Chapitre II

Calcul Différentiel

Sommaire

II.1 Différentielle, dérivées partielles

II.1.1 Définitions, premières propriétés

On sait qu'une fonction f, définie d'un intervalle ouvert $I \subset \mathbb{R}$ à valeurs dans $\mathbb{R}$, est dérivable en $x \in I$ si le taux de variation admet une limite, notée alors $f'(x)$, lorsque h tend vers 0 :

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \varepsilon(h),$$

où $\varepsilon(h)$ tend vers 0 quand h tend vers 0.

On a alors

$$f(x+h) = f(x) + f'(x)h + \varepsilon(h)\ |h|\,, \tag{II.1.1}$$

(on a juste changé le signe de $\varepsilon(h)$ dans le cas où h était négatif). Ce développement exprime le fait que la fonction peut être approchée à l'ordre 1 au voisinage de x par une application affine.

Inversement, on vérifie immédiatement que si une fonction f admet un développement limité du type de (II.1.1) :

$$f(x+h) = f(x) + \gamma h + \varepsilon(h)\ |h|\,,$$

alors la fonction est dérivable en x, et le coefficient du terme de premier ordre est $\gamma = f'(x)$, la dérivée de f en x.

Cette approche s'étend sans difficultés au cas où la fonction est à valeurs vectorielles : $f = (f_1, \ldots, f_m)$. On peut définir la dérivée de chacune des composantes f_i par rapport à la variable d'espace $x \in \mathbb{R}$, la dérivée $f'(x)$ s'écrit

$$f'(x) = (f_1'(x), f_2'(x), \ldots, f_m'(x)),$$

et le développement limité est simplement écrit dans $\mathbb{R}^m$.

Nous allons nous intéresser maintenant la généralisation de ces notions au cas où l'espace de départ lui-même peut être de dimension strictement supérieure à 1, l'objet typique étudié à partir de maintenant sera donc une application

$$f \ : \ x = (x_1 \ldots, x_n) \in \mathbb{R}^n \longmapsto f(x) = (f_1(x), \ldots, f_m(x)) \in \mathbb{R}^m$$

où chacune des m composantes f_i est une application de $\mathbb{R}^n$ dans $\mathbb{R}$.

Exemple II.1.1. (Champ de vecteurs)
Un champ de vecteurs dans l'espace physique est une application qui à chaque point $\mathbb{R}^3$ associe un vecteur de $\mathbb{R}^3$. On le note en général $u = (u_1, u_2, u_3)$ où chaque composante u_i est une fonction de $x = (x_1, x_2, x_3)$. Il peut encoder un champ de vitesses fluides à un instant donné, ou un champ de déformations infinitésimales au sein d'un objet élastique déformable. Un champ de vecteurs dans le plan (par exemple un champ de vitesses horizontales à la surface d'une eau bien plate) correspond au cas $n = m = 2$.

Exemple II.1.2. (Champ scalaire)
On parle d'un champ scalaire lorsque l'espace d'arrivée est $\mathbb{R}$ (cas $n = 3$ et $m = 1$ pour un champ de l'espace physique). Cela correspond par exemple au champ de température dans une zone de l'espace à un instant donné.

Exemple II.1.3. On peut considérer les versions dynamiques des exemples ci-dessus en rajoutant une variable de temps dans l'espace de départ. Par exemple un champ de vitesse variable en temps correspond au cas $n = 4$, $m = 3$, il peut être considéré comme une application $\mathbb{R}^3 \times \mathbb{R}$ dans $\mathbb{R}^3$, qui à chaque $(x, t) = (x_1, x_2, x_3, t)$ fait correspondre un vecteur $(u_1(x, t), u_2(x, t), u_3(x, t))$.

Si l'on cherche à écrire un développement limité du type (II.1.1), l'identité est à valeurs dans $\mathbb{R}^m$, et la variation h de la variable x de l'espace de départ vit dans $\mathbb{R}^m$. Le terme $f'(x)h$ doit être remplacé par un terme à valeurs dans $\mathbb{R}^m$, qui dépend linéairement du vecteur $h \in \mathbb{R}^m$, il s'écrit donc sous la forme d'une application linéaire appliquée à la variation h. Ces considérations conduisent à la définition suivante de la notion de *différentiabilité*, qui exprime que certaines applications peuvent être approchées localement par une application affine.

Notation II.1.1. (Image par une application linéaire et produit matrice vecteur)
Nous adoptons dans ce qui suit une convention courante dans le contexte du calcul différentiel (et en particulier en mécanique des fluides), qui est de noter $F \cdot x$ l'image par une application linéaire F d'un vecteur x. De la même manière, si A est une matrice, écrira $A \cdot x$ le produit matrice vecteur. Cette notation est issue de ce que l'on appelle le calcul tensoriel [1], qui n'est pas abordé en tant que tel dans ce cours.

