Chapitre 1

Fonctions et optimisation

En sciences, on utilise très souvent le concept de fonction, et en particulier le concept de fonctions à plusieurs variables. Par exemple, on peut considérer la température T dans une pièce comme fonction des trois variables d'espace x, y, z. Dans ce chapitre, nous commençons par rappeler le concept de fonction à une variable réelle, puis nous introduisons le concept de fonction à plusieurs variables réelles. Nous nous intéressons en particulier au problème très courant d'optimisation, c'est-à-dire la recherche de minimums ou de maximums d'une fonction.

1.1 Fonction d'une variable réelle

1.1.1 Définition

Définition 1. *(Fonction d'une variable réelle).* Une fonction f d'un ensemble $D \subset \mathbb{R}$ dans $\mathbb{R}$ associe à tout nombre réel $x \in D$ un nombre réel noté $f(x)$. On note :

$$\begin{aligned} f : D &\to \mathbb{R} \\ x &\mapsto f(x). \end{aligned}$$

L'ensemble de départ D est appelé domaine de définition de f.

Par abus de langage, on dit souvent que « $f(x)$ est une fonction ». Mais, en toute rigueur, $f(x)$ n'est pas une fonction mais la valeur de la fonction f au point x. Si on veut parler d'une fonction (un objet défini pour plusieurs valeurs de x), il est plus correct d'utiliser les notations f ou $x \mapsto f(x)$. Par exemple, il est plus correct de dire que « la fonction racine carrée est $x \mapsto \sqrt{x}$ » que de dire « la fonction racine carrée est $\sqrt{x}$ ».

1.1.2 Continuité et dérivabilité

Définition 2. *(Continuité).* On dit qu'une fonction $f : D \to \mathbb{R}$ est continue au point $x_0 \in D$ si

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Lorsqu'une fonction n'est pas continue en x_0, on dit qu'elle est discontinue en x_0. Si f est continue en tout point x_0 de D, alors on dit que f est continue sur le domaine D, ou plus simplement que f est continue. Pour faire simple, une fonction continue est telle que son graphe peut être tracé sans lever le crayon de la feuille.

Exemples :

— La fonction $f_1 : x \mapsto x^2$ est continue sur $\mathbb{R}$.

— La fonction $f_2 : x \mapsto |x|$ est continue sur $\mathbb{R}$.

— La fonction $f_3 : x \mapsto \begin{cases} 0 \text{ si } x < 0 \\ 1 \text{ si } x \geq 0 \end{cases}$ est discontinue en $x_0 = 0$.

Définition 3. *(Dérivabilité).* On dit qu'une fonction $f : D \to \mathbb{R}$ est dérivable en $x_0 \in D$ si le taux d'accroissement de f admet une limite finie quand $x \to x_0$

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

Cette limite est notée $f'(x_0)$ et s'appelle la dérivée de f en x_0.

Si cette limite est infinie ou si cette limite n'existe pas (par exemple, les limites à gauche et à droite sont différentes), alors f n'est pas dérivable en x_0.
Une formule alternative souvent utilisée est :

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0).$$

Si f est dérivable en tout point x_0 de D, alors on dit que f est dérivable sur le domaine D, ou plus simplement que f est dérivable.

Théorème 1. *(Tangente).* Soit $f : D \to \mathbb{R}$ une fonction dérivable. L'équation de la tangente à la courbe représentative de f en x_0 est : $y = f(x_0) + f'(x_0)(x - x_0)$.

Il s'agit simplement du développement limité de f en $x = x_0$ à l'ordre 1. La dérivée $f'(x_0)$ est donc le coefficient directeur (pente) de la droite tangente à la courbe de f à $x = x_0$.

Exemples :

— La fonction $f_1 : x \mapsto x^2$ est dérivable sur $\mathbb{R}$.

— La fonction $f_2 : x \mapsto |x|$ n'est pas dérivable en $x_0 = 0$.

— La fonction $f_3 : x \mapsto \begin{cases} 0 \text{ si } x < 0 \\ 1 \text{ si } x \geq 0 \end{cases}$ n'est pas dérivable en $x_0 = 0$.

Théorème 2. *(Dérivable $\Rightarrow$ continue).* Une fonction $f : D \to \mathbb{R}$ dérivable en $x_0 \in D$ est forcément continue en x_0.

La réciproque est fausse. Par exemple, la fonction $x \mapsto |x|$ est continue en $x_0 = 0$ mais pas dérivable en $x_0 = 0$.

La fonction $f' : x \mapsto f'(x)$ est appelée dérivée de f. On peut continuer à dériver, et on définit ainsi :

— la dérivée seconde, notée f'', comme la dérivée de la dérivée f', c'est-à-dire $f'' = (f')'$;

— la dérivée troisième, notée f''', comme la dérivée de la dérivée seconde f'', c'est-à-dire $f''' = (f'')'$;
— et, de manière générale, la dérivée d'ordre n (où n est un entier positif ou nul), notée $f^{(n)}$, comme la dérivée de la dérivée d'ordre $n-1$, c'est-à-dire $f^{(n)} = (f^{(n-1)})'$ pour $n \geq 1$. Par convention, la dérivée d'ordre 0 est juste la fonction $f^{(0)} = f$. Ainsi, la dérivée d'ordre 1 est $f^{(1)} = f'$, la dérivée d'ordre 2 est $f^{(2)} = f''$, etc.

La notation f', f'', f''',... ou $f^{(1)}$, $f^{(2)}$, $f^{(3)}$, ... pour les dérivées est appelée notation de Lagrange. C'est la notation la plus commune en mathématiques. En sciences, on utilise aussi souvent la notation de Leibniz

$$\frac{df}{dx}, \frac{d^2 f}{dx^2}, \frac{d^3 f}{dx^3}, \ldots$$

qui a le mérite de rappeler la définition de la dérivée avec le taux d'accroissement. Pour les fonctions du temps, $f : t \mapsto f(t)$, on utilise aussi parfois la notation de Newton :

$$\dot{f}, \ddot{f}, \dddot{f}, \ldots$$

Définition 4. *(Classe de régularité).* On dit qu'une fonction $f : D \to \mathbb{R}$ est de classe C^n (où n est un entier positif ou nul) si f est dérivable sur D jusqu'à l'ordre n et la dérivée d'ordre n, $f^{(n)}$, est continue sur D. On dit que f est de classe C^∞ si f est dérivable sur D une infinité de fois.

En particulier :
— Une fonction f de classe C^0 est une fonction continue.
— Une fonction f de classe C^1 est une fonction dérivable et sa dérivée f' est continue.
— Une fonction f de classe C^2 est une fonction deux fois dérivable et sa dérivée seconde f'' est continue.
— Une fonction f de classe C^∞ est une fonction infiniment dérivable. On dit aussi que c'est une fonction lisse ou régulière. Par exemple, la fonction $f : x \to e^{2x}$ est de classe C^∞ puisque la dérivée à l'ordre n existe pour tout entier $n \geq 0$ et est simplement $f^{(n)} : x \to 2^n e^{2x}$.

