# Comprendre les dérivées partielles et leurs notations

Kévin Santugini [*]

Ce mini-poly est destiné aux personnes déjà familières avec la notion de dérivation d'une fonction d'une seule variable. Le but de ce mini-poly est d'introduire la notion de différentiation des fonctions à plusieurs variables. Nous allons présenter la théorie dans un ordre unusuel. Nous allons commencer par la notion de dérivée partielle car dans les applications (en physique, mécanique ou autre) ce sont les dérivées partielles qui apparaissent le plus fréquemment. De plus, les notations usuelles pour les dérivées partielles sont très souvent déconcertantes quand elles sont vues pour la première fois. Aussi insisterons nous beaucoup sur la signification des différentes notations utilisées pour les dérivées partielles. Les notions plus élaborées, entre autres la différentielle, seront abordées dans un second temps.

## 1 Les dérivées partielles

### 1.1 Vision calculatoire

Nous commençons par montrer comment définir et calculer une dérivée partielle à partir la notion de dérivée d'une fonction d'une seule variable. Cela permettra de définir la notion de dérivée partielle, d'en expliquer les notations et surtout d'expliquer comment calculer rapidement une dérivée partielle [1]. Commençons par un exemple Soit $f$ la fonction

$$\begin{aligned} f\colon \Omega &\to E \\ (x,y) &\mapsto \sin(xy^2) \end{aligned}$$

Pour calculer la dérivée partielle de $f$ suivant la première variable $x$, on fixe $y$, puis on considère l'application $x \mapsto \sin(xy^2)$ puis on calcule sa dérivée que l'on note

$$\frac{\partial f}{\partial x}(x,y) = y^2\cos(xy^2).$$

De même, pour calculer la dérivée partielle de $f$ suivant la la deuxième variable $y$, on fixe $x$ puis on considère l'application $y \mapsto \sin(xy^2)$ puis on

* `Kevin.Santugini-Repiquetenseirb-matmeca.fr`

1. À la condition bien entendu de savoir calculer rapidement la dérivée d'une fonction d'une seule variable.

calcule sa dérivée :
$$\frac{\partial f}{\partial y}(x,y) = 2xy\cos(xy^2).$$

Étendons maintenant ce procédé. Soit $d$ un entier, $d \geqslant 1$. Considérons une fonction $f$ d'un ouvert $\Omega$ de $\mathbb{R}^d$ à valeur dans un espace vectoriel $E$ :

$$\begin{aligned} f\colon \Omega &\to E \\ (x_1,\dots,x_d) &\mapsto f(x_1,\dots,x_d). \end{aligned}$$

Soit $(x_1,\dots,x_d)$ dans $\Omega$. Fixons $i$ dans $[\![1,d]\!]$. Soit $(x_j)_{1\leqslant j\leqslant d}$ dans $\Omega$. Nous allons considérer l'application :

$$\begin{aligned} \Theta\colon \{s\in\mathbb{R} \mid (x_1,\dots,x_{i-1},s,x_{i+1},\dots,x_d)\in\Omega\} &\to E \\ s &\mapsto f(x_1,\dots,x_{i-1},s,x_{i+1},\dots,x_d) \end{aligned}$$

Si cette fonction d'une seule variable est dérivable en $s = x_i$, alors on dit que $f$ admet une dérivée partielle en $\boldsymbol{x} = (x_1,\dots,x_d)$ suivant sa $i^{\mathrm{e}}$ variable. On note $\frac{\partial f}{\partial x_i}$ cette dérivée partielle que l'on définit par

$$\frac{\partial f}{\partial x_i}(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_d) := \Theta'(x_i). \tag{1.1}$$

Cela revient à considérer que toutes les variables $x_j$ pour $j = 1,\dots,d$ et $j \neq i$ sont constantes et à dériver suivant la variable $x_i$ de la même manière que l'on dérive une fonction d'une seule variable scalaire.

Il est aussi courant d'utiliser la notation $\frac{\partial}{\partial x_i}(f)$ pour désigner cette même dérivée partielle :

$$\frac{\partial}{\partial x_i}(f)(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_d) := \frac{\partial f}{\partial x_i}(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_d), \tag{1.2}$$

par convention. Les parenthèses autour du $f$ sont parfois élidées. La notation (1.2) est surtout utilisée pour des raisons esthétiques, quand le numérateur serait trop long si on utilisait la notation (1.1). Nous utiliserons cette notation à la section §3.

Pour gagner de la place, une autre notation utilisée omet la barre de fraction et le dénominateur. Par convention,

$$\partial_{x_i}(f)(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_d) := \frac{\partial f}{\partial x_i}(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_d), \tag{1.3}$$

Les parenthèses autour du $f$ sont souvent élidées.

## 1.2 Dérivées partielles et notations

Nous allons maintenant nous atteler à une grande source de confusion dans l'apprentissage du calcul différentiel multivariables : les notations. Mais avant de les expliquer, nous devons expliquer pourquoi elles sont si difficiles à maîtriser. Elles sont difficiles car les notations des dérivées partielles ne sont pas conformes aux notations utilisées pour les fonctions. Commençons par un petit aparté sur les notions de notations positionnelles et désignationnelles [2] pour les fonctions.