1. On pourra se reporter à
`http://mms2.ensmp.fr/mmc_st_etienne_fort/calcul_tensoriel/polycop/tenseurs_poly.pdf`
pour une présentation détaillée de ces notions.

Définition II.1.2. (Différentielle (•))
Soit f une application définie d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}^m$. On dit que f est différentiable en $x \in U$ s'il existe une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, notée $df(x)$, telle que

$$f(x+h) = f(x) + df(x) \cdot h + \varepsilon(h)\,\|h\| \qquad \text{(II.1.2)}$$

où $\varepsilon(h)$ est une application de $\mathbb{R}^n$ dans $\mathbb{R}^m$, telle que $\|\varepsilon(h)\|$ tend vers 0 quand h tend vers 0. On pourra aussi utiliser la notation dite de *Landau* en écrivant $o(h)$ à la place de $\varepsilon(h)\,\|h\|$. On appelle alors cette application la *différentielle* de f en x.

Proposition II.1.3. Toute application différentiable en un point est continue en ce point.

Démonstration. C'est une conséquence directe du développement limité (II.1.2), qui assure que $f(x+h)$ tend vers $f(x)$ quand h tend vers 0. □

Définition II.1.4. (Continue différentiabilité (•))
Une application f d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}^m$ est dite *continûment différentiable* sur U si elle est différentiable en tout $x \in U$, et si l'application $x \mapsto df(x)$ est continue sur U (l'espace d'arrivée est muni canoniquement de la norme d'opérateur subordonnée à la norme euclidienne, voir proposition I.8.3, page 23).

Exercice II.1.1. (•) Montrer que toute application affine définie d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}^m$:

$$x \longmapsto f(x) = A \cdot x + b\,,\ A \in \mathcal{M}_{mn}(\mathbb{R})\,,\ b \in \mathbb{R}^m,$$

est continûment différentiable sur cet ouvert, et préciser sa différentielle.

Proposition II.1.5. (Linéarité de la différentiation)
Soient f et g des applications définies d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}^m$. Si f et g sont différentiables en $x \in U$, alors, pour tous λ, μ réels, l'application $\lambda f + \mu g$ est différentiable, et l'on a

$$d(\lambda f + \mu g) = \lambda df + \mu dg.$$

Démonstration. C'est une conséquence immédiate de la définition de la différentiabilité. □

Corollaire II.1.6. L'ensemble $C^1(U, \mathbb{R}^m)$ des applications continûment différentiables sur un ouvert U de $\mathbb{R}^n$ est un espace vectoriel.

Dérivées partielles

L'application différentielle, quand elle existe, peut être représentée par une matrice appelée *matrice jacobienne*, dont les éléments sont les *dérivées partielles* des composantes de f par rapport aux variables x_j, pour $j = 1, \ldots, n$.

Définition II.1.7. (Dérivées partielles, matrice jacobienne)
Soit f une application définie d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}^m$, et $x = (x_1, \ldots, x_n) \in U$. On dit que f admet en x une dérivée partielle par rapport à la variable x_j si l'application partielle

$$y \longmapsto f(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n) \in \mathbb{R}^m$$

définie d'un voisinage de x_j et à valeurs dans $\mathbb{R}^m$, est dérivable en $y = x_j$.

La dérivée de la i-ème composante de f par rapport à la variable x_j, telle que définie ci-dessus, est notée [2]

$$\frac{\partial f_i}{\partial x_j}(x) \text{ ou } \partial_{x_j} f_i(x) = \lim_{s\to 0} \frac{f_i(x+se_j)-f_i(x)}{s},$$

où l'on a noté e_j le j-ème vecteur de la base canonique de $\mathbb{R}^n$. Si toutes les dérivées partielles des f_i par rapport aux x_j existent, on appelle *matrice Jacobienne* la matrice

$$J = \begin{pmatrix} \partial_{x_1} f_1 & \partial_{x_2} f_1 & \dots & \partial_{x_n} f_1 \\ \partial_{x_1} f_2 & \partial_{x_2} f_2 & \dots & \partial_{x_n} f_2 \\ \vdots & \vdots & \vdots & \vdots \\ \partial_{x_1} f_m & \partial_{x_2} f_m & \dots & \partial_{x_n} f_m \end{pmatrix} \in \mathcal{M}_{mn}(\mathbb{R}).$$