Il est particulièrement agréable de travailler avec des fonctions de classe C^∞ puisque l'on peut les dériver sans problème autant de fois que l'on veut. Heureusement, en sciences, nous avons souvent à faire à de telles fonctions. En particulier, les fonctions usuelles $x \mapsto x^n$ (n entier positif ou négatif), $x \mapsto e^x$, $x \mapsto \ln x$, $x \mapsto \cos x$, $x \mapsto \sin x$ et leurs sommes, produits et compositions sont toutes de classe C^∞ sur leurs domaines de définition respectifs.

1.1.3 Règles de dérivation

Nous rappelons ici les règles usuelles de dérivation. Comme il est très courant de le faire, nous utiliserons parfois l'abus de notation « $[f(x)]'$ » ou « $(f(x))'$ » pour désigner la dérivée $f'(x)$ en x.

On a 4 règles générales très importantes à retenir. Soit f et g deux fonctions dérivables, partout où cela est bien défini, les fonctions suivantes sont dérivables et on a :

1. Somme de 2 fonctions : $[f(x) + g(x)]' = f'(x) + g'(x)$
2. Produit de 2 fonctions : $[f(x)\ g(x)]' = f'(x)\ g(x) + f(x)\ g'(x)$
3. Fonction à une puissance (r est une constante réelle) : $[f(x)^r]' = rf'(x)f(x)^{r-1}$
4. Composition de 2 fonctions : $[f(g(x))]' = g'(x)f'(g(x))$

En appliquant la règle 3 pour $r = -1$, on peut déduire : $[1/f(x)]' = -f'(x)/f(x)^2$.
En appliquant les règles 2 et 3, on peut déduire : $[f(x)/g(x)]' = [f'(x)\ g(x) - f(x)\ g'(x)]/g(x)^2$.
En appliquant la règle 3 pour $r = 1/2$, on peut déduire la dérivée de la fonction racine carrée : $(\sqrt{x})' = (x^{1/2})' = (1/2)x'x^{1/2-1} = (1/2)x^{-1/2} = 1/(2\sqrt{x})$.

Rappelons aussi les dérivées de quelques fonctions usuelles :
- $(e^x)' = e^x$
- $(\ln x)' = 1/x$
- $(\cos x)' = -\sin x$
- $(\sin x)' = \cos x$

En utilisant la dérivée $(f/g)'$, on peut déduire la dérivée de la fonction tangente : $(\tan x)' = (\sin x/\cos x)' = (\cos x \cos x + \sin x \sin x)/(\cos x)^2 = 1/(\cos x)^2$.

En utilisant la règle 4, on peut déduire les dérivées des fonctions composées suivantes :
- $(e^{g(x)})' = g'(x)e^{g(x)}$
- $(\ln g(x))' = g'(x)/g(x)$
- $(\cos g(x))' = -g'(x) \sin g(x)$
- $(\sin g(x))' = g'(x) \cos g(x)$

1.1.4 Extremums

On a souvent besoin de trouver les extremums d'une fonction, c'est-à-dire les maximums et les minimums.

Définition 5. *(Extremums).* Soit une fonction $f : D \to \mathbb{R}$. On dit que :
- f a un maximum global en x_0 si $f(x_0) \geq f(x)$ pour tout $x \in D$.
- f a un minimum global en x_0 si $f(x_0) \leq f(x)$ pour tout $x \in D$.
- f a un maximum local en x_0 s'il existe $\epsilon > 0$ tel que $f(x_0) \geq f(x)$ pour tout $x \in]x_0 - \epsilon, x_0 + \epsilon[$.
- f a un minimum local en x_0 s'il existe $\epsilon > 0$ tel que $f(x_0) \leq f(x)$ pour tout $x \in]x_0 - \epsilon, x_0 + \epsilon[$.

Un intervalle du type $]x_0 - \epsilon, x_0 + \epsilon[$ pour $\epsilon > 0$ s'appelle un voisinage de x_0.

Définition 6. *(Point stationnaire).* Soit une fonction $f : D \to \mathbb{R}$ dérivable dans un voisinage d'un point $x_0 \in D$. On dit que le point x_0 est un point stationnaire de f si la dérivée est nulle en ce point : $f'(x_0) = 0$.

En un point stationnaire, la tangente à la courbe représentative de f est donc horizontale.

Théorème 3. *(Classification des points stationnaires).* Soit une fonction $f : D \to \mathbb{R}$ de classe C^3 dans un voisinage d'un point stationnaire $x_0 \in D$ (on a donc $f'(x_0) = 0$). On considère trois possibilités :
- Si $f''(x_0) > 0$, alors f a minimum local en x_0.
- Si $f''(x_0) < 0$, alors f a maximum local en x_0.
- Si $f''(x_0) = 0$ et $f'''(x_0) \neq 0$, alors f a un point d'inflexion en x_0.

Exemples :
- La fonction $x \mapsto x^2$ a un minimum en $x_0 = 0$.
- La fonction $x \mapsto e^{-x^2}$ a un maximum en $x_0 = 0$.
- La fonction $x \mapsto x^3$ a un point d'inflexion en $x_0 = 0$.

Remarque 1 : Pour une fonction non dérivable, on peut aussi avoir des extremums en des points non stationnaires. Par exemple, la fonction $x \mapsto |x|$ a un minimum à $x_0 = 0$ mais $x_0 = 0$ n'est pas un point stationnaire (la dérivée n'est pas définie en $x_0 = 0$).

Remarque 2 : Quand on définie une fonction sur un domaine limité, on peut aussi avoir des extremums en des points non stationnaires qui sont sur un bord du domaine. Par exemple, si on définit la fonction $f : x \mapsto x$ sur le domaine $D = [0, 1]$, alors f a un minimum en $x_0 = 0$ (bord inférieur) et un maximum à $x_0 = 1$ (bord supérieur) mais ces points ne sont pas des points stationnaires.

Remarque 3 : On peut aussi avoir des points d'inflexion qui ne sont pas des points stationnaires, c'est-à-dire $f'(x_0) \neq 0$ et $f''(x_0) = 0$ et $f'''(x_0) \neq 0$.

1.2 Fonction de plusieurs variables réelles

Nous allons maintenant généraliser l'étude précédente à des fonctions de plusieurs variables réelles. Ces fonctions sont en effet omniprésentes en sciences.

