6 Cols, collines et étangs : les points critiques et extrema locaux

6.1 Définitions

Définition 6.1 (**Extremum local**). *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et $x_0 \in E$. On dit que f admet un minimum (resp. maximum) local en x_0 s'il existe un ouvert $\Omega \subset \mathbb{R}^n$ tel que*

$$\forall x \in \Omega \cap E, \quad f(x_0) \leq f(x) \quad (\text{resp. } f(x_0) \geq f(x)).$$

On dit que ce minimum (resp. maximum) est un minimum local strict s'il existe $\Omega \subset \mathbb{R}^n$ tel que

$$\forall x \in (\Omega \cap E) \setminus \{x_0\}, \quad f(x_0) < f(x) \quad (\text{resp. } f(x_0) > f(x)).$$

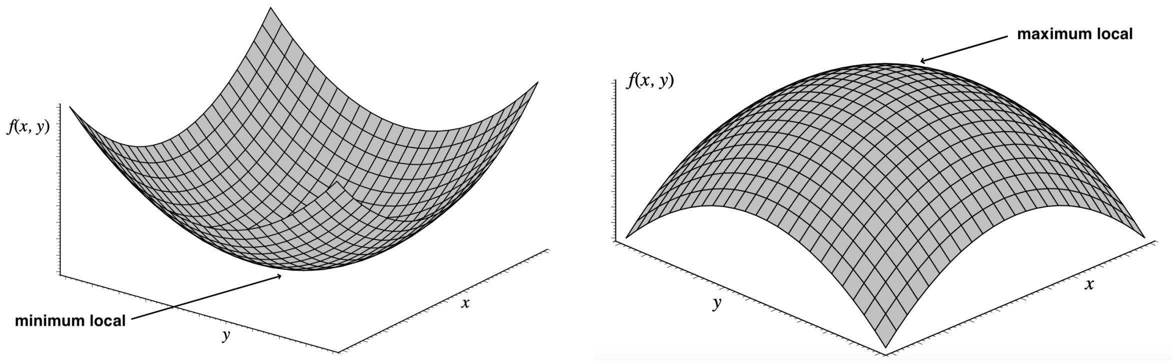

FIGURE 12 : Exemples de minimum et maximum locaux stricts

Remarque 6.2. On a donc ici deux cas importants :

- si x_0 est un point intérieur à E, alors Ω peut être considéré comme inclus dans E (quitte à le modifier), et $D_{x_0} f$ a un sens.
- si x_0 n'est pas un point intérieur à E, c'est-à-dire que l'on ne peut pas y centrer une boule ouverte incluse dans E, alors on a nécessairement que, pour tout ouvert $\Omega \subset \mathbb{R}^n$ contenant x_0, $\Omega \cap E^c \neq \emptyset$, c'est-à-dire que $x_0 \in \partial E$, et $D_{x_0} f$ n'a pas de sens ici.

Définition 6.3 (**Point critique/stationnaire**). *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et x_0 un point intérieur à E. On suppose que f est différentiable en x_0. Alors on dit que x_0 est un point critique (ou stationnaire) de f si $\nabla_{x_0} f = 0$.*

Définition 6.4 (**Point selle/col**). *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et x_0 un point intérieur à E. On suppose que f est différentiable en x_0 et que x_0 est un point critique de f. On dit que x_0 est un point selle (ou col) de f si, pour tout voisinage V de x_0, il existe $(x_1, x_2) \in (V \cap E)^2$, $f(x_1) > f(x_0)$ et $f(x_2) < f(x_0)$.*