### 1.2.1 Notations positionnelles et notations désignationnelles

En effet, en mathématique, le choix a été fait pour les fonctions d'utiliser une notation “positionnelle” : c'est la position des arguments qui compte et non la lettre employée pour l'argument dans la définition de la fonction. Ainsi poser $f\colon (x,y) \mapsto x-y$ ou $f\colon (y,x) \mapsto y-x$ est complètement équivalent. Par exemple, dans les deux cas, $f(1,2) = -1$. Il s'agit là d'une convention très largement respectée mais cette convention n'était pas la seule convention possible. Il aurait été tout à fait possible d'imaginer qu'une autre convention s'impose. Cette convention que nous appellerions notation désignationnelle utiliserait des expressions du style $f(x = 2, y = 1)$, ici $x$ et $y$ sont ce que l'on appelle des désignateurs. Dans cette convention, l'ordre des arguments n'aurait plus d'importance, la variable $x$ dans l'expression de $f$ serait remplacée par 2 et la variable $y$ par 1. Et dans cette notation, on aurait $f(x = 2, y = 1) = f(y = 1, x = 2)$. Ce qui implique que dans une notation désignationnelle, $f\colon (x,y) \mapsto x-y$ ou $f\colon (y,x) \mapsto y-x$, ne sont pas équivalentes. Bien sûr, les notations désignationnelles ne sont jamais employées en mathématiques mais elles le sont parfois en informatique, connues sous le nom de « named parameters », « pass-by-name », ou « keyword arguments » comme, entre autres, en `FORTRAN` [3].

Quel rapport avec les dérivées partielles ? Et bien, c'est très simple. **Alors que toutes les notations pour les fonctions sont positionnelles, la notation usuelle pour les dérivées partielle est désignationnelle**. Et c'est exactement à cause de cette incohérence que ceux qui viennent de découvrir les dérivées partielles s'emmêlent les pinceaux.

Prenons un exemple. Considérons une fonction de deux variables scalaires

$$f\colon (x,y) \mapsto f(x,y).$$

Dans $\frac{\partial f}{\partial x}(1,x)$, le $x$ entre parenthèses, dans la liste d'arguments, se réfère à la deuxième variable car les arguments d'une fonction suivent les notations positionnelles. Et inversement, le $x$ au dénominateur, dans $\partial x$, se réfère à la première variable car il suit une notation désignationnelle et dans la définition de $f$, le premier paramètre s'appelle $x$. La présence du $x$ entre parenthèses dans la liste d'arguments ne change pas le sens de $\partial x$ au dénominateur. Donc, dans la même formule, un $x$ se réfère à la première variable et un autre

2. Les mots « désignationnelles » et « positionnelles » ne sont pas standardisés en mathématiques pour les fonctions.

3. À partir du `FORTRAN` 90

se réfère à la seconde. Avec des notations aussi incohérentes, il n'est guère surprenant que les novices en calcul différentiel multivariables se sentent perdus. Malheureusement, ces notations sont maintenant trop ancrées dans l'usage pour les remplacer par des notations plus cohérentes donc il faudra faire avec.

*Remarque* 1.1. Si vous n'êtes pas encore convaincu de l'incohérence complète des notations usuelles pour les dérivées partielles, l'exemple suivant devrait vous convaincre. Considérons les deux fonctions suivantes :

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R} & \qquad g\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto \cos(x)\sin(y) & \qquad (y,x) &\mapsto \cos(y)\sin(x) \end{aligned}$$

Au sens des fonctions, $f = g$. Et donc la dérivée partielle de $f$ suivant la première variable est égale à la dérivée partielle de $g$ suivant la première variable. Mais la désignation de la première variable de $f$ est $x$ alors que la désignation de la première variable de $g$ est $y$. Ainsi la dérivée partielle de $f$ suivant la première variable est notée $\frac{\partial f}{\partial x}$ et celle de $g$ est notée $\frac{\partial g}{\partial y}$. Aussi a-t-on :

$$f = g, \qquad \frac{\partial f}{\partial x} = \frac{\partial g}{\partial y}, \qquad \frac{\partial f}{\partial y} = \frac{\partial g}{\partial x}.$$

Exactement le contraire d'une notation pratique, intuitive et cohérente. Mais, malheureusement, comme dit plus haut, ces notations sont maintenant standarts et nous devrons faire avec.

### 1.2.2 Comprendre la notation des dérivées partielles

Maintenant que nous avons expliqué le problème inhérent aux notations usuels pour les dérivées partielles, nous allons expliquer un moyen pratique de savoir suivant quel variable on dérive quand on rencontre la dérivée partielle d'une fonction. Pour cela, nous allons expliciter la notion de désignateur dans les notations. Un désignateur est juste un caractère qui désigne une position d'argument dans une fonction. Et par facilité, le caractère utilisé comme désignateur pour une position d'argument est le même caractère que celui utilisé pour la variable se situant à cette position dans la définition de la fonction. Explicitons tout cela dans une notation maison :

$$\begin{aligned} f\colon \mathbb{R}^3 &\to \mathbb{R} \\ (x,y,z) &\mapsto f(\underset{\substack{\uparrow\\x}}{x}, \underset{\substack{\uparrow\\y}}{y}, \underset{\substack{\uparrow\\z}}{z}) \end{aligned}$$