1.2.1 Définition

Définition 7. *(Fonction de plusieurs variables réelles).* Une fonction f d'un ensemble $D \subset \mathbb{R}^d$ (où $d \geq 1$ est un entier) dans $\mathbb{R}$ associe à d nombres réels $(x_1, x_2, ..., x_d) \in D$ un nombre réel noté $f(x_1, x_2, ..., x_d)$. On note :

$$f : \begin{array}{ccc} D & \to & \mathbb{R} \\ (x_1, x_2, ..., x_d) & \mapsto & f(x_1, x_2, ..., x_d). \end{array}$$

Dans la suite, on s'intéressera au cas de deux variables ($d = 2$) et on prendra le domaine entier $D = \mathbb{R}^2$ pour faire simple. On appellera le plus souvent les variables (x, y) et la fonction sera donc notée $f : (x, y) \mapsto f(x, y)$.

Exemple : La fonction $f : (x, y) \mapsto e^{-(x^2+y^2)}$ est une fonction de deux variables réelles. On peut représenter cette fonction en traçant la surface $z = f(x, y)$ dans l'espace à 3 dimensions x, y, z (voir figure 1.1 à gauche). On peut aussi tracer les lignes de niveau $f(x, y) = c$ dans l'espace à 2 dimensions x, y, où c est la valeur déterminant la ligne de niveau (voir figure 1.1 à droite).

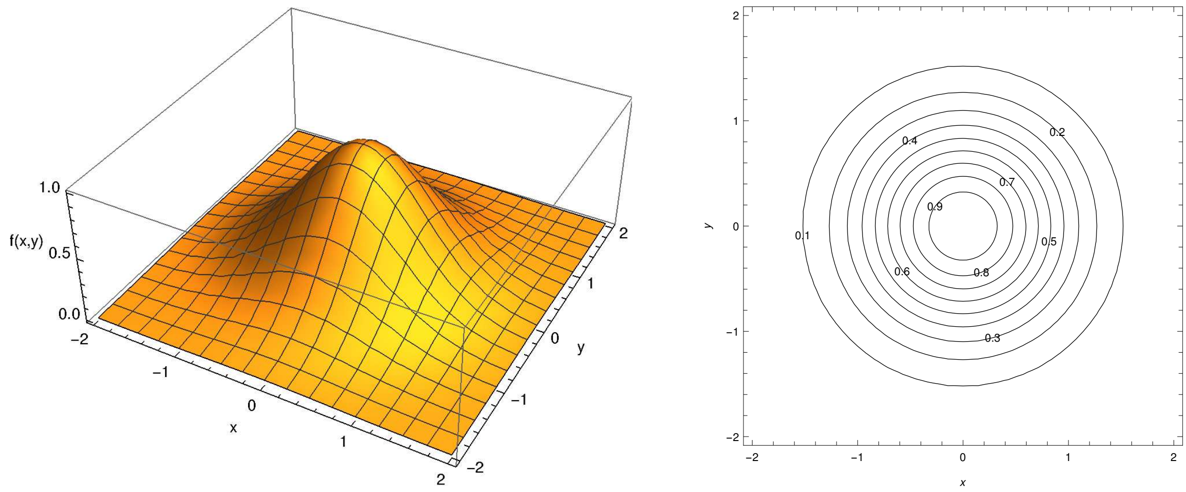

FIGURE 1.1 – *Deux représentations graphiques de la fonction $f : (x, y) \mapsto e^{-(x^2+y^2)}$. A gauche : surface $z = f(x, y)$ dans l'espace à 3 dimensions x, y, z. A droite : lignes de niveau $f(x, y) = c$ dans l'espace à 2 dimensions x, y (les valeurs de c sont indiquées sur le graphe).*

Remarque 1 : Le cas de trois variables ($d = 3$) est aussi très important puisqu'il permet de décrire des fonctions des 3 variables d'espace (x, y, z), c'est-à-dire $f : (x, y, z) \mapsto f(x, y, z)$. Nous ne le traiterons pas explicitement. Néanmoins, les concepts que nous allons introduire pour le cas $d = 2$ se généralisent sans trop de difficulté à $d = 3$.

Remarque 2 : On peut aussi considérer des fonctions de $\mathbb{R}^d$ dans $\mathbb{R}^k$ (où $k \geq 1$ est un entier). Par exemple, pour $d = 2$ et $k = 2$, $f(x, y)$ est alors un couple de deux nombres réels, c'est-à-dire $f(x, y) = (f_1(x, y), f_2(x, y))$ où $f_1(x, y)$ et $f_2(x, y)$ sont deux nombres réels. On peut alors l'interpréter comme un vecteur à deux dimensions, c'est-à-dire $\vec{f}(x, y) = f_1(x, y)\vec{i} + f_2(x, y)\vec{j}$ où $\vec{i}$ et $\vec{j}$ sont deux vecteurs orthonormaux formant un base du plan $\mathbb{R}^2$. On parle alors de fonction vectorielle ou de champ de vecteurs.

1.2.2 Continuité et dérivabilité

La définition de la continuité d'une fonction de deux variables réelles est une simple généralisation de la continuité d'une fonction d'une variable réelle.

Définition 8. *(Continuité).* On dit qu'une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ est continue au point $(x_0, y_0) \in \mathbb{R}^2$ si

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0).$$

Exemple : La fonction $f : (x, y) \mapsto e^{-(x^2+y^2)}$ est continue sur $\mathbb{R}^2$.

Définition 9. *(Dérivées partielles)*. Soit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$. La dérivée partielle de f par rapport à x au point (x_0, y_0) est, si elle existe,

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h},$$

et la dérivée partielle de f par rapport à y au point (x_0, y_0) est, si elle existe,

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}.$$

Si ces deux dérivées partielles existent (et sont donc finies) alors on dit que f est dérivable au point (x_0, y_0).

Les dérivées partielles définissent de nouvelles fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$:

$$\begin{array}{rccl} \dfrac{\partial f}{\partial x} : & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x, y) & \mapsto & \dfrac{\partial f}{\partial x}(x, y) \end{array} \qquad \text{et} \qquad \begin{array}{rccl} \dfrac{\partial f}{\partial y} : & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x, y) & \mapsto & \dfrac{\partial f}{\partial y}(x, y). \end{array}$$

La dérivée partielle $\partial f / \partial x$ est donc la dérivée de f par rapport à la variable x en considérant l'autre variable y comme une constante. De manière symétrique, la dérivée partielle $\partial f / \partial y$ est la dérivée de f par rapport à la variable y en considérant l'autre variable x comme une constante.

D'un point de vue géométrique, la dérivée partielle $\frac{\partial f}{\partial x}(x_0, y_0)$ est le coefficient directeur de la tangente à la surface représentative de f au point (x_0, y_0) dans la direction x. De même, la dérivée partielle $\frac{\partial f}{\partial y}(x_0, y_0)$ est le coefficient directeur de la tangente à la surface représentative de f au point (x_0, y_0) dans la direction y.