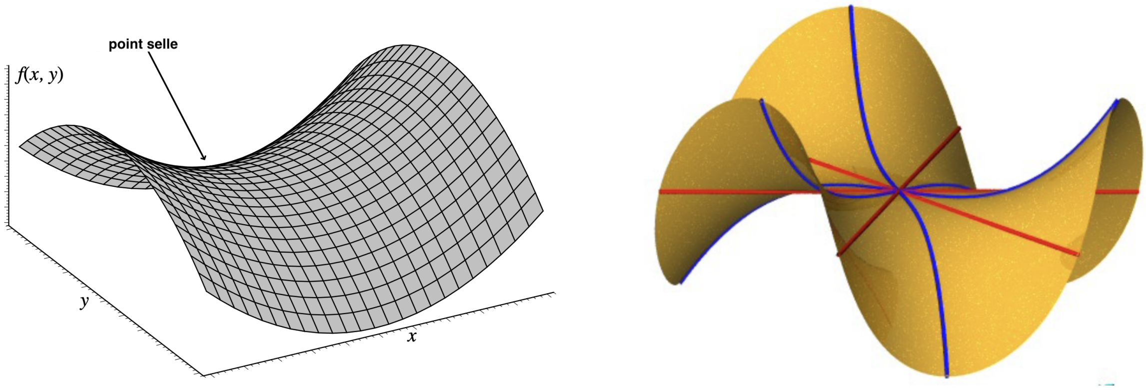

FIGURE 13 : Exemples de points selles

6.2 Propriétés du vecteur gradient

Considérons une fonction $f : E \subset \mathbb{R}^n \to \mathbb{R}$ continue et x_0 un point intérieur à E tel que f est différentiable en x_0. Alors :

1. **Le gradient $\nabla_{x_0} f$ donne la direction le long de laquelle f varie le plus vite à partir de x_0.** En effet, si on cherche $v \in \mathbb{R}^n \backslash \{0\}$ de norme fixée (sinon le problème suivant n'a pas de solution) tel que $\left|\frac{\partial f}{\partial v}(x_0)\right|$ soit maximale, alors comme, par Cauchy-Schwarz,
$$\left|\frac{\partial f}{\partial v}(x_0)\right| = \left|\langle \nabla_{x_0} f, v\rangle\right| \leq \|v\|_2 \|\nabla_{x_0} f\|_2$$
avec égalité si et seulement si v et $\nabla_{x_0} f$ sont colinéaires, on en déduit que $v = \lambda \nabla_{x_0} f$ avec $\lambda \neq 0$ de telle sorte que $\frac{\partial f}{\partial v}(x_0) = \lambda \|\nabla_{x_0} f\|_2^2$ et donc que la direction de plus forte croissance soit celle de $\nabla_{x_0} f$ et celle de plus forte décroissance soit celle de $-\nabla_{x_0} f$.

2. **Plus la norme du gradient est grande, plus la fonction varie vite dans sa direction.** Cela vient du fait que,d ans la direction optimale v déterminée au point précédent, $\frac{\partial f}{\partial v}(x_0) = \lambda \|\nabla_{x_0} f\|_2^2$ avec $\lambda \neq 0$.

3. **Le gradient en x_0 est orthogonal à toute courbe de niveau passant par x_0.** Soit L une courbe de niveau de f, c'est-à-dire un ensemble de type $\{x \in \Omega : f(x) = \ell\}$, telle que $x_0 \in L$. Soit v un vecteur tangent à L, alors comme f est constante sur L $(= \ell)$, on a
$$0 = \frac{\partial f}{\partial v}(x_0) = \langle \nabla_{x_0} f, v\rangle.$$
Ainsi, $\nabla_{x_0} f$ est orthogonal à tout vecteur tangent à L, il est donc orthogonal à L.