Les caractères sous les flèches représentent les désignateurs. Les désignateurs ne changent pas et ce quel que soit les arguments que l'on « appelle ». Aussi, écrira-t-on :

$$f(\underset{\substack{\uparrow\\x}}{x}, \underset{\substack{\uparrow\\y}}{y}, \underset{\substack{\uparrow\\z}}{z}), \quad f(\underset{\substack{\uparrow\\x}}{1}, \underset{\substack{\uparrow\\y}}{1}, \underset{\substack{\uparrow\\z}}{2}), \quad \frac{\partial f}{\partial x}(\underset{\substack{\uparrow\\x}}{2}, \underset{\substack{\uparrow\\y}}{3}, \underset{\substack{\uparrow\\z}}{x}), \quad \frac{\partial f}{\partial y}(\underset{\substack{\uparrow\\x}}{1}, \underset{\substack{\uparrow\\y}}{xy}, \underset{\substack{\uparrow\\z}}{x^2}), \quad \frac{\partial f}{\partial z}(\underset{\substack{\uparrow\\x}}{z}, \underset{\substack{\uparrow\\y}}{x}, \underset{\substack{\uparrow\\z}}{y}). \tag{1.4}$$

Les $\partial x$, $\partial y$ et $\partial z$ aux dénominateurs se réfèrent toujours à la position de l'argument désignée par le caractère sous la flèche qui est celui employé dans la définition de la fonction. Ainsi $\frac{\partial f}{\partial x}$ se réfère (pour cette fonction $f$) toujours à la dérivée partielle suivant la première variable. Le $x$ dans le $\frac{\partial}{\partial x}$ n'est ni un réel ni un élément d'un ensemble quelconque, **c'est une chaîne de caractères associée à la position d'un argument** de la fonction dont on souhaite calculer une dérivée partielle.

Avec ces notations explicites, tout est clair. Mais ces notations étant non conventionnelles, il faut éviter de les écrire ailleurs que sur un brouillon. Le mieux est d'être capable de s'en passer et de se contenter de les rajouter mentalement chaque fois que l'on rencontre une dérivée partielle de fonction.

## 1.3 Dérivation partielle d'expressions

Une notation très couramment utilisée consiste à dériver non une fonction mais une expression mathématique. Il n'y a aucune différence avec la notion de dérivation d'une fonction. Il s'agit simplement d'une notation ou plus exactement d'un abus de notation permettant d'éviter de définir une fonction au préalable (et de lui réserver une lettre) avant de calculer sa dérivée partielle. Par expression mathématique, nous entendons juste la partie après le $\mapsto$ d'une fonction. Par exemple, $xyz$, $\sin(x)\cos(y)$, $yx^2 + xz$ et $\tan(xy)$ peuvent être vu comme des expressions mathématiques. La dérivée partielle d'une expression se calcule exactement comme la dérivée partielle d'une fonction. Regardons un exemple :

$$\frac{\partial x^2yz^2}{\partial x} = 2xyz^2. \tag{1.5}$$

Le résultat est obtenu en considérant toutes les variables présentes numérateur excepté $x$ comme fixes. Il s'agit d'un raccourci et d'un abus de notation pour la dérivée partielle suivant $x$ de la fonction $(x, y, z) \mapsto x^2yz^2$ calculée au point $(x, y, z)$. Cet abus de notation permet de gagner en concision et est très répandu. Il est aussi courant de rencontrer l'expression, non au numérateur mais à droite de la fraction, en ne laissant qu'un « $\partial$ » au numérateur de la fraction :

$$\frac{\partial}{\partial x}(x^2yz^2) = 2xyz^2. \tag{1.6}$$

Les notations (1.5) et (1.6) ont exactement la même signification. On choisit en général l'une ou l'autre de ces notations en fonction de raisons esthétiques. Typiquement, on emploiera la notation (1.6) si l'expression est très longue. Notez que c'est la présence d'une expression entre parenthèses qui distingue la dérivée partielle d'expression (1.6) et la dérivée partielle de fonction (1.2). Dans cette dernière, seul des lettres qui ont déjà été définies comme fonctions apparaissent. Alors, que dans la première apparaissent une ou des lettres qui n'ont pas été préalablement définies comme des fonctions.

On peut aussi introduire des fonctions dans l'expression. Soit

$$\begin{aligned} f\colon \mathbb{R}^3 &\to \mathbb{R} \\ (x, y, z) &\mapsto f(x, y, z) \end{aligned}$$

On peut alors écrire les dérivées partielles d'expressions suivantes :

$$\frac{\partial f(u^2, uv, \cos(uv))}{\partial u}, \quad \frac{\partial f(u^2, uv, \cos(uv))}{\partial v}.$$

Ici, contrairement aux dérivées partielle de fonctions, la liste d'argument est au dessus de la barre de fraction et non à côté. Quand on dérive une expression, le $\partial u$ au dénominateur se réfère aux $u$ présent au numérateur. Si on pose :

$$\begin{aligned} \psi\colon \mathbb{R}^2 &\to \mathbb{R}, \\ (u, v) &\mapsto (u^2, uv, \cos(uv)). \end{aligned}$$

alors, par définition, on a les égalités suivantes entre dérivées partielles de fonctions et dérivées partielles d'expressions :

$$\frac{\partial(f \circ \psi)}{\partial u} = \frac{\partial f(u^2, uv, \cos(uv))}{\partial u}, \qquad \frac{\partial(f \circ \psi)}{\partial v} = \frac{\partial f(u^2, uv, \cos(uv))}{\partial v.}$$

ou si on préfère l'autre notation :

$$\frac{\partial(f \circ \psi)}{\partial u} = \frac{\partial}{\partial u}(f(u^2, uv, \cos(uv))), \quad \frac{\partial(f \circ \psi)}{\partial v} = \frac{\partial}{\partial v}(f(u^2, uv, \cos(uv))).$$

Si jamais une expression ne dépend que d'une unique variable, alors, on remplace le « $\partial$ » par un « d » dans la notation et on parle de dérivée totale. Cela ne change rien au calcul. Par exemple, on notera préférentiellement $\frac{\mathrm{d}f(t,t,t)}{\mathrm{d}t}$, respectivement $\frac{\mathrm{d}}{\mathrm{d}t}(f(t,t,t))$, au lieu de $\frac{\partial f(t,t,t)}{\partial t}$, respectivement $\frac{\partial}{\partial t}(f(t,t,t))$.