Attention, contrairement au cas à une seule variable, une fonction de plusieurs variables dérivable en un point n'est pas forcément continue en ce point.

Exemple 1 : Soit la fonction $f : (x, y) \mapsto x^2 + y^2 + xe^y$. Les dérivées partielles sont : $\frac{\partial f}{\partial x}(x, y) = 2x + e^y$ et $\frac{\partial f}{\partial y}(x, y) = 2y + xe^y$.

Exemple 2 : L'énergie interne U d'un gaz parfait monoatomique (par exemple, un gaz fait d'atomes argon assez dilué) est une fonction du nombre de moles n et de la température T :

$$U : (n, T) \mapsto U(n, T) = \frac{3}{2} nRT,$$

où R est la constante des gaz parfaits. La capacité thermique c est définie comme la dérivée partielle de U par rapport à T :

$$c = \frac{\partial U}{\partial T}(n, T) = \frac{3}{2} nR.$$

Définition 10. *(Gradient).* Soit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ dérivable. Le gradient de f, noté $\vec{\nabla} f$ ou $\overrightarrow{\text{grad}} f$, est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par :

$$\begin{aligned} \vec{\nabla} f : \ & \mathbb{R}^2 \ \to \ \mathbb{R}^2 \\ & (x,y) \mapsto \vec{\nabla} f(x,y) = \left(\frac{\partial f}{\partial x}(x,y) \ , \ \frac{\partial f}{\partial y}(x,y) \right) = \frac{\partial f}{\partial x}(x,y)\vec{i} + \frac{\partial f}{\partial y}(x,y)\vec{j}, \end{aligned}$$

où $\vec{i}$ et $\vec{j}$ sont les vecteurs unitaires dans les directions x et y, respectivement.

Le gradient $\vec{\nabla} f$ est donc une fonction vectorielle ou champ de vecteurs. Le symbole $\vec{\nabla}$ s'appelle "nabla". La direction de $\vec{\nabla} f(x,y)$ donne, en partant du point (x,y), la direction suivant laquelle la fonction f augmente le plus vite.

Exemple : On considère une plaque métallique de température inhomogène. A chaque instant, la température de la plaque $T : (x,y) \mapsto T(x,y)$ est une fonction des deux variables d'espace. A chaque point (x,y), le flux de chaleur en ce point est dans la direction opposée au gradient de la température $\vec{\nabla} T(x,y)$.

On peut continuer à dériver les dérivées partielles par rapport à x ou à y. On définit ainsi :

— la dérivée partielle seconde où l'on dérive deux fois par rapport à x :

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right);$$

— la dérivée partielle seconde où l'on dérive deux fois par rapport à y :

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right);$$

— les dérivées partielles secondes croisées où l'on dérive d'abord par rapport à y puis par rapport à x, et inversement :

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

De la même manière, on peut définir 8 dérivées partielles troisièmes, et ainsi de suite.

Définition 11. *(Classe de régularité).* Une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ est de classe C^n si toutes les dérivées partielles jusqu'à l'ordre n existent et sont continues.

En général, on a donc 4 dérivées partielles secondes. Heureusement, pour des fonctions suffisamment régulières, on a le théorème suivant qui apporte une simplification :

Théorème 4. *(Théorème de Schwarz).* Soit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$. Si la fonction f est de classe C^2 alors les dérivées partielles secondes croisées sont égales :

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Autrement dit, la dérivée partielle seconde croisée ne dépend pas de l'ordre dans lequel on dérive, ce qui est très pratique.

Exemple : Reprenons la fonction $f : (x, y) \mapsto x^2 + y^2 + xe^y$. La fonction f est de classe C^∞ (donc en particulier de classe C^2). Les dérivées partielles secondes sont : $\frac{\partial^2 f}{\partial x^2} = 2$, $\frac{\partial^2 f}{\partial y^2} = 2 + xe^y$, $\frac{\partial^2 f}{\partial x \partial y} = e^y$ et $\frac{\partial^2 f}{\partial y \partial x} = e^y$. On trouve bien l'égalité des dérivées partielles secondes croisées.

Définition 12. *(Différentielle)*. Soit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^1. La différentielle de f au point (x, y) est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par :

$$\begin{array}{rccl} df_{(x,y)} : & \mathbb{R}^2 & \to & \mathbb{R} \\ & (h_1, h_2) & \mapsto & df_{(x,y)}(h_1, h_2) = \dfrac{\partial f}{\partial x}(x, y)h_1 + \dfrac{\partial f}{\partial y}(x, y)h_2. \end{array}$$

Dans cette définition, x et y sont des paramètres fixés, alors que h_1 et h_2 sont les variables. La différentielle $df_{(x,y)}(h_1, h_2)$ représente la variation de f au premier ordre induite par une variation de x d'une quantité h_1 et par une variation de y d'une quantité h_2, en partant du point (x, y). Autrement dit, il s'agit du terme du premier ordre dans le développement limité de f par rapport à x et y. Puisque l'on pense habituellement à h_1 et h_2 comme des petites variations de x et y, on note très souvent $h_1 = dx$ et $h_2 = dy$, et on écrit en sciences la différentielle en notation compacte sans forcément préciser le point (x, y) auquel on la calcule :

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy.$$

Exemple : Reprenons la fonction $f : (x, y) \mapsto x^2 + y^2 + xe^y$. La différentielle de f s'écrit :

$$df = (2x + e^y)dx + (2y + xe^y)dy.$$

Comme le gradient, la différentielle permet de réunir dans un seul objet mathématique toutes les dérivées partielles premières d'une fonction. De plus, les mêmes règles de dérivation s'appliquent pour les différentielles. Par exemple, si f et g sont deux fonctions de deux variables réelles de classe C^1, on a (partout où cela est bien défini) :

- Somme de 2 fonctions : $d(f + g) = df + dg$
- Produit de 2 fonctions : $d(f \times g) = df \times g + f \times dg$
- Fonction à une puissance (r est une constante réelle) : $d(f^r) = r\ f^{r-1}\ df$

Remarque : On peut aussi définir une différentielle pour une fonction d'une variable $f : \mathbb{R} \to \mathbb{R}$ de classe C^1. La différentielle de f au point x est alors une fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par :

$$\begin{array}{rccl} df_x : & \mathbb{R} & \to & \mathbb{R} \\ & h & \mapsto & df_x(h) = f'(x)h. \end{array}$$

ou, dans la notation compacte habituellement utilisée en sciences,

$$df = f'(x)dx.$$

Ici, la différentielle df représente la variation de f au premier ordre induite par une variation de x d'une quantité dx. Mais, pour une fonction d'une seule variable réelle, le concept de différentielle n'apporte pas grand chose de plus que le concept de dérivée.

En thermodynamique, les différentielles de fonctions à plusieurs variables sont énormément utilisées. Voici un exemple.

