Chapitre 5

Dérivées d'ordres supérieurs. Application à l'étude d'extrema.

On s'intéresse dans ce chapitre aux dérivées d'ordre 2 ou plus d'une fonction de plusieurs variables. Comme pour une fonction d'une seule variable, la dérivée première (la différentielle ici) permet d'avoir la meilleure approximation d'une fonction en un point par une fonction affine, mais on peut avoir besoin d'être plus précis.

L'application que l'on va détailler ici est l'étude des extremums d'une fonction à valeurs réelles. Comme pour une fonction d'une variable, on va voir que la différentielle première s'annule là où la fonction atteint ses extremums locaux, mais on aura besoin de la dérivée seconde pour voir s'il s'agit effectivement d'un extremum local et, le cas échéant, s'il s'agit d'un maximum ou d'un minimum. Cela a évidemment de nombreuses applications, les problèmes où on cherche à optimiser une grandeur en fonction de divers paramètres étant très nombreux.

5.1 Dérivées partielles successives

Soit f une fonction d'un ouvert $\mathcal{U}$ de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$.

Définition 5.1.
- On dit que f est de classe C^2 sur $\mathcal{U}$ si elle est de classe C^1 et que toutes ses dérivées partielles sont de classe C^1 sur $\mathcal{U}$.
- Par récurrence, on dit que f est de classe C^k sur $\mathcal{U}$ si elle est de classe C^1 et que toutes ses dérivées partielles sont de classe C^{k-1} sur $\mathcal{U}$.
- On dit que f est de classe C^∞ sur $\mathcal{U}$ si f est de classe C^k sur $\mathcal{U}$ pour tout $k \in \mathbb{N}^*$.

Si f est de classe C^2, alors pour $j, k \in [\![1, n]\!]$ on note

$$\frac{\partial^2 f}{\partial x_j^2} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_j}\right) \quad \text{et} \quad \frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right).$$

On a également des notations analogues pour les dérivées partielles à tout ordre.

***Exercice* 5.1.** Calculer, en tout point (x, y) où elles sont définies, toutes les dérivées partielles secondes des fonctions de deux variables suivantes :

$$f_1 : (x, y) \mapsto x^2 \quad ; \quad f_2 : (x, y) \mapsto x^2 \cos(y) \quad ; \quad f_3 : (x, y) \mapsto x^y.$$

Les résultats de l'exercice précédent semblent indiquer que les dérivées croisées $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$ sont égales. C'est effectivement le cas pour une fonction de classe C^2. Et c'est une très bonne nouvelle.

Théorème 5.2 (Théorème de Schwarz). *On suppose que f est de classe C^2 sur $\mathcal{U}$. Alors pour tous $j,k \in [\![1,n]\!]$ on a sur $\mathcal{U}$:*

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j}.$$

Remarque 5.3. Attention tout de même ! Il existe des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ telles que les dérivées secondes $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$ existent en un point a mais prennent des valeurs distinctes (voir exercice 5.6). Une telle fonction n'est donc pas de classe C^2.

Démonstration. Soient $a \in \mathcal{U}$ et $j,k \in [\![1,n]\!]$. Soit $\delta_0 > 0$ tel que $a + h_j e_j + h_k e_k \in \mathcal{U}$ pour tout $h_j, h_k \in [0,\delta_0]$. Pour $h_j, h_k \in]0,\delta_0]$ on note

$$\phi(h_j,h_k) = \frac{1}{h_j h_k}\big(f(a+h_je_j+h_ke_k) - f(a+h_je_j) - f(a+h_ke_k) + f(a)\big).$$

On a

$$\begin{aligned}
\phi(h_j,h_k) &= \frac{1}{h_jh_k}\left(\int_0^{h_k} \frac{\partial f}{\partial x_k}(a+h_je_j+te_k)\,dt - \int_0^{h_k}\frac{\partial f}{\partial x_k}(a+te_k)\,dt\right)\\
&= \frac{1}{h_j}\int_0^1 \left(\frac{\partial f}{\partial x_k}(a+h_je_j+sh_ke_k) - \frac{\partial f}{\partial x_k}(a+sh_ke_k)\right)ds\\
&= \int_0^1\int_0^1 \frac{\partial^2 f}{\partial x_j\partial x_k}(a+\tau h_je_j+sh_ke_k)\,ds\,d\tau.
\end{aligned}$$