Pour pouvoir calculer la dérivée partielle d'une expression constituée d'une fonction dont les arguments sont des expressions non triviale, comme par exemple $\frac{\partial f(u^2,uv,\cos(uv))}{\partial u}$, il faut faire appel à la règle de dérivation en chaîne qui exprime les dérivées partielles de la composition de deux fonctions en fonction des dérivées partielles de chacune des deux fonctions. Règle que nous donnons à la section 2.

## 2 Dérivées partielles et changement de variable

Pour calculer les dérivées partielles de compositions de fonctions ou après un changement de variable, on peut utiliser la formule de dérivation en chaîne. Avant d'énoncer cette formule, il est utile de disposer de la définition suivante

**Définition 2.1.** *Soit $p$ dans $\mathbb{N}^*$. Soit $\Omega$ un ouvert de $\mathbb{R}^p$. Soit*

$$\begin{aligned} \boldsymbol{f}\colon \Omega &\to \mathbb{R}^m, \\ (y_1,\dots,y_p) &\mapsto f(y_1,\dots,y_p). \end{aligned}$$

*une fonction. La fonction $f$ est dite de classe $\mathcal{C}^1$ sur $\Omega$ si elle est continue sur $\Omega$, admet des dérivées partielles $\frac{\partial \boldsymbol{f}}{\partial y_k}$ suivant chacune de ses variables $y_k$ en tout point de $\Omega$, et si ces dérivées partielles $\frac{\partial \boldsymbol{f}}{\partial y_k}$ sont continues sur $\Omega$.*

Aussi, nous admettrons le résultat suivant.

**Proposition 2.2: Règle de dérivation en chaîne**

Soit $m, n, p$ trois entiers naturels. Soit $\Omega_f$ un ouvert de $\mathbb{R}^p$ et soit $\Omega_\psi$ un ouvert de $\mathbb{R}^n$. Soient deux fonctions continues :

$$\begin{aligned} \boldsymbol{f}\colon \Omega_f &\to \mathbb{R}^m \\ (y_1,\dots,y_p) &\mapsto f(y_1,\dots,y_p) \end{aligned} \qquad \begin{aligned} \boldsymbol{\psi}\colon \Omega_\psi &\to \Omega_f \\ (x_1,\dots,x_d) &\mapsto \psi(x_1,\dots,x_d) \end{aligned}$$

Si $f$ est de classe $\mathcal{C}^1$ et si $\psi$ admet une dérivée partielle suivant $x_i$ au point $\boldsymbol{x} = (x_1,\dots,x_n)$ alors $f\circ\psi$ admet une dérivée partielle en $\boldsymbol{x} = (x_1,\dots,x_n)$ suivant $x_i$ et :

$$\frac{\partial(\boldsymbol{f}\circ\boldsymbol{\psi})}{\partial x_i} = \sum_{k=1}^{p} \frac{\partial \boldsymbol{f}}{\partial y_k}(\boldsymbol{\psi}(\boldsymbol{x}))\frac{\partial \psi_k}{\partial x_i}(\boldsymbol{x}). \tag{2.1}$$

*Démonstration.* On se contente de l'idée de la preuve[4]. On pose pour $h > 0$ $\boldsymbol{x}_h = (x_1,\dots,x_{i-1}, x_i+h, x_{i+1}\dots,x_n)$. On écrit :

$$\begin{aligned} &\boldsymbol{f}(\psi_1(\boldsymbol{x}_h)) - \boldsymbol{f}(\psi_1(\boldsymbol{x})) \\ = &\sum_{k=1}^{p} \boldsymbol{f}(\psi_1(\boldsymbol{x}_h),\dots,\psi_k(\boldsymbol{x}_h),\psi_{k+1}(\boldsymbol{x}),\dots,\psi_p(\boldsymbol{x})) \\ &\qquad - \boldsymbol{f}(\psi_1(\boldsymbol{x}_h),\dots,\psi_{k-1}(\boldsymbol{x}_h),\psi_k(\boldsymbol{x}),\dots,\psi_p(\boldsymbol{x})) \end{aligned}$$

On divise par $h$ et on fait tendre $h$ vers 0. En utilisant la compacité locale de $\Omega_\psi$, la continuité de $\boldsymbol{\psi}$, et la continuité des $\frac{\partial \boldsymbol{f}}{\partial y_k}$, on démontre que le $k^{\text{e}}$ terme de la somme divisé par $h$ converge vers le $k^{\text{e}}$ terme de (2.1) quand $h$ tend vers 0. □

Il est très important de savoir appliquer cette formule. Parmi toutes les formules sur la différentiation d'une composition de fonctions, c'est à la fois

4. Pour faire la preuve rigoureusement, il est préférable d'utiliser la notion de différentielle que nous n'abordons pas dans ce mini-poly .