Exemple : En thermodynamique, pour une substance (gaz, liquide ou solide) dans un récipient fermé de volume V et à la température T, une quantité centrale est l'énergie libre F qui est une fonction des deux variables V et T (s'il n'y a pas de réaction chimique) :

$$F : (V, T) \mapsto F(V, T).$$

La différentielle de F s'écrit :

$$dF = \frac{\partial F}{\partial V} dV + \frac{\partial F}{\partial T} dT.$$

On peut interpréter dF comme la petite variation de F induite par une petite variation de V d'une quantité dV et d'une petite variation de T d'une quantité dT. La quantité $P = \partial F/\partial V$ correspond à la pression et la quantité $c_V = \partial F/\partial T$ correspond à la capacité calorifique à volume constant.

1.2.3 Extremums

La définition d'un maximum ou d'un minimum se généralise de manière évidente à une fonction de plusieurs variables réelles.

Donnons la généralisation de la définition d'un point stationnaire.

Définition 13. *(Point stationnaire).* Soit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ dérivable dans un voisinage d'un point $(x_0, y_0) \in \mathbb{R}^2$. On dit que le point (x_0, y_0) est un point stationnaire de f si les deux dérivées partielles premières sont nulles en ce point :

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

Avec le gradient, on peut écrire ces deux conditions de manière plus compacte : $\vec{\nabla} f(x_0, y_0) = \vec{0}$. Un point stationnaire de f est donc un point où le gradient de f s'annule. Également, on peut écrire que la différentielle de f calculée en ce point (x_0, y_0) est la fonction nulle : $df_{(x_0,y_0)} = 0$.

Définition 14. *(Déterminant Hessien).* Soit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^2 dans un voisinage d'un point (x_0, y_0). On note :

$$r = \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \quad \text{et} \quad t = \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$$
$$\text{et} \quad s = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0).$$

Le déterminant Hessien de f en (x_0, y_0) est

$$D = r\,t - s^2.$$

Attention à ne pas oublier que r, t, s et D dépendent du point (x_0, y_0).

Théorème 5. *(Classification des points stationnaires).* Soit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^2 dans un voisinage d'un point stationnaire (x_0, y_0) (on a donc $\frac{\partial f}{\partial x}(x_0, y_0) = 0$ et $\frac{\partial f}{\partial y}(x_0, y_0) = 0$). On considère trois possibilités :
— Si $D > 0$ et $r > 0$, alors f a minimum local en (x_0, y_0).
— Si $D > 0$ et $r < 0$, alors f a maximum local en (x_0, y_0).
— Si $D < 0$, alors f a un point selle (ou point col) en (x_0, y_0), c'est-à-dire un maximum dans une direction et un minimum dans une autre.

Si $D = 0$, alors on ne peut pas conclure sur la nature du point stationnaire.

Exemple : Avec les critères précédents, on peut vérifier que :
— la fonction $f : (x, y) \mapsto x^2 + y^2$ a un minimum en $(0, 0)$ (figure 1.2 à gauche) ;
— la fonction $f : (x, y) \mapsto -x^2 - y^2$ a un maximum en $(0, 0)$ (figure 1.2 au milieu) ;
— la fonction $f : (x, y) \mapsto x^2 - y^2$ a un point selle en $(0, 0)$ (figure 1.2 à droite).

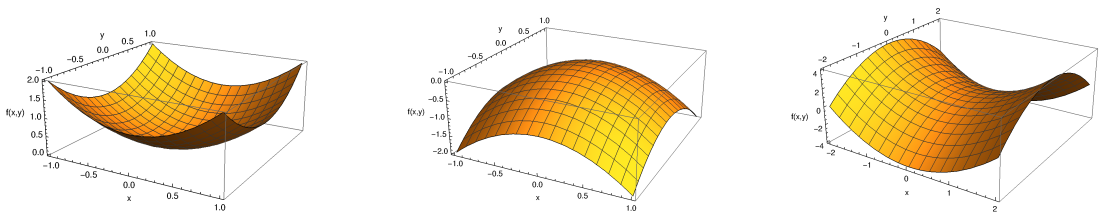

FIGURE 1.2 – *A gauche,* $f : (x, y) \mapsto x^2 + y^2$ *a un minimum en* $(0, 0)$. *Au milieu,* $f : (x, y) \mapsto -x^2 - y^2$ *a un maximum en* $(0, 0)$. *A droite,* $f : (x, y) \mapsto x^2 - y^2$ *a un point selle en* $(0, 0)$.

1.3 Optimisation

Les problèmes d'optimisation sont extrêmement courants en sciences en général et en chimie en particulier. Il s'agit de trouver les valeurs des variables qui maximisent ou minimisent une fonction donnée. Il y a deux grands types de problèmes d'optimisation : sans et avec contraintes.

1.3.1 Optimisation sans contraintes

Commençons par le cas d'une fonction d'une variable réelle $f : x \mapsto f(x)$ supposée dérivable. Pour chercher les maximums ou les minimums de f, on commence par chercher les points stationnaires de f (c'est-à-dire les points x_0 tels que $f'(x_0) = 0$), puis on identifie la nature de ces points stationnaires en utilisant le théorème 3. Les maximums correspondent aux points pour lesquels $f''(x_0) < 0$ et les minimums correspondent aux points pour lesquels $f''(x_0) > 0$. S'il y a plusieurs maximums ou minimums, on peut calculer la valeur de la fonction $f(x_0)$ en ces points pour identifier le maximum global ou le minimum global.

Exercice 1. *(Vitesse de réaction).* En fonction de la valeur du pH, la vitesse $v(\mathrm{pH})$ (en mmol.L^{-1}/min) de la réaction enzymatique de métabolisme du lactose en glucose suit l'expression suivante (à une certaine température fixée) :

$$v(\mathrm{pH}) = \frac{1}{1 + 0.1(\mathrm{pH} - 7.5) + 0.1(\mathrm{pH} - 7.5)^2}.$$

Déterminer la valeur du pH pour laquelle la vitesse est maximale.

Exercice 2. *(Potentiel de Morse).* L'énergie potentielle d'une molécule diatomique en fonction de la distance d entre les deux atomes est bien modélisée par le potentiel de Morse :

$$V(d) = V_0 + E\left(1 - e^{-a(d-d_0)}\right)^2,$$

où V_0, E, a, d_0 sont des constantes réelles (E, a et d_0 sont strictement positifs). Déterminer la valeur de la distance d pour laquelle l'énergie potentielle est minimale.