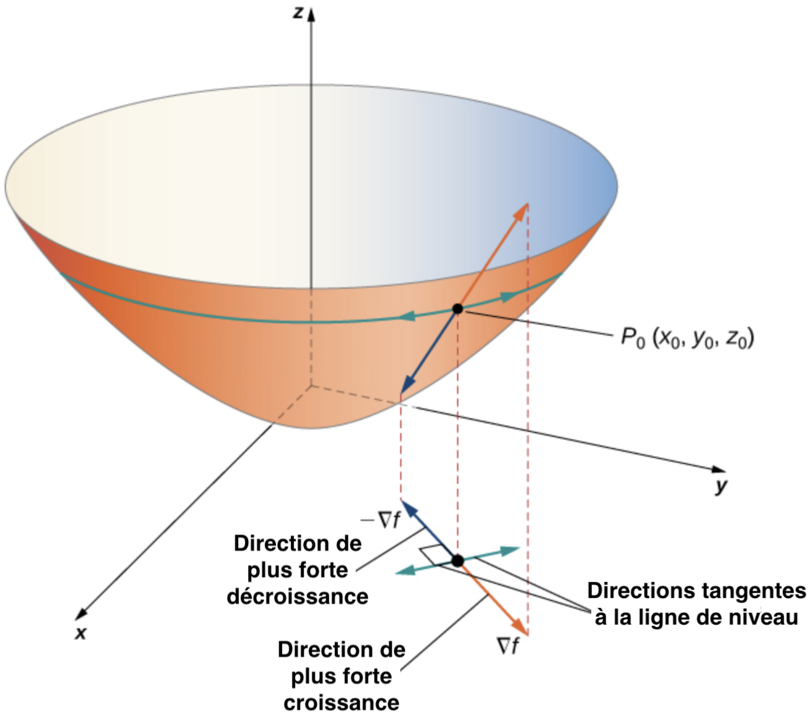

FIGURE 14 : Illustration des trois propriétés pour $f : \mathbb{R}^2 \to \mathbb{R}$ au point (x_0, y_0) et $z_0 = f(x_0, y_0)$.

6.3 Conditions Nécessaires et Suffisantes d'ordre 1 et 2

Proposition 6.5 (Condition nécessaire d'ordre 1 pour un extremum). *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et x_0 est un point intérieur à E. Si f est différentiable en x_0 et admet en x_0 un extremum local (minimum ou maximum), alors x_0 est un point critique de f.*

Démonstration. Comme x_0 est un extremum local de f, les n fonctions $g_i : t \mapsto f(x_0 + te_i)$ admettent un extremum local en $t = 0$, donc leurs dérivées en 0, qui sont les dérivées partielles de f en x_0 sont toutes nulles. On a donc bien $\nabla_{x_0} f = 0$ et x_0 est un point critique de f. □

Remarque 6.6. Si $x_0 \notin \mathring{E}$ est un extremum local de f, on n'a pas nécessairement $\nabla_{x_0} f = 0$. Pensez par exemple à la fonction $f : [-1, 1] \to \mathbb{R}$, $x \mapsto x^2$. Alors f admet son maximum en -1 et 1 mais $f'(1^-) = 2 \neq 0$ et $f'(-1^+) = -2 \neq 0$. Par contre 0 est intérieur à $[-1, 1]$ et est le minimum de f sur $[-1, 1]$, donc $f'(0) = 0$, ce qui est effectivement le cas.

Définition 6.7 (Matrice/forme quadratique (définie) positive). *Soit $q : \mathbb{R}^n \to \mathbb{R}$ une forme quadratique et $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ sa matrice symétrique réelle associée. On dit que*

- *q et A sont positives si $\forall x \in \mathbb{R}^n$, $q(x) = \langle Ax, x\rangle \geq 0$;*
- *q et A sont définies positives si $\forall x \in \mathbb{R}^n \setminus \{0\}$, $q(x) = \langle Ax, x\rangle > 0$;*
- *q et A sont (définies) négatives si $-q$ et $-A$ sont (définies) positives.*

Remarque 6.8. On rappelle que, pour tout $x \in \mathbb{R}^n$ et toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, ${}^t x A x = \langle Ax, x\rangle$.

Proposition 6.9 (**La hessienne est diagonalisable**)**.** *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et x_0 un point intérieur à E. Si f est deux fois différentiable en x_0, alors $H_f(x_0)$ est diagonalisable.*

Démonstration. C'est évident car $H_f(x_0)$ est une matrice symétrique réelle, donc diagonalisable. □

Proposition 6.10 (**Positivité et valeurs propres**)**.** *Une matrice A symétrique réelle est (définie) positive si et seulement si toutes ses valeurs propres sont (strictement) positives.*

Démonstration. La matrice A est symétrique réelle, donc diagonalisable. Soit $(\lambda_1, \dots, \lambda_n)$ les valeurs propres de A (comptées sans leur multiplicité). Soit $(u_1, \dots, u_n)$ une base orthonormée de vecteurs propres de A, alors pour tout $x \in \mathbb{R}^n$, on a $x = \sum_{i=1}^n x_i u_i$ avec $(x_1, \dots, x_n) \in \mathbb{R}^n$, et ainsi, comme, pour tout $(i, j) \in \{1, \dots, n\}^2$, $\langle u_i, u_j \rangle = \delta_{i,j}$, on a

$$\langle Ax, x \rangle = \left\langle \sum_{i=1}^n x_i \lambda_i u_i, \sum_{i=1}^n x_i u_i \right\rangle = \langle \lambda_1 x_1 u_1 + \dots + \lambda_n x_n u_n, x_1 u_1 + \dots + x_n u_n \rangle = \sum_{i=1}^k \lambda_i x_i^2.$$