Soit $\varepsilon > 0$. Par continuité de $\frac{\partial^2 f}{\partial x_j \partial x_k}$ il existe $\delta \in]0,\delta_0]$ tel que pour tous $h_j,h_k \in [0,\delta]$ on a

$$\left\|\frac{\partial^2 f}{\partial x_j\partial x_k}(a+h_je_j+h_ke_k) - \frac{\partial^2 f}{\partial x_j\partial x_k}(a)\right\| < \varepsilon$$

Ainsi pour $h_j,h_k \in [0,\delta]$ on a

$$\left\|\phi(h_j,h_k) - \frac{\partial^2 f}{\partial x_j\partial x_k}(a)\right\| \leqslant \int_0^1\int_0^1 \left\|\frac{\partial^2 f}{\partial x_j\partial x_k}(a+\tau h_je_j+sh_ke_j) - \frac{\partial^2 f}{\partial x_j\partial x_k}(a)\right\| ds\,d\tau \leqslant \varepsilon.$$

Cela prouve que

$$\phi(h_j,h_k) \xrightarrow[(h_j,h_k)\to(0,0)]{} \frac{\partial^2 f}{\partial x_j\partial x_k}(a).$$

Mais d'autre part on a

$$\begin{aligned}
\phi(h_j,h_k) &= \frac{1}{h_jh_k}\left(\int_0^{h_j}\frac{\partial f}{\partial x_j}(a+te_j+h_ke_k)\,dt - \int_0^{h_j}\frac{\partial f}{\partial x_j}(a+te_j)\,dt\right)\\
&= \int_0^1\int_0^1 \frac{\partial^2 f}{\partial x_k\partial x_j}(a+\tau h_je_j+sh_ke_j)\,d\tau\,ds,
\end{aligned}$$

et on obtient de la même manière que

$$\phi(h_j,h_k) \xrightarrow[(h_j,h_k)\to(0,0)]{} \frac{\partial^2 f}{\partial x_k\partial x_j}(a).$$

D'où le résultat par unicité de la limite. ☐

Par récurrence, on peut montrer que si f est de classe C^k, alors l'ordre de dérivation des dérivées jusqu'à l'ordre k n'a pas d'importance.

La somme des dérivées secondes par rapport à chacune des variables joue un rôle très important dans beaucoup de problèmes physiques. Cela lui vaut un nom et une notation spéciale :

Définition 5.4. Soit f une fonction de classe C^2 de $\mathcal{U} \subset \mathbb{R}^n$ dans $\mathbb{R}$. On appelle laplacien de f la fonction définie sur $\mathcal{U}$ par

$$\Delta f : x \mapsto \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j^2}(x).$$

5.2 Formule de Taylor-Young à l'ordre 2

Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^n$, f une fonction de $\mathcal{U}$ dans $\mathbb{R}^p$, et $a \in \mathcal{U}$. Pour $h \in \mathbb{R}^n$ et $t \in \mathbb{R}$ assez petit on note

$$\varphi(t) = f(a + th). \tag{5.1}$$

On rappelle que si f est différentiable sur $\mathcal{U}$ alors φ est dérivable et

$$\varphi'(t) = d_a f(a + th) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a + th).$$

Si les dérivées partielles de f sont elle-mêmes différentiables, alors pour tout $j \in [\![1, n]\!]$ on obtient de même

$$\frac{d}{dt}\left(\frac{\partial f}{\partial x_j}(a + th)\right) = \sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k \partial x_j}(a + th).$$

Si f est de classe C^2, la fonction φ est donc elle-même de classe C^2 avec

$$\varphi''(t) = \sum_{j,k=1}^{n} h_j h_k \frac{\partial f}{\partial x_j \partial x_k}(a + th).$$

Plus généralement, si f est de classe C^k, on peut vérifier que φ est de classe C^k et calculer explicitement ses dérivées. Ainsi, en écrivant les formules de Taylor usuelles pour φ, on peut obtenir des formules analogues pour la fonction de plusieurs variables f. On utilisera cette idée dans la section suivante.