Prenons maintenant le cas d'une fonction de deux variables réelles $f : (x, y) \mapsto f(x, y)$ supposée dérivable. Pour chercher les maximums ou les minimums, on commence par chercher les points stationnaires de f (c'est-à-dire les points (x_0, y_0) tels que $\frac{\partial f}{\partial x}(x_0, y_0) = 0$ et $\frac{\partial f}{\partial y}(x_0, y_0) = 0$), puis on identifie la nature de ces points stationnaires en utilisant le théorème 5. Les maximums correspondent aux points pour lesquels $D > 0$ et $r < 0$, et les minimums correspondent aux points pour lesquels $D > 0$ et $r > 0$. S'il y a plusieurs maximums ou minimums, on peut calculer la valeur de la fonction $f(x_0, y_0)$ en ces points pour identifier le maximum global ou le minimum global.

Exercice 3. *(Molécule d'eau).* Proche de la géométrie d'équilibre, l'énergie potentielle d'un molécule d'eau en fonction de la distance d de la liaison O-H et de l'angle θ entre les deux liaisons O-H peut être modélisé par :

$$V(d, \theta) = V_0 + k_1(d - d_0)^2 + k_2(\theta - \theta_0)^2,$$

où V_0, k_1, k_2, d_0, θ_0 sont des constantes réelles (k_1, k_2, d_0 et θ_0 sont strictement positifs). Déterminer les valeurs de la distance d et de l'angle θ pour lesquelles l'énergie potentielle est minimale.

1.3.2 Optimisation avec contraintes : méthode des multiplicateurs de Lagrange

Dans un problème d'optimisation avec contraintes, on cherche encore les valeurs des variables qui maximisent ou minimisent une fonction donnée mais on ajoute en plus une ou plusieurs contraintes que les variables doivent satisfaire. On ne peut plus chercher simplement les points stationnaires de la fonction sans prendre en compte les contraintes. Il existe cependant une méthode générale et très utilisée pour résoudre ce problème : la méthode des multiplicateurs de Lagrange.

Nous donnons ici la méthode des multiplicateurs de Lagrange pour le cas du problème de maximisation ou minimisation d'une fonction de deux variables réelles avec une contrainte.

Théorème 6. *(Méthode des multiplicateurs de Lagrange).* On cherche les points (x_0, y_0) qui maximisent ou minimisent une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^1 avec une contrainte de la forme $g(x_0, y_0) = 0$ où $g : \mathbb{R}^2 \to \mathbb{R}$ est aussi une fonction de classe C^1. On introduit la fonction de Lagrange (ou lagrangien) $L : \mathbb{R}^3 \to \mathbb{R}$ définie par

$$L(x, y, \lambda) = f(x, y) + \lambda\, g(x, y),$$

où la nouvelle variable introduite λ s'appelle « multiplicateur de Lagrange ». Les solutions (x_0, y_0) du problème initial de maximisation ou minimisation de $f(x, y)$ avec la contrainte $g(x_0, y_0) = 0$ sont parmi les points stationnaires de la fonction de Lagrange, c'est-à-dire déterminées par le système d'équations

$$\begin{cases} \dfrac{\partial L}{\partial x}(x_0, y_0, \lambda_0) &= \dfrac{\partial f}{\partial x}(x_0, y_0) + \lambda_0 \dfrac{\partial g}{\partial x}(x_0, y_0) = 0 \\ \dfrac{\partial L}{\partial y}(x_0, y_0, \lambda_0) &= \dfrac{\partial f}{\partial y}(x_0, y_0) + \lambda_0 \dfrac{\partial g}{\partial y}(x_0, y_0) = 0 \\ \dfrac{\partial L}{\partial \lambda}(x_0, y_0, \lambda_0) &= g(x_0, y_0) = 0. \end{cases}$$

La méthode des multiplicateurs de Lagrange transforme donc un problème d'optimisation avec contraintes en un problème d'optimisation sans contraintes sur la fonction de Lagrange L et on peut alors résoudre le problème en cherchant les points stationnaires L. Ceci se fait au prix de l'introduction d'une nouvelle variable fictive : le multiplicateur de Lagrange λ. On résolvant le système d'équations, on peut trouver, en plus de x_0 et y_0, la valeur du multiplicateur de Lagrange au point stationnaire λ_0, mais seuls x_0 et y_0 nous intéressent.

Une fois que l'on a trouvé les points stationnaires de L, il reste en principe à identifier celui ou ceux qui correspondent à un maximum ou à un minimum du problème initial. Pour cela, on pourrait calculer les dérivées partielles secondes de L et utiliser une généralisation du théorème 5 aux fonctions de trois variables. Nous ne le ferons pas. On se contentera de calculer les valeurs $f(x_0, y_0)$ de la fonction aux points stationnaires et de choisir le maximum ou le minimum voulu.

<u>Remarque 1 :</u> Il y a en fait une hypothèse supplémentaire dans la méthode des multiplicateurs de Lagrange : $\frac{\partial g}{\partial x}(x_0, y_0)$ et $\frac{\partial g}{\partial y}(x_0, y_0)$ ne doivent pas être simultanément nuls au point stationnaire qui nous intéresse.

<u>Remarque 2 :</u> Nous avons traité uniquement le cas d'une seule contrainte. Si on veut imposer deux contraintes, il faut introduire deux multiplicateurs de Lagrange, et ainsi de suite.

<u>Exemple :</u> On cherche les points (x_0, y_0) qui maximisent ou minimisent la fonction f :

$(x, y) \mapsto x + y$ avec la contrainte $x_0^2 + y_0^2 = 1$. La contrainte peut être écrite comme $g(x_0, y_0) = 0$ avec $g(x, y) = x^2 + y^2 - 1$. On introduit la fonction de Lagrange

$$L(x, y, \lambda) = f(x, y) + \lambda\, g(x, y),$$

et on cherche les points stationnaires (x_0, y_0, λ_0) de L

$$\begin{cases} \dfrac{\partial L}{\partial x}(x_0, y_0, \lambda_0) & = 1 + \lambda_0\,(2x_0) = 0 \\ \dfrac{\partial L}{\partial y}(x_0, y_0, \lambda_0) & = 1 + \lambda_0\,(2y_0) = 0 \\ \dfrac{\partial L}{\partial \lambda}(x_0, y_0, \lambda_0) & = x_0^2 + y_0^2 - 1 = 0. \end{cases}$$

Les deux premières équations impliquent $x_0 = y_0$ et, avec la troisième équation, on trouve deux solutions :
— la solution $(x_0, y_0) = (1/\sqrt{2}, 1/\sqrt{2})$ donnant $f(1/\sqrt{2}, 1/\sqrt{2}) = 2/\sqrt{2}$;
— la solution $(x_0, y_0) = (-1/\sqrt{2}, -1/\sqrt{2})$ donnant $f(-1/\sqrt{2}, -1/\sqrt{2}) = -2/\sqrt{2}$.
La première solution correspond à un maximum et la deuxième à un minimum. La solution à ce problème est représentée sur la figure 1.3.