Proposition II.1.8. Soit f une application définie d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}^m$. Si f est différentiable en $x \in U$, alors sa différentielle admet pour représentation matricielle la matrice jacobienne J définie ci-dessus : c'est-à-dire f admet le développement limité

$$f(x+h) = f(x) + J(x)\cdot h + o(h).$$

Démonstration. Si f est différentiable en x, on peut écrire le développement limité composante par composante), en prenant la variation h de la forme se_j, où s est un réel et e_j un vecteur unitaire de la base canonique $\mathbb{R}^n$: pour tout $i = 1,\dots, m$, tout $j = 1, \dots, m$,

$$f_i(x+se_j) = f_i(x) + s(df(x)\cdot e_j)_i + \varepsilon_i(se_j)\,|s|\,.$$

On a donc

$$(df(x)\cdot e_j)_i = \lim_{s\to 0} \frac{f_i(x+se_j)-f_i(x)}{s} = \frac{\partial f_i}{\partial x_j}$$

d'après la définition II.1.7), c'est-à-dire le coefficient (i,j) de la matrice jacobienne J. Ce coefficient s'identifie donc à la composante i de l'image du j-ème vecteur de la base canonique par la différentielle, ce qui termine la preuve. □

Remarque II.1.9. (Très important)
On prendra garde au fait que, dès que la dimension n de l'espace de départ est strictement plus grande que 1, l'existence d'une matrice jacobienne (c'est à dire l'existence de toutes les dérivées partielles) *n'implique pas* la différentiabilité. Une application peut même admettre une matrice jacobienne en un point sans pour autant être continue en ce point (voir exercice ci-dessous). On verra néanmoins que si une application admet sur un ouvert U des dérivées partielles qui sont toutes continues sur U, alors l'application est continûment différentiable sur U

2. Nous commettons ici un abus de notation si courant qu'il nous paraît préférable de le commettre en connaissance de cause, plutôt que de le contourner. Dans ce qui suit x_j dans $\partial f_i/\partial x_j$ encode le fait que l'on dérive par rapport à la j–ème variable. Mais quand on écrit $x = (x_1, \dots, x_j, \dots, x_n)$, x_j désigne un réel, qui est la valeur particulière de la j–ième composante du point x.

Exercice II.1.2. Montrer que la fonction

$$(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\} \longmapsto \frac{xy}{x^2+y^2}\,,\ f(0,0)=0,$$

admet en $(0,0)$ des dérivées partielles, mais n'est pas continue en ce point (et donc non différentiable d'après la proposition II.1.3.

Lorsqu'une fonction différentiable est à valeurs dans $\mathbb{R}$, la différentielle est une *forme linéaire*, c'est-à-dire une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$. Elle peut alors s'exprimer[3] à l'aide du produit scalaire canonique sur $\mathbb{R}^n$, et s'identifie par ce biais à un vecteur de $\mathbb{R}^n$. C'est ce vecteur que l'on appelle gradient de f.

Définition II.1.10. (Gradient)
Soit f une application définie d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}$. On suppose que f est différentiable en $x \in U$. Il existe alors un unique vecteur, noté $\nabla f(x)$, tel que

$$df(x)\cdot h = \langle \nabla f(x) \,|\, h \rangle \quad \forall h \in \mathbb{R}^n,$$

où $\langle \nabla f(x) \,|\, h\rangle$ représente le produit scalaire canonique sur $\mathbb{R}^n$.
Conformément à la proposition II.1.8, ce gradient s'écrit à l'aide des dérivées partielles

$$\nabla f(x) = (\partial_{x_1} f(x), \partial_{x_2} f(x), \dots, \partial_{x_n} f(x)) \in \mathbb{R}^n.$$

Il sera utile dans certaines applications d'utiliser la notion de *gradient partiel*. Cela consiste simplement à considérer une fonction de n variables comme une fonction d'une partie de ces variables, les autres étant gelées. La définition ci-dessous précise cette notion dans le cas général, que nous illustrons au préalable sur un cas particulier. Considérons une fonction de $\mathbb{R}^3$ dans $\mathbb{R}$, différentiable en un point $x=(x_1,x_2,x_3)$. Son gradient en x est un vecteur de $\mathbb{R}^3$. Si on la considère maintenant comme une fonction de (x_1,x_2), avec x_3 fixé à sa valeur correspondant à x, le gradient de cette nouvelle fonction est un vecteur de $\mathbb{R}^2$, que l'on pourra noter $\nabla_{x_1x_2} f$, ou $\nabla_{x_{12}} f$.

Définition II.1.11. (Gradient partiel)
Soit f une application définie d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}$, différentiable en un point $x \in U$. On écrit $\mathbb{R}^n = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_p}$, de telle sorte que x peut s'écrire[4] $x=(x_1,\dots,x_p)$, avec $x_i \in \mathbb{R}^{n_i}$. Pour i entre 1 et p, on considère la fonction partielle qui ne dépend que du vecteur x_i, les autres étant figées. Le gradient de cette fonction partielle est un vecteur de $\mathbb{R}^{n_i}$, on le note $\nabla_{x_i} f$.