On commence par donner une preuve pour la formule de Taylor-Young à l'ordre 2 :

Théorème 5.5 (Taylor-Young à l'ordre 2)**.** *Soit f une fonction de classe C^2 de $\mathcal{U}$ dans $\mathbb{R}^p$. Soit $a \in \mathcal{U}$. Alors on a*

$$f(a + h) = f(a) + \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) + \frac{1}{2} \sum_{j,k=1}^{n} h_j h_k \frac{\partial^2 f}{\partial x_j \partial x_k}(a) + \underset{h \to 0}{o}\left(\|h\|^2\right).$$

Démonstration. Pour $h \in \mathbb{R}^n$ assez petit on note

$$R_2(h) = f(a + h) - f(a) - \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) - \frac{1}{2} \sum_{k,l=1}^{n} h_k h_l \frac{\partial^2 f}{\partial x_k \partial x_l}(a).$$

On doit montrer que

$$\frac{R_2(h)}{\|h\|^2} \xrightarrow[h \to 0]{} 0. \tag{5.2}$$

On rappelle que si g est une fonction différentiable en a on a pour $h \in \mathbb{R}^n$ assez petit

$$g(a + h) = g(a) + \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) + \underset{h \to 0}{o}\left(\|h\|\right).$$

Soit $j \in [\![1, n]\!]$. La dérivée partielle $\frac{\partial f}{\partial x_j}$ est différentiable en a, donc

$$\frac{\partial f}{\partial x_j}(a+h) - \frac{\partial f}{\partial x_j}(a) - \sum_{k=1}^{n} h_k \frac{\partial^2 f}{\partial x_k \partial x_j}(a) = \underset{h \to 0}{o}\left(\|h\| \right).$$

Or

$$\frac{\partial R_2}{\partial h_j}(h) = \frac{\partial f}{\partial x_j}(a+h) - \frac{\partial f}{\partial x_j}(a) - \sum_{k=1}^{n} h_k \frac{\partial^2 f}{\partial x_k \partial x_j}(a).$$

On en déduit que

$$|||d_h R_2||| = \underset{h \to 0}{o}\left(\|h\| \right).$$

Soit $\varepsilon > 0$. Il existe donc $\delta > 0$ tel que pour $h \in B(0, \delta)$ on a $|||d_h R_2||| \leqslant \varepsilon \|h\|$. D'après l'inégalité des accroissements finis on a alors

$$\|R_2(h)\| = \|R_2(h) - R_2(0)\| \leqslant \varepsilon \|h\| \, \|h\| = \varepsilon \|h\|^2 .$$

Cela prouve (5.2) et conclut la démonstration. □

Si f est une fonction de plusieurs variables à valeurs réelles, on regroupe des dérivées partielles dans un vecteur appelé gradient. Les dérivées secondes sont quant à elles collectées dans une matrice, appelée matrice Hessienne :

Définition 5.6. Soit f une fonction de classe C^2 de $\mathcal{U} \subset \mathbb{R}^n$ dans $\mathbb{R}$ et $a \in \mathcal{U}$. On appelle Hessienne de f en a la matrice (symétrique)

$$\mathrm{Hess}_a(f) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{1 \leqslant i,j \leqslant n} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) & \dots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix} \in M_n(\mathbb{R}).$$

Remarque 5.7. Si f admet des dérivées partielles secondes mais n'est pas de classe C^2, alors la matrice Hessienne peut être définie mais ce n'est plus nécessairement une matrice symétrique.

On peut alors ré-écrire la formule de Taylor-Young en utilisant la matrice Hessienne. Pour $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$ on a

$$f(a+h) = f(a) + d_a f(h) + \frac{{}^t h \cdot \mathrm{Hess}_a(f) \cdot h}{2} + \underset{h \to 0}{o}\left(\|h\|^2 \right).$$

5.3 Formules de Taylor à tout ordre

Si la fonction f de la partie précédente est en fait de classe C^k pour $k \in \mathbb{N}^*$, alors la fonction φ définie en (5.1) est de classe C^k et on a

$$\varphi^{(k)}(t) = \sum_{j_1, \dots, j_k = 1}^{n} h_{j_1} \dots h_{j_k} \frac{\partial^k f}{\partial x_{j_1} \dots \partial x_{j_k}}(a + th).$$

Pour $k \in \mathbb{N}$ et $h \in \mathbb{R}^n$ on note

$$D^k f(a)(h)^k = \sum_{j_1, \dots, j_k = 1}^{n} h_{j_1} \dots h_{j_k} \frac{\partial^k f}{\partial x_{j_1} \dots \partial x_{j_k}}(a).$$

Attention, cette notation est plus compacte (ce qui est sans aucun doute une bonne chose), mais pas forcément plus claire ...