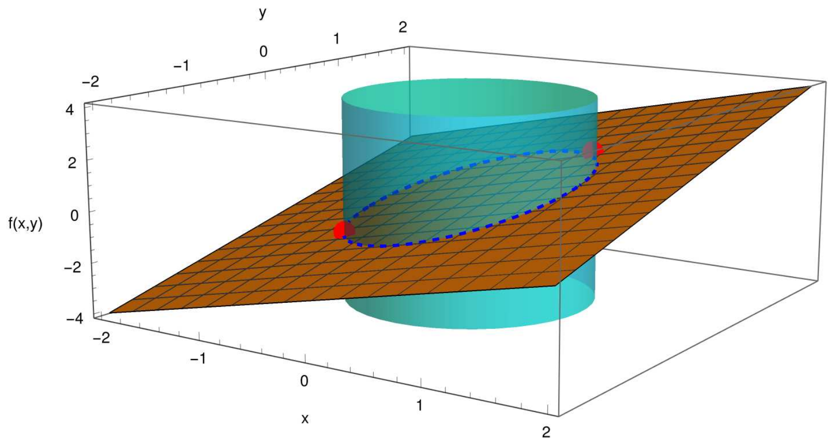

FIGURE 1.3 – *Illustration d'un problème d'optimisation avec contrainte. Le plan orange représente la fonction $f : (x, y) \mapsto x+y$. Le cylindre cyan représente la contrainte $x^2+y^2 = 1$. L'ellipse en pointillés bleu est l'intersection du plan avec le cylindre. Le maximum et le minimum de $f(x, y)$ avec la contrainte $x^2 + y^2 = 1$ correspondent aux points rouges.*

Exercice 4. *(Modèle de Hückel de l'éthylène).* Dans le modèle de Hückel de l'éthylène (C_2H_4), une orbitale π est exprimée comme combinaison linéaire des deux orbitales $p_{z,1}$ et $p_{z,2}$ des atomes de carbone

$$\pi = c_1\, p_{z,1} + c_2\, p_{z,2},$$

où les coefficients (réels) c_1 et c_2 doivent respecter la contrainte de normalisation

$$c_1^2 + c_2^2 = 1.$$

L'énergie de cette orbitale π est alors donnée par

$$E(c_1, c_2) = \alpha(c_1^2 + c_2^2) + 2\beta c_1 c_2,$$

où $\alpha < 0$ et $\beta < 0$ sont des constantes (énergie d'une orbitale p_z et énergie d'interaction entre deux orbitales p_z, respectivement). L'énergie de l'orbitale π liante correspond à la valeur minimale de l'énergie $E(c_1, c_2)$ quand on fait varier les coefficients c_1 et c_2 avec la contrainte $c_1^2 + c_2^2 = 1$. Déterminer cette valeur à l'aide de la méthode des multiplicateurs de Lagrange.

Corrigé des exercices

Exercice 1

Il s'agit d'un problème d'optimisation sans contrainte pour la fonction d'une variable réelle $v : \mathrm{pH} \mapsto v(\mathrm{pH}) = 1/f(\mathrm{pH})$ avec $f(\mathrm{pH}) = 1+0.1(\mathrm{pH}-7.5)+0.1(\mathrm{pH}-7.5)^2$. Commençons par chercher un point stationnaire de v. La dérivée de v est :

$$v'(\mathrm{pH}) = -\frac{f'(\mathrm{pH})}{f(\mathrm{pH})^2} = -\frac{0.1 + 0.2(\mathrm{pH} - 7.5)}{f(\mathrm{pH})^2}.$$

On détermine la valeur pH_0 pour laquelle la dérivée s'annule :

$$\begin{aligned} v'(\mathrm{pH}_0) = 0 &\Leftrightarrow 0.1 + 0.2(\mathrm{pH}_0 - 7.5) = 0 \\ &\Leftrightarrow \mathrm{pH}_0 = 7.0. \end{aligned}$$

Vérifions à présent qu'il s'agit d'un maximum. Pour cela, on calcule la dérivée seconde :

$$\begin{aligned} v''(\mathrm{pH}) &= -\frac{f''(\mathrm{pH})f(\mathrm{pH})^2 - f'(\mathrm{pH})2f'(\mathrm{pH})f(\mathrm{pH})}{f(\mathrm{pH})^4} \\ &= -\frac{f''(\mathrm{pH})f(\mathrm{pH}) - 2f'(\mathrm{pH})^2}{f(\mathrm{pH})^3} \\ &= -\frac{0.2f(\mathrm{pH}) - 2f'(\mathrm{pH})^2}{f(\mathrm{pH})^3}. \end{aligned}$$

Au point stationnaire, $\mathrm{pH} = \mathrm{pH}_0 = 7.0$, on a :

$$\begin{aligned} v''(\mathrm{pH}_0) &= -\frac{0.2f(7.0) - 2f'(7.0)^2}{f(7.0)^3} \\ &= -\frac{0.2}{0.975^2} = -0.210... \end{aligned}$$

La dérivée seconde est strictement négative : il s'agit donc bien d'un maximum. En conclusion, la valeur maximale de la vitesse de réaction est atteinte pour $\mathrm{pH} = 7.0$ et vaut :

$$v_{\max} = v(7.0) = \frac{1}{0.975} = 1.0256...\ \mathrm{mmol.L^{-1}/min}.$$

Le graphe représentatif de la fonction $v : \mathrm{pH} \mapsto v(\mathrm{pH})$ est donné dans la figure 1.4.

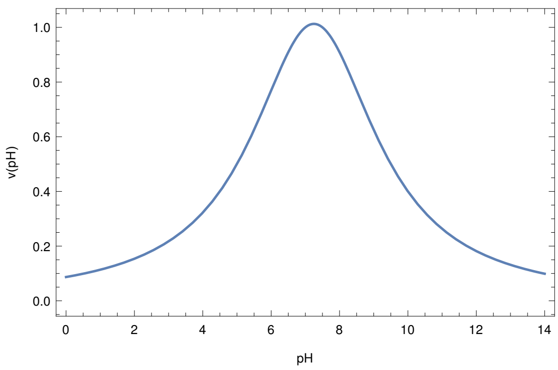

FIGURE 1.4 – Graphe représentatif de la fonction $v : \mathrm{pH} \mapsto v(\mathrm{pH})$.