$$\frac{\partial(\boldsymbol{f} \circ \boldsymbol{\psi})}{\partial x_i} = \sum_{k=1}^{p} \frac{\partial \boldsymbol{f}}{\partial y_k}(\boldsymbol{\psi}(\boldsymbol{x}))\frac{\partial \psi_k}{\partial x_i}(\boldsymbol{x}).$$

— Désignateur pour $k^{\text{e}}$ variable de $\boldsymbol{f}$ / $k^{\text{e}}$ composante de $\boldsymbol{\psi}$

— Même désignateur, *i.e.*, même position de variable

FIGURE 1 – La formule de dérivation en chaîne expliquée

la plus rapide à appliquer et la plus susceptible d'être appliquée de manière incorrecte. Premièrement, la notion de dérivée partielle étant une notion purement locale, une formule donnant la dérivée partielle de $\boldsymbol{f} \circ \boldsymbol{\psi}$ en $\boldsymbol{x}$ ne pourra faire intervenir que des dérivées partielles de $f$ *calculées en* $\boldsymbol{\psi}(\boldsymbol{x})$ et des dérivées partielles de $\boldsymbol{\psi}$ *calculées en* $\boldsymbol{x}$. Aussi, les points où sont calculées les dérivées partielles sont souvent omis dans l'expression de la règle de dérivation en chaîne, en particulier dans les ouvrages de physique[5]. La façon dont la formule est organisée est expliquée à l'aide de flèches à la figure 1. Nous allons maintenant appliquer cette formule sur plusieurs exemples. Nous commençons par un changement de coordonnées polaire :

*Exemple* 2.1. Prenons $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto f(x, y)$ et $\boldsymbol{\psi}\colon \mathbb{R}^2 \to \mathbb{R}^2, (r, \theta) \mapsto (r\cos\theta, r\sin\theta)$. On note $\boldsymbol{\psi} = (\psi_x, \psi_y)$. Par la règle de dérivation en chaîne, on a

$$\begin{aligned}\frac{\partial(f \circ \psi)}{\partial r} &= \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)\frac{\partial \psi_x}{\partial r}(r, \theta) + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)\frac{\partial \psi_y}{\partial r}(r, \theta)\\ &= \cos\theta\frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta\frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta).\end{aligned}$$

Dans l'exemple suivant, nous utilisons l'abus de notation de dérivée d'expression.

*Exemple* 2.2. Soit

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R}\\ (u, v) &\mapsto f(u, v)\end{aligned}$$

On a alors :

$$\frac{\partial f(u, u+w)}{\partial u} = \frac{\partial f}{\partial u}(u, u+w) + \frac{\partial f}{\partial v}(u, u+w), \quad \frac{\partial f(u, u+w)}{\partial w} = \frac{\partial f}{\partial v}(u, u+w)$$

5. En omettant les points où sont calculées les dérivées partielles, la règle de dérivation en chaîne s'écrit alors $\frac{\partial(f\circ\psi)}{\partial x_i} = \sum_{k=1}^{p} \frac{\partial f}{\partial y_k}\frac{\partial \psi_k}{\partial x_i}$.

En physique, on dirait que si une fonction mathématique $f$ représente une "grandeur physique $X$" selon les variables $(u,v)$ alors $(u,w) \mapsto f(u,u+w)$ représente la même "grandeur physique $X$" selon les coordonnées $(u,w)$ avec comme changement de variables $u=u$ et $w=v-u$. On remarque que $\frac{\partial f(u,u+w)}{\partial u}$ est différent de $\frac{\partial f}{\partial u}(u,u+w)$ tant bien même que $u$ n'a pas été modifiée par le changement de variable. C'est pourquoi les physiciens, particulièrement en thermodynamique, parlerons de dérivée partielle de la "grandeur physique $X$" suivant $u$ à $v$ constant, et de dérivée partielle suivant $u$ à $w$ constant et les noterons respectivement :

$$\left(\frac{\partial X}{\partial u}\right)_{v=a}(u,v) = \frac{\partial f}{\partial u}(u,v), \qquad \left(\frac{\partial X}{\partial u}\right)_{w=a}(u,w) = \frac{\partial f(u,u+w)}{\partial u}.$$

Dans cette notation, très courante en thermodynamique, "la grandeur physique $X$" représente des fonctions différentes suivant le contexte, toutes pouvant s'obtenir les unes des autres par changement de variables. Ce genre de notation est rare en mathématiques : on préfèrera expliciter le changement de variables. Elles sont par contre très courantes en physique, particulièrement en thermodynamique.

Un cas particulier très important est celui où $\boldsymbol{\psi}$ est fonction d'une seule variable. Dans ce cas, $f \circ \psi$ est aussi fonction d'une seule variable. On peut donc calculer sa dérivée totale :

*Exemple* 2.3. Soit $f$

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R} \\ (u,v) &\mapsto f(u,v) \end{aligned}$$

de classe $\mathcal{C}^1$. Soit $\boldsymbol{\psi} : \mathbb{R} \to \mathbb{R}^2$ de classe $\mathcal{C}^1$. Alors, en appliquant la règle de dérivation en chaîne (2.1), on obtient

$$\frac{\mathrm{d}(f \circ \boldsymbol{\psi})}{\mathrm{d}t} = \frac{\partial f}{\partial u}(\boldsymbol{\psi}(t))\frac{\mathrm{d}\psi_1(t)}{\mathrm{d}t} + \frac{\partial f}{\partial v}(\boldsymbol{\psi}(t))\frac{\mathrm{d}\psi_2(t)}{\mathrm{d}t}.$$