Une autre notation possible exploite le fait que dans la somme précédente beaucoup de termes sont en fait égaux (par le théorème de Schwarz). Pour $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ on note $|\alpha| = \alpha_1 + \dots + \alpha_n$, $h^\alpha = h_1^{\alpha_1} \dots h_n^{\alpha_n}$ et

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x^\alpha} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_{n-1}}}{\partial x_{n-1}^{\alpha_{n-1}}} \frac{\partial^{\alpha_n} f}{\partial x_n^{\alpha_n}}.$$

On note également $\alpha! = \alpha_1! \dots \alpha_n!$. Pour $k \in \mathbb{N}$ on a alors :

$$\frac{1}{k!} D^k f(a)(h)^k = \sum_{|\alpha|=k} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha.$$

Une fois toutes ces notations absorbées, on peut énoncer les formules de Taylor. On donne ici la formule de Taylor avec reste intégral. Comme pour les fonctions d'une seule variable, c'est la plus précise.

Théorème 5.8 (Formule de Taylor avec reste intégral)**.** *Soit $k \in \mathbb{N}$. Soient $f : \mathcal{U} \to \mathbb{R}^p$ une fonction de classe C^{k+1}, $a \in \mathcal{U}$ et $h \in \mathbb{R}^n$ tels que le segment $[a, a+h]$ est contenu dans $\mathcal{U}$. Alors on a*

$$f(a+h) = f(a) + Df(a)(h) + \dots + \frac{1}{k!} D^k f(a)(h)^k + \int_0^1 \frac{(1-t)^k}{k!} D^{k+1} f(a+th)(h)^{k+1} \, dt.$$

Démonstration. Il suffit d'appliquer entre 0 et 1 la formule de Taylor avec reste intégral usuelle à la fonction φ définie par (5.1). □

On peut déduire la formule de Taylor-Young de la formule avec reste intégral en estimant simplement le reste. On observe qu'on a alors besoin d'une fonction de classe C^{k+1} au lieu de C^k. On peut donc préférer généraliser la démonstration du théorème 5.5 pour obtenir le résultat suivant.

Théorème 5.9 (Formule de Taylor-Young)**.** *Soit $f : \mathcal{U} \to \mathbb{R}$ une fonction de classe C^k et $a \in \mathcal{U}$. Alors pour $h \in \mathbb{R}^n$ on a*

$$f(a+h) = f(a) + Df(a)(h) + \dots + \frac{1}{k!} D^k f(a)(h)^k + \underset{h \to 0}{o}\left(\|h\|^k \right).$$

5.4 Application à l'étude des extremums locaux

Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^n$ et f une fonction de $\mathcal{U}$ dans $\mathbb{R}$. On cherche maintenant à trouver les extremums locaux de f.

Définition 5.10. Soit $a \in \mathcal{U}$.

- On dit que f admet un maximum (minimum) local en a s'il existe $r > 0$ tel que pour tout $x \in B(a, r)$ on a $f(x) \leqslant f(a)$ $(f(x) \geqslant f(a))$.
- On dit que f admet un maximum (minimum) local strict en a s'il existe $r > 0$ tel que pour tout $x \in B(a, r) \setminus \{a\}$ on a $f(x) < f(a)$ $(f(x) > f(a))$.
- On dit que f admet un extremum local en a si elle admet un maximum ou un minimum local en a.

