} = \cos(\theta).\vec{i} + \sin(\theta)\vec{j}$, et : $\overrightarrow{u_\theta}' = -\sin(\theta).\vec{i} + \cos(\theta).\vec{j}$, avec (i,j) la base orthonormale naturelle de $\mathbb{R}^2$, dans laquelle les coordonnées sont notées x et y, alors :

$\forall$ (x,y) $\in$ U, $\overrightarrow{grad}(f)(x,y) = \overrightarrow{grad}(f)(\rho.\cos(\theta),\rho.\sin(\theta)) = \dfrac{\partial F}{\partial \rho}(\rho,\theta).\overrightarrow{u_\theta} + \dfrac{1}{\rho}.\dfrac{\partial F}{\partial \theta}(\rho,\theta).\overrightarrow{u_\theta}'$.

Démonstration :

On remarque tout d'abord que U ne contenant pas (0,0), la valeur de ρ est toujours non nulle.
Puis si on reprend l'expression obtenue dans l'exemple 2.1 pour les dérivées partielles de F, on constate qu'on peut « inverser » ces équations et obtenir :

- $\dfrac{\partial f}{\partial x}(x,y) = \dfrac{\partial F}{\partial \rho}(\rho,\theta).\cos(\theta) - \dfrac{1}{\rho}.\dfrac{\partial F}{\partial \theta}(\rho,\theta).\sin(\theta)$, et :
- $\dfrac{\partial f}{\partial y}(x,y) = \dfrac{\partial F}{\partial \rho}(\rho,\theta).\sin(\theta) + \dfrac{1}{\rho}.\dfrac{\partial F}{\partial \theta}(\rho,\theta).\cos(\theta)$.

Alors :

$$\begin{aligned}\overrightarrow{grad}(f)(x,y) &= \left(\frac{\partial F}{\partial \rho}(\rho,\theta).\cos(\theta) - \frac{1}{\rho}.\frac{\partial F}{\partial \theta}(\rho,\theta).\sin(\theta)\right).\vec{i} + \left(\frac{\partial F}{\partial \rho}(\rho,\theta).\sin(\theta) + \frac{1}{\rho}.\frac{\partial F}{\partial \theta}(\rho,\theta).\cos(\theta)\right).\vec{j} \\ &= \frac{\partial F}{\partial \rho}(\rho,\theta).(\cos(\theta).\vec{i} + \sin(\theta)\vec{j}) + \frac{1}{\rho}.\frac{\partial F}{\partial \theta}(\rho,\theta).(-\sin(\theta).\vec{i} + \cos(\theta).\vec{j}) \\ &= \frac{\partial F}{\partial \rho}(\rho,\theta).\overrightarrow{u_\theta} + \frac{1}{\rho}.\frac{\partial F}{\partial \theta}(\rho,\theta).\overrightarrow{u_\theta}'.\end{aligned}$$

3. Dérivées d'ordre 2, étude d'extrema de fonctions à valeurs réelles.

Définition 3.1 : dérivées partielles d'ordre 2 d'une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$ en un point

Soit f une fonction définie d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$ et de classe C^1, et soit : a $\in$ U.

Si pour : $1 \leq i,j \leq p$, la fonction $\dfrac{\partial f}{\partial x_i}$ admet une $j^{\text{ième}}$ dérivée partielle en a, on note $\dfrac{\partial}{\partial x_j}\left(\dfrac{\partial f}{\partial x_i}\right)(a)$ ou $\partial_{j,i} f(a)$, cette dérivée partielle seconde, et pour : i = j, on écrit : $\dfrac{\partial}{\partial x_i}\left(\dfrac{\partial f}{\partial x_i}\right)(a) = \dfrac{\partial^2 f}{\partial x_i^{\ 2}}(a) = \partial_{i,i} f(a)$.

Définition 3.2 : fonction de classe C^2 de $\mathbb{R}^p$ dans $\mathbb{R}$

Soit f une fonction d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$.
On dit que la fonction f est de classe C^2 sur U si et seulement si f est de classe C^1 sur U et si toutes ses dérivées partielles sont de classe C^1.

Théorème 3.1 : de Schwarz

Soit f une fonction de classe C^2 d'un ouvert U de $\mathbb{R}^p$ dans $\mathbb{R}$.

Alors : $\forall$ a $\in$ U, $\forall$ $1 \leq i \neq j \leq n$, $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$.

Démonstration : (hors programme)

• on va commencer par démontrer le résultat pour les fonctions à valeurs réelles et donc on considère f de classe C^2 de U dans $\mathbb{R}$.
Soit : a $\in$ U, $(e_1, \dots, e_n)$ la base $\mathscr{B}$ de E et : $(i,j) \in \mathbb{N}_n^{\ 2}$.
On note : $V = \{(h_i,h_j) \in \mathbb{R}^2,\ (a + h_i.e_i + h_j.e_j) \in U\}$.

Soit par ailleurs φ définie de [0,1] dans $\mathbb{R}$ par :

$\varphi(t) = f(a + t.h_i.e_i + h_j.e_j) - f(a + t.h_i.e_i)$.

φ est alors définie, continue et dérivable sur [0,1], et sa dérivée vaut :

$\forall$ t $\in$ [0,1], $\varphi'(t) = h_i.\dfrac{\partial f}{\partial x_i}(a + t.h_i.e_i + h_j.e_j) - h_i.\dfrac{\partial f}{\partial x_i}(a + t.h_i.e_i)$.