C'est à dire

$$(f \circ \psi)' = \left(\frac{\partial f}{\partial u} \circ \psi\right)\psi_1' + \left(\frac{\partial f}{\partial v} \circ \psi\right)\psi_2'.$$

# 3 Dérivées partielles d'ordre 2 ou plus

Comme pour les fonctions d'une seule variable, il est possible de définir des dérivées partielles d'ordre supérieur. Par exemple, on peut définir $\frac{\partial^2 f}{\partial x_i \partial x_j}$ comme

$$\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$$

quand le terme de droite a un sens et existe. Regardons un exemple. Soit

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto \sin(xy^2) \end{aligned}$$

Alors

$$\frac{\partial f}{\partial x}(x,y) = y^2\cos(xy^2), \qquad \frac{\partial f}{\partial y}(x,y) = 2xy\cos(xy^2).$$

Et

$$\begin{aligned} &\frac{\partial^2 f}{\partial x \partial x}(x,y) = -y^4\sin(xy^2), && \frac{\partial^2 f}{\partial y \partial x}(x,y) = 2y\cos(xy^2) - 4xy^3\sin(xy^2), \\ &\frac{\partial^2 f}{\partial x \partial y}(x,y) = 2y\cos(xy^2) - 2xy^3\sin(xy^2), && \frac{\partial^2 f}{\partial y \partial y}(x,y) = 2x\cos(xy^2) - 4x^2y^2\sin(xy^2). \end{aligned}$$

On constate dans ce cas précis que

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

L'ordre de dérivation n'a pas changé la valeur de la dérivée partielle. Il s'agit d'un cas particulier de l'égalité de Schwarz qui est vrai dès que certaines conditions sur la fonction $f$ sont vérifiées. Nous donnerons ces conditions plus loin.

Pour éviter de répéter le même nom de variable plusieurs fois, on utilise les notations suivantes :

$$\frac{\partial^2 f}{\partial x^2}(x,y) := \frac{\partial^2 f}{\partial x \partial x}(x,y), \qquad \frac{\partial^2 f}{\partial y^2}(x,y) := \frac{\partial^2 f}{\partial y \partial y}(x,y).$$

Nous définissons maintenant rigoureusement par récurrence les dérivées d'ordre supérieur :

**Définition 3.1.** *Soit $d$ dans $\mathbb{N}^*$. Soit $\Omega$ un ouvert de $\mathbb{R}^d$. Soit*

$$\begin{aligned} \boldsymbol{f}\colon \Omega &\to \mathbb{R}^m, \\ (x_1,\ldots,x_d) &\mapsto f(x_d,\ldots,x_d). \end{aligned}$$

*une fonction. Soit $m$ un entier naturel supérieur ou égal à 2. Soit $(i_j)_{1\leqslant j\leqslant m}$ une suite finie dans $[\![1,d]\!]^m$. On dit que $f$ admet une dérivée partielle d'ordre $m$ suivant les variables $x_{i_1},\ldots,x_{i_m}$ sur $\Omega$ si*

— *La fonction $f$ admet une admet une dérivée partielle d'ordre $m-1$ suivant les variables $x_{i_2},\ldots,x_{i_m}$ sur $\Omega$ que l'on note*

$$\frac{\partial^{m-1} f}{\partial x_{i_2}\ldots\partial x_{i_m}}.$$

— *La fonction* $\frac{\partial^{m-1} f}{\partial x_{i_2} \ldots \partial x_{i_m}}$ *admet une dérivée partielle suivant* $x_{i_1}$. *On note alors*

$$\frac{\partial^m f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_m}} := \frac{\partial}{\partial x_{i_1}} \left( \frac{\partial^{m-1} f}{\partial x_{i_2} \ldots \partial x_{i_m}} \right).$$

Pour pouvoir établir des résultats sur les fonctions admettant des dérivées partielles, il est en général utile de supposer que les dérivées partielles sont des fonctions continues jusqu'à un certain ordre. Aussi introduisons-nous la définition suivante qui généralise la Définition 2.1 :

**Définition 3.2.** *Soit $d$ dans $\mathbb{N}^*$. Soit $\Omega$ un ouvert de $\mathbb{R}^d$. Soit*

$$\boldsymbol{f} \colon \Omega \to \mathbb{R}^m,$$
$$(x_1, \ldots, x_d) \mapsto f(x_d, \ldots, x_d).$$

*une fonction. Soit $m$ un entier naturel supérieur ou égal à 1. On dit que $f$ est de classe $\mathcal{C}^m$ sur $\Omega$ si $f$ est continue sur $\Omega$ et si toutes ses dérivées partielles d'ordre inférieur ou égal à $m$ existent et sont continues.*

Quand une fonction est de classe $\mathcal{C}^m$, les dérivées partielles d'ordre $m$ ou moins ne dépendent pas de l'ordre de dérivation.