Proposition 5.11. *On suppose que f est de classe C^1 et admet un extremum local en a. Alors $d_a f = 0$ ou, ce qui est équivalent, $\nabla f(a) = 0$.*

Démonstration. Soit $j \in [\![1, d]\!]$. L'application $t \mapsto f(a + te_j)$ est définie et dérivable au voisinage de 0 et admet un extremum en 0. Cela implique que sa dérivée en 0, c'est-à-dire $\frac{\partial f}{\partial x_j}(a)$, est nulle. □

Définition 5.12. On dit que f admet un point critique en a si $\nabla f(a) = 0$.

Remarque 5.13. Comme en dimension 1, le fait que a soit un point critique de f n'implique pas que f admet un extremum local en a.

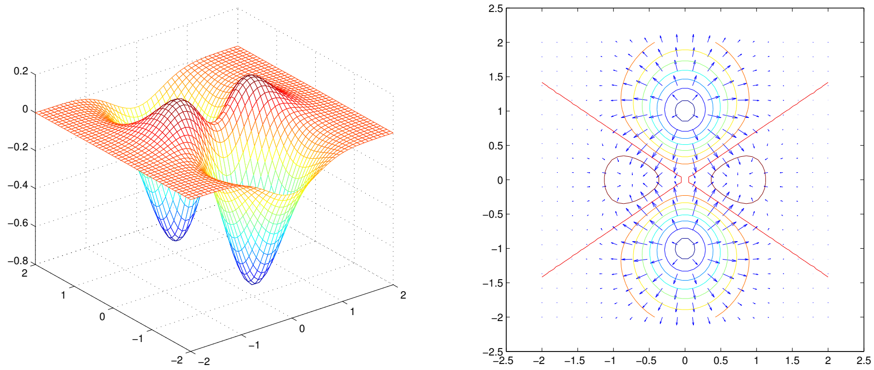

FIGURE 5.1 – La fonction $(x, y) \mapsto (x^2 - 2y^2)e^{-2x^2-y^2}$ admet deux maximums, deux minimums, et un point critique qui n'est pas un extremum local.

On suppose maintenant que f est de classe C^2 et que a est un point critique de f. D'après la formule de Taylor-Young on a

$$f(a + h) = f(a) + Q_a f(h) + \underset{h \to 0}{o}\left(\|h\|^2\right),$$

où on a noté

$$Q_a f(h) = \frac{{}^t h \cdot \mathrm{Hess}_a(f) \cdot h}{2}.$$

Le signe de la quantité $Q_a f(h)$ va nous permettre de déterminer si pour h petit $f(a+h)$ est plus petit ou plus grand que $f(a)$, ou si les deux cas se présentent.

Définition 5.14. On dit que (la forme quadratique) Q_a est
- positive si $Q_a(h) \geqslant 0$ pour tout $h \in \mathbb{R}^n$,
- négative si $Q_a(h) \leqslant 0$ pour tout $h \in \mathbb{R}^n$,
- définie positive si $Q_a(h) > 0$ pour tout $h \in \mathbb{R}^n \setminus \{0\}$,
- définie négative si $Q_a(h) < 0$ pour tout $h \in \mathbb{R}^n \setminus \{0\}$.

On observe que la forme quadratique Q_a peut n'être ni positive ni négative. La proposition suivante donne une condition nécessaire pour qu'un point critique soit un minimum ou un maximum local.

Proposition 5.15. *Soit a un point critique de f.*
- *Si f admet un minimum local en a, alors $Q_a(f)$ est positive.*
- *Si f admet un maximum local en a, alors $Q_a(f)$ est négative.*

Démonstration. On montre la première propriété. La deuxième s'obtient en appliquant la première à $-f$. On suppose donc que f admet un minimum local en a et on considère $h \in \mathbb{R}^n$. Pour $t \in \mathbb{R}$ assez petit on a

$$f(a + th) - f(a) \geqslant 0.$$

Or

$$\frac{1}{t^2}\big(f(a+th)-f(a)\big) = Q_a(f)(h) + \underset{t\to 0}{o}(1) \xrightarrow[t\to 0]{} Q_a(f)(h).$$

Cela prouve que $Q_a(f)(h) \geqslant 0$. □

***Exercice* 5.2.** Montrer que l'application $(x,y) \mapsto x^2 - y^2$ admet un point critique mais n'admet aucun extremum local (voir figure 5.2)

Malheureusement la réciproque de la proposition 5.15 n'est pas vraie. Ainsi cette proposition permet de dire qu'un point critique n'est pas un maximum local ou pas un minimum local, mais elle ne donne pas de conclusion positive.