On a donc

$$\forall x \in \mathbb{R}^n, \langle Ax, x \rangle \geq 0 \iff \forall 1 \leq i \leq n, \lambda_i \geq 0,$$

ainsi que

$$\forall x \in \mathbb{R}^n \backslash \{0\}, \langle Ax, x \rangle > 0 \iff \forall 1 \leq i \leq n, \lambda_i > 0.$$

□

Proposition 6.11 (**Condition nécessaire d'ordre 2 pour un extremum**)**.** *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et x_0 un point intérieur à E. Si f est deux fois différentiable en x_0 et x_0 est un point critique de f, alors :*

1. *si f admet en x_0 un minimum local, alors la forme quadratique $h \mapsto D^2_{x_0} f(h, h)$ (et donc $H_f(x_0)$) est positive.*
2. *si f admet en x_0 un maximum local, alors la forme quadratique $h \mapsto D^2_{x_0} f(h, h)$ (et donc $H_f(x_0)$) est négative.*

Remarque 6.12 (Point selle)**.** Si la forme quadratique $h \mapsto D^2_{x_0} f(h, h)$ (ou la matrice $H_f(x_0)$) n'est ni positive, ni négative, le point x_0 n'est ni un maximum, ni un minimum de f, il s'agit d'un point selle de f et $H_f(x_0)$ admet au moins une valeur propre strictement positive et un autre strictement négative.

Démonstration. Supposons que f admette un minimum local en x_0. Alors il existe un ouvert $\Omega \subset E$ tel que $f(x_0)$ soit le minimum de f sur Ω, et donc pour tout $h \in \mathbb{R}^n$ et $t \in \mathbb{R}$ suffisamment petit tel que $x_0 + th \in \Omega$, on a

$$f(x_0 + th) \geq f(x_0).$$

Supposons que $D^2_{x_0} f \neq 0$, sinon le résultat est évident car $H_f(x_0)$ sera à la fois positive et négative. Alors, comme f est deux fois différentiable en $x_0 \in \mathring{E}$, d'après la formule de Taylor-Young à l'ordre 2, on a, comme $o(\| th \|^2) = o(\| h \|^2 t^2) = o(t^2)$ quand $t \to 0$,

$$f(x_0 + th) = f(x_0) + t D_{x_0} f(h) + \frac{t^2}{2} D^2_{x_0} f(h, h) + o(t^2) = f(x_0) + \frac{t^2}{2} D^2_{x_0} f(h, h) + o(t^2).$$

Ainsi, on a, quand $t \to 0,\ t \neq 0$,

$$0 \leq \frac{f(x_0 + th) - f(x_0)}{t^2} = \frac{1}{2} D^2_{x_0} f(h,h) + o(1) \to \frac{1}{2} D^2_{x_0} f(h,h),$$

et donc $D^2_{x_0} f$ est positive. On montre de façon analogue le deuxième point, dans le cas d'un maximum local. □

Lemme 6.13 (Forme quadratique définie positive et première valeur propre). *Soit $q : \mathbb{R}^n \to \mathbb{R}$ une forme quadratique, A sa matrice associée et λ_1 la plus petite valeur propre de A. Alors on a*

$$\min_{\substack{x \in \mathbb{R}^n, \\ \|x\|_2 = 1}} \langle Ax, x \rangle = \lambda_1.$$