**Théorem 3.3: Égalité de Schwarz**

Soit $A$ un ouvert de $\mathbb{R}^2$. Soit $f \colon A \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$, de classe $\mathcal{C}^2$. Alors

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

Soit $m$ un entier naturel non nul. Soit $d$ et $n$ deux entiers naturel non nul. Soit $\Omega$ un ouvert de $\mathbb{R}^d$. Soit $g \colon A \to \mathbb{R}^n$, $(x_1, x_2, \ldots, x_d) \mapsto f(x, y)$, de classe $\mathcal{C}^m$. Soit $k$ un entier naturel non nul inférieur ou égal à $m$. Alors, si $(i_1, \ldots, i_k)$ est une permutation de $(j_1, \ldots, j_k)$ :

$$\frac{\partial^k g}{\partial x_{i_1} \ldots \partial x_{i_k}} = \frac{\partial^k g}{\partial x_{j_1} \ldots \partial x_{j_k}}.$$

*Démonstration.* Pour la première partie du théorème, on part de l'égalité

$$\frac{1}{h} \left( \frac{f(h,k) - f(h,0)}{k} - \frac{f(0,k) - f(0,0)}{k} \right) = \frac{1}{k} \left( \frac{f(h,k) - f(,k)}{h} - \frac{f(h,0) - f(0,0)}{h} \right).$$

On applique le principe fondamental de l'analyse dans les deux cas puis on étudie soigneusement la convergence quand $h$ et $k$ tendent vers zéro. La deuxième partie du théorème est une conséquence de la première partie et du raisonnement par récurrence. □

Dans la majorité des cas, nous ne rencontrerons que des fonctions de classe $\mathcal{C}^m$ pour lesquelles l'égalité de Schwarz s'applique [6]. Aussi, en général, on ne fait pas attention à l'ordre dans lequel on dérive quand on manipule des dérivées partielles. Ainsi, en général, on ne répètera pas un même nom de variable au dénominateur d'une dérivée partielle. À la place on utilisera un exposant pour indiquer combien de fois on dérive par rapport à cette dérivée. Ainsi, si $f$ est une fonction de 3 variables $x$, $y$, et $z$ de classe $\mathcal{C}^\infty$, on préfèrera la notation $\frac{\partial^6 f}{\partial x^3 \partial y^2 \partial z}$ à la notation $\frac{\partial^6 f}{\partial x \partial y \partial x \partial y \partial z \partial x}$.

# 4 Matrice Jacobienne, gradient et dérivée par rapport à une variable vectorielle

Dans cette section, nous allons définir la Jacobienne d'une fonction de classe $\mathcal{C}^1$ et la notation de dérivée partielle par rapport à une variable vectoriel.

## 4.1 La Jacobienne et le gradient

En général, la définition de la Jacobienne est donné à partir de la notion plus abstraite de différentielle. Puisque ce mini-poly se concentre sur les dérivées partielles et les aspects calculatoires, nous préférons ici définir la Jacobienne à partir des dérivées partielles. Cette définition est plus pratique quand il s'agit de calculer des Jacobiennes. L'inconvénient de cette définition est qu'elle rend la démonstration de certains résultats moins aisé et plus calculatoire.

6. L'égalité de Schwarz reste vraie sous des hypothèses moins fortes. Elle n'est cependant pas vraie dans le cas général. Il existe des contre-exemples.

**Définition 4.1: La Jacobienne**

Soit $m$ et $n$ deux entiers naturel non nul. Soit $\Omega$ un ouvert de $\mathbb{R}^n$. Soit

$$\begin{aligned} \boldsymbol{f} \colon \Omega &\to \mathbb{R}^m, \\ (x_1, \dots, x_n) &\mapsto \boldsymbol{f}(x_1, \dots, x_n) \end{aligned}$$

une fonction de classe $\mathcal{C}^1$ sur $\Omega$. On note $f_i$ la $i^{\text{e}}$ composante de $\boldsymbol{f}$. On appelle Jacobienne de $f$ en $\boldsymbol{x} = (x_1, \dots, x_d)$ la matrice

$$\mathrm{J}f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{x}) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial f_2}{\partial x_n}(\boldsymbol{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\boldsymbol{x}) & \frac{\partial f_n}{\partial x_2}(\boldsymbol{x}) & \dots & \frac{\partial f_n}{\partial x_n}(\boldsymbol{x}) \end{bmatrix}. \tag{4.1}$$

Ou, écrit autrement,

$$\mathrm{J}f(\boldsymbol{x}) = \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}. \tag{4.2}$$

Cette définition correspond à ce que les mathématiciens appellent la Jacobienne. Dans d'autres disciplines, des conventions différentes peuvent être employées et il est possible de rencontrer des définitions de Jacobienne qui la définissent comme la transposée de la matrice 4.1.

Pour représenter la Jacobienne, d'autres notations sont parfois employées. Le membre droit de l'Eq (4.2) peut être utilisé comme notation. On rencontre parfois dans la littérature

$$\left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant m}}$$

au lieu de $\mathrm{J}f(\boldsymbol{x})$ pour dénoter la Jacobienne. Dans d'autres disciplines, il arrive même que cette notation soit abrégée et de trouver juste

$$\frac{\partial f_i}{\partial x_j}$$

pour dénoter toute la Jacobienne. Cette dernière notation est ambiguë car elle peut représenter toute la Jacobienne ou une seule de ses composantes. Seul le contexte permet de lever cette ambigüïté. Dans un ouvrage de mathématiques, il est très rare que cette notation soit utilisée pour dénoter autre chose qu'une dérivée partielle. Dans un ouvrage de mécanique ou de physique, il est assez probable que cette notation représente toute la Jacobienne.