On admet pour ce cours le résultat suivant :

Proposition 5.16. *La matrice* $\mathrm{Hess}_a f$ *est orthodiagonalisable : cela signifie qu'il existe une matrice diagonale réelle* $D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$ *et* $P \in GL_n(\mathbb{R})$ *tels que* $P^{-1} = {}^tP$ *et*

$$\mathrm{Hess}_a(f) = P^{-1}DP = {}^tPDP.$$

Avec ce résultat on est capable de donner une condition suffisante pour qu'un point critique soit un minimum ou un maximum local :

Proposition 5.17. *Soit a un point critique de f.*
- *Si $Q_a(f)$ est définie positive, alors f admet un minimum local strict en a.*
- *Si $Q_a(f)$ est définie négative, alors f admet un maximum local strict en a.*

Remarque 5.18. Si $Q_a(f)$ est positive mais pas définie positive, ou négative mais pas définie négative, alors ni la proposition 5.15 ni la proposition 5.17 ne permettent de dire si f admet ou non un extrumum en a.

Démonstration. On suppose que $Q_a(f)$ est définie positive. On a $D = P\,\mathrm{Hess}_a(f)\,{}^tP$. Donc pour tout $j \in [\![1,n]\!]$ on a

$$\lambda_j = {}^te_jDe_j = {}^t({}^tPe_j)\,\mathrm{Hess}_a(f)\,{}^tPe_j = 2Q_a(f)({}^tPe_j) > 0.$$

On note $\lambda = \min(\lambda_1,\dots,\lambda_n) > 0$. Soit $h \in \mathbb{R}^n$. On note $Ph = \begin{pmatrix} \tilde{h}_1 \\ \vdots \\ \tilde{h}_n \end{pmatrix}$. On a alors

$$\begin{aligned} Q_a(f)(h) = {}^th\,\mathrm{Hess}_a(f)h = {}^t(Ph)D(Ph) = \sum_{j=1}^n \lambda_j(\tilde{h}_j)^2 \\ \geqslant \lambda\sum_{j=1}^n(\tilde{h}_j)^2 = \lambda\,{}^t(Ph)\cdot(Ph) = \lambda\,{}^th\cdot h = \lambda\,\|h\|_2^2. \end{aligned}$$

Et comme

$$f(a+h) - f(a) = Q_a(f)(h) + \underset{h\to 0}{o}\big(\|h\|_2^2\big),$$

il existe $r > 0$ tel que $f(a+h) - f(h) > 0$ pour tout $h \in B(0,r)\setminus\{0\}$, et donc f admet un minimum local strict en a. □

Définition 5.19. Soit a un point critique de f. On dit que a est un point selle de f si $Q_a(f)(h)$ prend des valeurs strictement positives et strictement négatives.

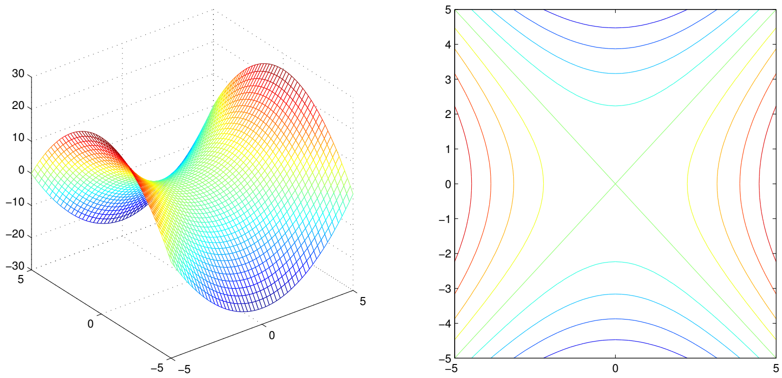

FIGURE 5.2 – Le point (0,0) est un point selle pour la fonction $(x, y) \mapsto x^2 - y^2$.

Remarque 5.20. Si a est un point selle de f, alors f n'admet pas d'extremum local en a.