Démonstration. Tout d'abord, on sait que ce minimum est atteint car $x \mapsto \langle Ax, x\rangle$ est une fonction continue et la sphère pour la norme euclidienne $\{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ est un compact. Considérons une base orthonormée $(u_1, \dots, u_n)$ de vecteurs propres de A. Quitte à renuméroter, on suppose que les valeurs propres de A (comptées sans leur multiplicité) sont $\lambda_1 \leq \dots \leq \lambda_n$. Pour tout $x \in \mathbb{R}^n$, on peut écrire $x = \sum_{i=1}^n x_i u_i$ et donc $\langle Ax, x \rangle = \sum_{i=1}^n \lambda_i x_i^2$. On obtient pour tout $x \in \mathbb{R}^n$ tel que $\|x\|_2 = 1$, c'est-à-dire $\sum\limits_{i=1}^n x_i^2 = 1$,

$$\langle Ax, x \rangle = \sum_{i=1}^n \lambda_i x_i^2 \geq \lambda_1 \sum_{i=1}^n x_i^2 = \lambda_1,$$

atteint pour $x = (1, 0, \dots, 0)$, ce qui prouve le résultat souhaité. □

Proposition 6.14 (Condition suffisante d'ordre 2 pour un extremum local). *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ et x_0 un point intérieur à E. Si f est deux fois différentiable en x_0 et x_0 est un point critique de f, alors :*

1. *si la forme quadratique $h \mapsto D^2_{x_0} f(h,h)$ (ou bien $H_f(x_0)$) est définie positive, alors f admet un minimum local strict en x_0.*
2. *si la forme quadratique $h \mapsto D^2_{x_0} f(h,h)$ (ou bien $H_f(x_0)$) est définie négative, alors f admet un maximum local strict en x_0.*

Démonstration. Supposons que $h \mapsto D^2_{x_0} f(h,h)$ soit définie positive. Les valeurs propres de la matrice hessienne $H_f(x_0)$ sont donc toutes strictement positives. Soit λ_1 la plus petite de ces valeurs propres, alors d'après le lemme précédent, on a

$$\inf_{h \in \mathbb{R}^n : \|h\|_2 = 1} D^2_{x_0} f(h,h) = \lambda_1 > 0.$$

Or, d'après la formule de Taylor-Young à l'ordre 2, on a, avec ε une fonction de limite nulle quand $h \to 0,\ h \neq 0$,

$$f(x_0 + h) - f(x_0) = \frac{1}{2} D^2_{x_0} f(h,h) + \|h\|^2 \varepsilon(h) = \|h\|^2 \left(\frac{1}{2} D^2_{x_0} f\left(\frac{h}{\|h\|}, \frac{h}{\|h\|} \right) + \varepsilon(h) \right),$$

car x_0 est un point critique et par bilinéarité de $D^2_{x_0} f$. Ainsi, si $h \neq 0$ est suffisamment petit, $|\varepsilon(h)| < \frac{\lambda_1}{2}$ et donc, comme $\frac{h}{\|h\|} \in S_{\|\cdot\|_2}(0,1)$,

$$f(x_0 + h) - f(x_0) = \|h\|^2 \left(\frac{1}{2} D^2_{x_0} f\left(\frac{h}{\|h\|}, \frac{h}{\|h\|} \right) + \varepsilon(h) \right) > \|h\|^2 \left(\frac{1}{2} D^2_{x_0} f\left(\frac{h}{\|h\|}, \frac{h}{\|h\|} \right) - \frac{\lambda_1}{2} \right) > 0,$$

ce qui veut dire que f admet un minimum local strict en x_0 (on a égalité si et seulement si $h = 0$). On démontre de manière analogue le cas du maximum local strict. $\square$

6.4 Position par rapport au plan tangent

On considère de nouveau ici un ouvert $\Omega \subset \mathbb{R}^2$ une fonction $f : \Omega \to \mathbb{R}$ de classe C^2 sur Ω et $(x_0, y_0) \in \Omega$. Dans ce cas, le graphe de f

$$\mathscr{G}_f := \{(x, y, f(x, y)) : (x, y) \in \Omega\} \subset \mathbb{R}^3$$

est une nappe/surface de $\mathbb{R}^3$ d'équation $z = f(x, y)$. On note de nouveau

$$P : (x, y) \mapsto f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

l'approximation linéaire de f au voisinage de (x_0, y_0), $X_0 = (x_0, y_0, z_0)$ et $z_0 = f(x_0, y_0)$.