Lorsque $f$ est une fonction scalaire de classe $\mathcal{C}^1$ sur un ouvert de $\mathbb{R}^m$, on définit le gradient de $f$ comme

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}. \tag{4.3}$$

On remarque que dans ce cas que le gradient $\nabla f$ est la transposée de la Jacobienne.

## 4.2 Dérivée partielle par rapport à une variable vectorielle

Nous introduisons maintenant la notion de dérivée partielle par rapport à une variable vectorielle. Cette notion ressemble beaucoup à la notion de Jacobienne. Nous allons juste l'illustrer sur un exemple.

$$\begin{aligned} \boldsymbol{f} \colon \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 &\to \mathbb{R}^3, \\ \big(t, \underbrace{(x_1, x_2, x_3)}_{\boldsymbol{x}}, \underbrace{(v_1, v_2, v_3)}_{\boldsymbol{v}}\big) &\mapsto \boldsymbol{f}\big(t, (x_1, x_2, x_3), (v_1, v_2, v_3)\big). \end{aligned}$$

Par $\frac{\partial f}{\partial \boldsymbol{v}}$, on dénote la matrice constituée des trois dernières colonnes de la jacobienne. Par $\frac{\partial f}{\partial \boldsymbol{x}}$ on dénote la matrice constituée des colonnes 2 à 4 de la jacobienne. Ainsi

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{v}} = \begin{bmatrix} \frac{\partial f_1}{\partial v_1} & \frac{\partial f_1}{\partial v_2} & \frac{\partial f_1}{\partial v_3} \\ \frac{\partial f_2}{\partial v_1} & \frac{\partial f_2}{\partial v_2} & \frac{\partial f_2}{\partial v_3} \\ \frac{\partial f_3}{\partial v_1} & \frac{\partial f_3}{\partial v_2} & \frac{\partial f_3}{\partial v_3} \end{bmatrix}.$$

Et

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{bmatrix}.$$

Ainsi la dérivée partielle d'une fonction $f$ par rapport à une variable vectorielle $\boldsymbol{q}$ est la matrice dont la $i^{\text{e}}$ colonne est la dérivée partielle de la fonction $f$ par rapport à la $i^{\text{e}}$ composante de la variable vectorielle $\boldsymbol{q}$. La dérivée partielle par rapport à une variable vectorielle est toujours une sous-matrice de la Jacobienne. On l'obtient à partir de la Jacobienne en ne gardant que les colonnes qui correspondent à des composantes de la variable vectorielle.

De même, on peut définir le gradient d'une fonction scalaire par rapport à une variable vectorielle. Soit une fonction

$$\begin{aligned} \rho \colon \mathbb{R} \times \mathbb{R}^3 &\to \mathbb{R}, \\ (t, \boldsymbol{x}) &\mapsto \rho(t, \boldsymbol{x}). \end{aligned}$$

On pose alors

$$\nabla_{\boldsymbol{x}} \rho(t, \boldsymbol{x}) := \left(\frac{\partial \rho}{\partial \boldsymbol{x}}\right)^{\top} = \begin{bmatrix} \frac{\partial \rho}{\partial x_1} \\ \frac{\partial \rho}{\partial x_2} \\ \frac{\partial \rho}{\partial x_3} \end{bmatrix}. \tag{4.4}$$

Ce genre de notations est très utilisée en physique pour indiquer que l'on ne calcule certains opérateurs différentiels (gradient, divergence, rotationnel) que suivant les variables spatiales, *i.e.*, à temps fixé. Nous ne rappelons pas dans ce mini-poly les définitions de rotationnel et de divergence.

# 5 Exercices

Certains de ces exercices sont purement calculatoires. C'est voulu et il est recommandé de ne pas faire l'impasse sur ces exercices.

**Exercice 1.**

Soit $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto \sin(\cos(x))$. Calculer $f'$, $f''$, $f'''$ et $f''''$.

**Exercice 2.**

Calculer les dérivées partielles suivant $x$ et $y$ des fonctions suivantes : (a) $f\colon (x,y) \mapsto y - \sin(x)$, (b) $f\colon (x,y) \mapsto \frac{x+y^2}{x+y}$, (c) $f\colon (x,y) \mapsto \sin(\cos(xy))$.

**Exercice 3.**

Soit :

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto x - y \end{aligned}$$

(i) Calculer $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$.

(ii) Calculer $\frac{\partial f(u,v)}{\partial u}$ et $\frac{\partial f(u,v)}{\partial v}$.

(iii) Calculer $\frac{\partial f(u,v)}{\partial x}$ et $\frac{\partial f(u,v)}{\partial y}$.

**Exercice 4.**

Soit

$$\begin{aligned} f\colon \mathbb{R}^3 &\to \mathbb{R}, \\ (x,y,z) &\mapsto f(x,y,z). \end{aligned}$$

de classe $\mathcal{C}^1$. Calculer en fonction de $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, et de $\frac{\partial f}{\partial z}$,

a) $\dfrac{\partial f(u,v,w)}{\partial u}$,

b) $\dfrac{\partial f(u^2, uv, v^2+u)}{\partial u}$,

c) $\dfrac{\partial f(y,z,x)}{\partial x}$,

d) $\dfrac{\mathrm{d} f(t,t,t)}{\mathrm{d} t}$,

e) $\dfrac{\mathrm{d} f(x,x,x)}{\mathrm{d} x}$,

f) $\dfrac{\partial f(y^2, yu, x^2+u)}{\partial y}$.

Ne pas oublier de préciser en quels points doivent être calculées les dérivées partielles de la fonction $f$.