Exercice 2
Il s'agit d'un problème d'optimisation sans contrainte pour la fonction d'une variable réelle $V : d \mapsto V(d) = V_0 + E\left(1 - e^{-a(d-d_0)}\right)^2$. Commençons par chercher un point stationnaire de V. La dérivée de V est :

$$V'(d) = 2E\,a\,e^{-a(d-d_0)}\left(1 - e^{-a(d-d_0)}\right).$$

On voit que la dérivée s'annule pour $d = d_0$. Vérifions à présent qu'il s'agit d'un minimum. Pour cela, on calcule la dérivée seconde :

$$V''(d) = -2E\,a^2\,e^{-a(d-d_0)}\left(1 - e^{-a(d-d_0)}\right) + 2E\,a^2\,e^{-2a(d-d_0)}.$$

Au point stationnaire, $d = d_0$, on a :

$$V''(d_0) = 2E\,a^2 > 0.$$

La dérivée seconde est strictement positive : il s'agit donc bien d'un minimum. En conclusion, la valeur minimale de l'énergie potentielle est atteinte pour $d = d_0$ et vaut :

$$V_{\max} = V(d_0) = V_0.$$

Le graphe représentatif de la fonction $V : d \mapsto V(d)$ est donné dans la figure 1.5.

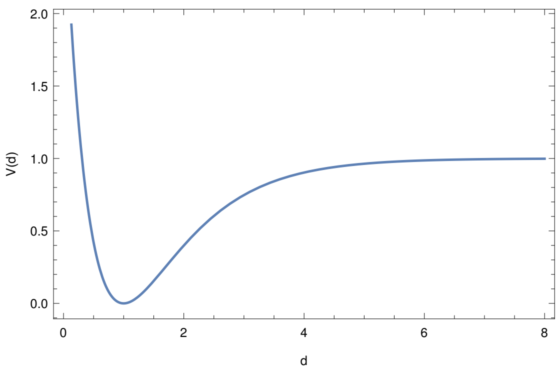

FIGURE 1.5 – Graphe représentatif de la fonction $V : d \mapsto V(d)$ pour $V_0 = 0$, $E = 1$, $a = 1$ et $d_0 = 1$.

Exercice 3
Il s'agit d'un problème d'optimisation sans contrainte pour une fonction de deux variables réelles $V : (d, \theta) \mapsto V(d, \theta) = V_0 + k_1(d - d_0)^2 + k_2(\theta - \theta_0)^2$. Commençons par chercher un point stationnaire de V. Les dérivées partielles premières de V sont :

$$\frac{\partial V}{\partial d}(d, \theta) = 2k_1(d - d_0) \;\text{ et }\; \frac{\partial V}{\partial \theta}(d, \theta) = 2k_2(\theta - \theta_0).$$

On voit que les deux dérivées partielles s'annulent pour $d = d_0$ et $\theta = \theta_0$. Vérifions à présent qu'il s'agit d'un minimum. Pour cela, on calcule les dérivées secondes partielles :

$$\frac{\partial^2 V}{\partial d^2}(d, \theta) = 2k_1 \;\text{ et }\; \frac{\partial^2 V}{\partial \theta^2}(d, \theta) = 2k_2 \;\text{ et }\; \frac{\partial^2 V}{\partial d \partial \theta}(d, \theta) = 0.$$

Pour ce problème, les dérivées secondes ne dépendent donc pas des variables d et θ. Au point stationnaire, on a donc $r = 2k_1$, $t = 2k_2$, $s = 0$ et le déterminant Hessien est

$D = r\,t - s^2 = 4k_1k_2$. On a $D > 0$ et $r > 0$: il s'agit donc bien d'un minimum. En conclusion, la valeur minimale de l'énergie potentielle est atteinte pour $d = d_0$ et $\theta = \theta_0$, et vaut :

$$V_{\max} = V(d_0, \theta_0) = V_0.$$

Exercice 4
Il s'agit d'un problème d'optimisation avec contrainte. On recherche le minimum de la fonction $E : (c_1, c_2) \mapsto E(c_1, c_2) = \alpha(c_1^2 + c_2^2) + 2\beta c_1 c_2$ avec la contrainte $g(c_1, c_2) = c_1^2 + c_2^2 - 1 = 0$. On utilise la méthode des multiplicateurs de Lagrange. On introduit donc la fonction de Lagrange

$$L(c_1, c_2, \lambda) = E(c_1, c_2) + \lambda \left(c_1^2 + c_2^2 - 1\right),$$

et on cherche un point stationnaire de L. Ce point stationnaire doit annuler les dérivées partielles premières de L

$$\begin{cases} \dfrac{\partial L}{\partial c_1} = 2\alpha c_1 + 2\beta c_2 + 2\lambda c_1 = 0 \\ \dfrac{\partial L}{\partial c_2} = 2\alpha c_2 + 2\beta c_1 + 2\lambda c_2 = 0 \\ \dfrac{\partial L}{\partial \lambda} = c_1^2 + c_2^2 - 1 = 0. \end{cases}$$

En supposant $c_1 \neq 0$, on extrait λ de la première équation : $\lambda = -(\alpha c_1 + \beta c_2)/c_1$. De même, en supposant $c_2 \neq 0$, on extrait λ de la deuxième équation : $\lambda = -(\alpha c_2 + \beta c_1)/c_2$. On doit donc avoir

$$-\frac{\alpha c_1 + \beta c_2}{c_1} = -\frac{\alpha c_2 + \beta c_1}{c_2},$$

ce qui donne

$$\alpha c_1 c_2 + \beta c_2^2 = \alpha c_2 c_1 + \beta c_1^2,$$

et qui se simplifie en

$$c_1^2 = c_2^2.$$

On a donc deux possibilités : $c_1 = c_2$ ou $c_1 = -c_2$. En utilisant maintenant la troisième équation du système, on trouve les quatre solutions :

$$(c_1, c_2) = (1/\sqrt{2}, 1/\sqrt{2}) \text{ ou } (c_1, c_2) = (-1/\sqrt{2}, -1/\sqrt{2}),$$
$$\text{ou } (c_1, c_2) = (1/\sqrt{2}, -1/\sqrt{2}) \text{ ou } (c_1, c_2) = (-1/\sqrt{2}, 1/\sqrt{2}).$$

Pour les deux premières solutions, l'énergie vaut

$$E(1/\sqrt{2}, 1/\sqrt{2}) = E(-1/\sqrt{2}, -1/\sqrt{2}) = \alpha + \beta.$$

Pour les deux dernières solutions, l'énergie vaut

$$E(1/\sqrt{2}, -1/\sqrt{2}) = E(-1/\sqrt{2}, 1/\sqrt{2}) = \alpha - \beta.$$

Comme β est négatif, les deux premières solutions donnent l'énergie minimale. Les deux premières solutions correspondent à deux façons équivalentes d'exprimer l'orbitale π liante (on combine les orbitales $p_{z,1}$ et $p_{z,2}$ avec des coefficients c_1 et c_2 de même signe). Les deux dernières solutions donnent l'énergie maximale. Les deux dernières solutions correspondent à deux façons équivalentes d'exprimer l'orbitale π antiliante (on combine les orbitales $p_{z,1}$ et $p_{z,2}$ avec des coefficients c_1 et c_2 de signes opposés).