On a les résultats suivants (qui découlent de ce que l'on a déjà démontré) :

- Si (x_0, y_0) est un point critique de f, alors P est la fonction constante égale à $f(x_0, y_0)$ et le plan tangent en X_0 à $\mathscr{G}_f$ a pour équation $z = f(x_0, y_0)$, c'est un plan horizontal, c'est-à-dire parallèle au plan (xOy).
- Si (x_0, y_0) est un point critique de f et $H_f(x_0, y_0)$ est définie positive, alors $\mathscr{G}_f$ est localement, au voisinage de X_0, au-dessus de ce plan tangent;
- Si (x_0, y_0) est un point critique de f et $H_f(x_0, y_0)$ est définie négative, alors $\mathscr{G}_f$ est localement, au voisinage de X_0, en-dessous de ce plan tangent;
- Si (x_0, y_0) est un point selle de f, $\mathscr{G}_f$ traverse localement, sur tout voisinage de X_0, ce plan tangent.

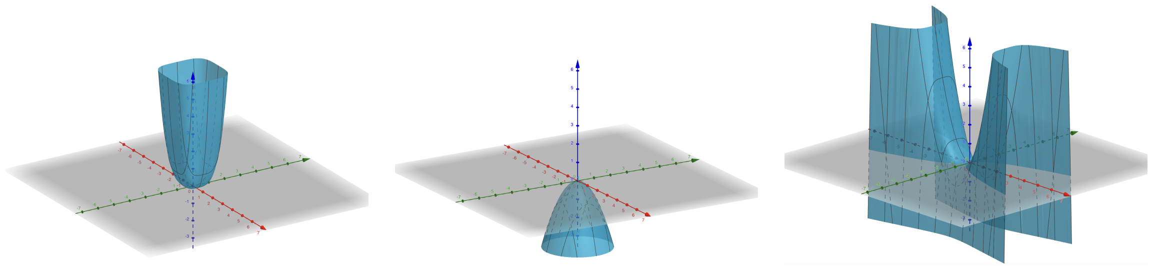

FIGURE 15 : Minimum local, maximum local, point selle et plan tangent en $(0, 0)$, pour les fonctions $(x, y) \mapsto x^4 + y^4$, $(x, y) \mapsto -x^2 - y^2$ et $(x, y) \mapsto x^2 - y^4$.

6.5 Optimalité globale

On termine cette partie par un résultat de minimalité/maximalité globale.

Proposition 6.15 (Conditions suffisantes de minimalité globale). *Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$ continue avec E un fermé. On suppose que f admet $N \in \mathbb{N}^*$ minima locaux $\{x_1, \ldots, x_N\} \subset E$. Alors*

$$\min_{x \in E} f(x) = \min_{1 \leq i \leq N} f(x_i)$$

et les x_i qui vérifient cette égalité sont des minima globaux de f si l'une des conditions suivantes est réalisée :

1. *E est compact;*
2. *E est non-borné et f coercive, c'est-à-dire $\lim\limits_{\substack{\|x\|_2 \to +\infty \\ x \in E}} f(x) = +\infty$;*
3. *$\lim\limits_{\substack{\|x\|_2 \to +\infty \\ x \in E}} f(x) = a \in \mathbb{R}$ et il existe $x_0 \in E$ tel que $f(x_0) < a$.*

Démonstration. Si E est compact, c'est évident. Si f est coercive, alors soit $a \in E$ tel que $f(a) \geq \min\limits_{1 \leq i \leq N} f(x_i)$, il existe donc $R > 0$ tel que, si $x \in E$ et $\|x\|_2 > R$, alors $f(x) > f(a)$. Maintenant, sur le compact $E \cap \overline{B}_{\|\cdot\|_2}(0, R)$, f admet un minimum par le théorème de Weierstrass, qui sera donc $\min\limits_{1 \leq i \leq N} f(x_i) \leq f(a)$, d'où le résultat puisque le minimum de f ne peut donc pas être sur $\overline{B}_{\|\cdot\|_2}(0, R)^c$.
Dans le dernier cas, la limite permet de conclure que f est bornée, et l'existence de x_0 assure que f atteint bien sa borne inférieure (sinon elle pourrait être simplement atteinte asymptotiquement en l'infini), puisque il existe $R > 0$ tel que, pour tout $x \in E$, $\|x\| > R \Rightarrow f(x) > f(x_0)$ et donc f atteint son minimum sur le compact $E \cap \overline{B}_{\|\cdot\|_2}(0, R)$ d'après le théorème de Weierstrass. □