Proposition 5.21. *On suppose que* $n = 2$. *Soit* a *un point critique de* f.
- *Si* $\det \mathrm{Hess}_a(f) < 0$, *alors* a *est un point selle de* f.
- *Si* $\det \mathrm{Hess}_a(f) > 0$, *alors* a *admet un extremum local strict en* a.
  - *Si* $\Delta f(a) \geqslant 0$, *c'est un minimum local.*
  - *Si* $\Delta f(a) \leqslant 0$, *c'est un maximum local.*

Démonstration. On note λ_1 et λ_2 les valeurs propres de $\mathrm{Hess}_a(f)$. Alors avec les notations précédentes on a $\det \mathrm{Hess}_a(f) = \det D = \lambda_1 \lambda_2$ et $\Delta f(a) = \mathrm{Tr}\, \mathrm{Hess}_a(f) = \mathrm{Tr}\, D = \lambda_1 + \lambda_2$. Si $\lambda_1 \lambda_2 < 0$, on en déduit que $Q_a(f)(Pe_1) = \lambda_1$ et $Q_a(f)(Pe_2) = \lambda_2$ sont non nuls et de signes distincts. Le point a est donc un point selle de f. Si $\lambda_1 \lambda_2 > 0$, les deux valeurs propres sont non nulles et de même signe, positif si $\Delta f(a) \geqslant 0$ et négatif si $\Delta f(a) \geqslant 0$. Le calcul de la preuve précédente montre que f admet alors en f un minimum local strict (respectivement maximum local strict). □

5.5 Exercices

***Exercice* 5.3.** Étudier sur $\mathbb{R}^2$ les extrema locaux et globaux des fonctions définies par

$$f_1 : (x, y) \mapsto x^4 + y^4 \quad ; \quad f_2 : (x, y) \mapsto x^4 - y^4 \quad ; \quad f_3 : (x, y) \mapsto x^4 \quad ; \quad f_4 : (x, y) \mapsto -(x^4 + y^4).$$

***Exercice* 5.4.** Étudier les extrema locaux et globaux des fonctions définies par
- $f_1(x, y) = x^3 + y^3 - 3xy$ sur $\mathbb{R}^2$,
- $f_2(x, y) = x^4 + y^4 - 4xy$ sur $\mathbb{R}^2$,
- $f_3(x, y) = y(x^2 + (\ln y)^2)$ sur $\mathbb{R} \times \mathbb{R}_+^*$,
- $f_4(x, y) = y^2 + x^2 + xy + 2x - 2y$ sur $\mathbb{R}^2$.

***Exercice* 5.5.** Soit $A \in M_n(\mathbb{R})$ une matrice symétrique définie positive et $b \in \mathbb{R}^n$. On admet (ou pas) que l'application

$$\phi : x \mapsto \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle$$

admet un minimum global sur $\mathbb{R}^n$. Montrer que ce minimum est atteint au point $A^{-1}b$.

***Exercice* 5.6** (Contre-exemple pour le théorème de Schwarz)**.** On considère l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie pour $(x, y) \in \mathbb{R}^2$ par

$$f(x, y) = \begin{cases} y^2 \sin\left(\frac{x}{y}\right) & \text{si } y \neq 0 \\ 0 & \text{sinon.} \end{cases}$$

1. Étudier la continuité de f.
2. Étudier l'existence et éventuellement la valeur des dérivées premières et secondes de f.
3. Déterminer le plus grand domaine de $\mathbb{R}^2$ sur lequel f est C^1, C^2.

***Exercice* 5.7.** Soit f une fonction de classe C^2 de $\mathbb{R}$ dans $\mathbb{R}$ telle que $f(0) = 0$ et $f'(0) \neq 0$. Pour $(x, y) \in \mathbb{R}^2$ on pose $F(x, y) = f(x)f(y)$.
1. Montrer que F est de classe C^2 sur $\mathbb{R}^2$.
2. La fonction F admet-elle un extremum local en (0,0) ?

***Exercice* 5.8.** Déterminer

$$\inf_{\substack{x>0 \\ y>0}} \left(\frac{1}{x} + \frac{1}{y} + xy \right).$$