On peut alors lui appliquer le théorème des accroissements finis, pour obtenir :

$\exists\ \theta_\varphi \in\ ]0,1[,\ \varphi(1)-\varphi(0)=(1-0).\varphi'(\theta_\varphi)=h_i.\frac{\partial f}{\partial x_i}(a+\theta_\varphi.h_i.e_i+h_j.e_j)-h_i.\frac{\partial f}{\partial x_i}(a+\theta_\varphi.h_i.e_i)$.

Pour cette valeur θ_φ fixée, notons alors ψ la fonction de [0,1] dans ℝ, définie par :

$\forall\ t \in [0,1],\ \psi(t)=\frac{\partial f}{\partial x_i}(a+\theta_\varphi.h_i.e_i+t.h_j.e_j)$.

Elle est également définie, continue et dérivable sur [0,1] et le théorème des accroissements finis à nouveau donne :

$\exists\ \theta_\psi \in\ ]0,1[, \psi(1)-\psi(0)=(1-0).\psi'(\theta_\psi)=h_j.\frac{\partial^2 f}{\partial x_j \partial x_i}(a+\theta_\varphi.h_i.e_i+\theta_\psi.h_j.e_j)$.

Globalement, on a donc :

$$\begin{aligned}\varphi(1)-\varphi(0) &= (f(a+h_i.e_i+h_j.e_j)-f(a+h_i.e_i))-(f(a+h_j.e_j)-f(a))\\ &= h_i.\left(\frac{\partial f}{\partial x_i}(a+\theta_\varphi.h_i.e_i+h_j.e_j)-\frac{\partial f}{\partial x_i}(a+\theta_\varphi.h_i.e_i)\right)\\ &= h_i.(\psi(1)-\psi(0))\\ &= h_i.h_j.\frac{\partial^2 f}{\partial x_j \partial x_i}(a+\theta_\varphi.h_i.e_i+\theta_\psi.h_j.e_j).\end{aligned}$$

En intervertissant les rôles de i et de j, on peut également trouver θ'_φ et θ'_ψ dans]0,1[, tels que :

$(f(a+h_i.e_i+h_j.e_j)-f(a+h_j.e_j))-(f(a+h_i.e_i)-f(a))=h_i.h_j.\frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta'_\varphi.h_i.e_i+\theta'_\psi.h_j.e_j)$.

Donc, en divisant par $(h_i.h_j)$, pour : $h_i.h_j \neq 0$, on aboutit à (puisqu'on vient d'évaluer deux fois la même quantité : $\frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta'_\varphi.h_i.e_i+\theta'_\psi.h_j.e_j)=\frac{\partial^2 f}{\partial x_j \partial x_i}(a+\theta_\varphi.h_i.e_i+\theta_\psi.h_j.e_j)$.

Enfin, lorsque l'on fait tendre (h_i, h_j) vers (0,0), puisque les deux dérivées partielles secondes sont continues en a, on obtient à la limite : $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)=\frac{\partial^2 f}{\partial x_j \partial x_i}(a)$.

Définition 3.3 : extremum d'une fonction à valeurs réelles

Soit U un ouvert de $\mathbb{R}^p$ et soit f une fonction de U dans ℝ.
On dit que f présente un minimum local (respectivement maximum local) en a si et seulement si il existe un voisinage ouvert V de a (ou une boule ouverte centrée en a) et inclus dans U tel que f(a) est le minimum (respectivement maximum) de f sur V.
On dit de même que f présente un minimum global (respectivement maximum local) en a si et seulement si f(a) est le minimum (respectivement maximum) de f sur U.

Définition 3.4 : point critique d'une fonction de classe C^1 à valeurs réelles

Soit f une fonction d'un ouvert U de $\mathbb{R}^p$ dans ℝ, de classe C^1 sur U.
On dit que a est un point critique de f si et seulement si : $df(a)=0$.

Théorème 3.2 : condition nécessaire d'extremum pour une fonction de classe C^1 sur un ouvert

Soit f une fonction d'un ouvert U de $\mathbb{R}^p$ dans ℝ, de classe C^1 sur U.
Si f présente un extremum local ou global en un point a de U, alors a est un point critique de f et on a donc : $df(a)=0$.

Démonstration :

Soit $\|.\|_\infty$ la norme infinie dans $\mathbb{R}^p$.

Puisque U est un ouvert, il existe : r > 0, tel que : B(a,r) ⊂ U.
En particulier : $\forall\ (h_1, \ldots, h_n) \in \mathbb{R}^n$, $(\forall\ 1 \leq i \leq n, |h_i| < r) \Rightarrow ((a+h_1.e_1+\ldots+h_n.e_n) \in U)$

De plus, si f présente un extremum local en a, on peut considérer, quitte à prendre une valeur de r plus petite que f présente un extremum global en a sur la boule B(a,r).

Puis pour : $1 \leq i \leq n$, la fonction partielle φ_i de]-R,+R[dans ℝ, définie par :

$\forall t \in$]-R,+R[, $\varphi_i(t) = f(a + t.e_i)$,

présente un extremum en 0.

Donc, puisque de plus φ_i est dérivable en 0, on a : $\varphi_i'(0) = 0$, soit : $\frac{\partial f}{\partial x_i}(a) = 0$.

Et comme toutes les dérivées partielles de f en a sont nulles, la différentielle de f en a est bien nulle.

Remarque :

Si f est définie de : $A \subset \mathbb{R}^p$, dans ℝ où A n'est pas un ouvert, f étant de classe C^1 sur A, et si f présente un extremum (local ou global) en : $a \in A$, alors :

- $df(a) = 0$, si : a est intérieur à A,
- ou a est « au bord » de A, c'est-à-dire que a est dans A, mais pas intérieur à A.