Cette notion de gradient partiel est notamment très utile en pratique lorsque l'on définit un potentiel d'interaction sur un système de particules localisées en q_1, q_2, …, q_N, chacun des q_i

3. Lorsque l'on travaille sur $\mathbb{R}^n$, l'usage de la base orthonormée canonique et du produit scalaire canonique sont tellement naturels que l'on a tendance à identifier spontanément la différentielle et le gradient. On prendra cependant garde au fait que le gradient n'est pas défini de façon intrinsèque. Contrairement à la différentielle, qui est une application définie de façon non ambigüe par le développement limité, ce gradient dépend du produit scalaire choisi. Ce point est particulièrement sensible dans certaines situations, notamment en dimension infinie, où plusieurs produits scalaires "naturels" peuvent co-exister.

4. Attention, x_i désigne dans ce qui suit non plus une variable scalaire, mais un groupe de n_i variables scalaires.

étant un point de l'espace physique $\mathbb{R}^3$. Si l'on définit un potentiel d'interaction sur le système comme une fonction Φ de $q = (q_1, \dots, q_N) \in \mathbb{R}^{3N}$, la force exercée sur la particule i dérivant de ce potentiel d'interaction est simplement $-\nabla_{q_i}\Phi \in \mathbb{R}^3$. On se reportera à l'exercice II.3.7, page 48, pour une étude plus approfondie de ces systèmes de particules en interaction, et l'utilisation dans ce cadre de la notion de gradient partiel.

Remarque II.1.12. Si l'on considère le gradient partiel d'une fonction de $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ vis-à-vis d'une unique variable x_i, on retrouve la notion de dérivée partielle par rapport à x_i déjà introduite.

Définition II.1.13. (Différentielle partielle)
On définit de façon tout à fait analogue une notion de différentielle partielle, pour des applications de $\mathbb{R}^n$ vers $\mathbb{R}^m$. Si l'on écrit comme précédemment

$$x = (x_1, \dots, x_p) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_p},$$

la différentielle partielle par rapport à x_i, que l'on notera[5] $\partial_{x_i} f$, est alors une application linéaire de $\mathbb{R}^{n_i}$ dans $\mathbb{R}^m$.

Exercice II.1.3. (Dépendance du gradient vis-à-vis du produit scalaire)
Comme indiqué précédemment, le gradient dépend du produit scalaire sous-jacent. Dans le cadre de ce cours, nous utiliserons cette notion essentiellement en lien avec le produit scalaire canonique de $\mathbb{R}^n$, parfois sans re-préciser qu'il s'agit bien de ce produit scalaire. Cette exercice illustre le fait qu'il peut être naturel, dans certains contextes, de travailler avec d'autres produits scalaires, et permet de comprendre comment le gradient se voit modifié.
On se place dans $\mathbb{R}^{3N} = \mathbb{R}^3 \times \cdots \times \mathbb{R}^3$ pour représenter les vitesses dans l'espace physique de N particules, de masses m_1, m_2, ..., m_N toutes strictement positives. On considère la fonction qui à un jeu de vitesses associe l'énergie cinétique

$$E(u) = E(u_1, \dots, u_N) = \frac{1}{2} \sum_{n=1}^{N} m_i \left| u_n \right|^2.$$

Montrer que E est différentiable, et calculer son gradient pour le produit scalaire canonique, puis pour le produit scalaire $\langle \cdot \mid \cdot \rangle_m$ pondéré par la collection de masses $m = (m_1, \dots, m_N)$, défini par

$$\langle u \mid v \rangle_m = \sum_{n=1}^{N} m_i \langle u_n \mid v_n \rangle,$$

où $\langle u_n \mid v_n \rangle$ représente le produit scalaire canonique sur $\mathbb{R}^3$.

Définition II.1.14. (Point critique / stationnaire)
Soit f une application continûment différentiable d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}$. On appelle point *critique* ou point *stationnaire* tout point $x \in U$ en lequel le gradient s'annule.

5. Cette notation est la plus couramment utilisée, et nous en recommandons l'usage, tout en reconnaissant qu'il aurait été assez naturel d'utiliser la notation d_{x_i}, puisqu'il s'agit d'une application linéaire, définie de façon intrinsèque comme associant à tout vecteur un vecteur, indépendamment du choix d'une base. La notation '∂' a pour l'instant été utilisée pour représenter des dérivées partielles, qui repose sur le choix d'un système de coordonnées. Dans le cas présent, on est un peu entre les deux : on souhaite représenter une application linéaire, mais définie sur un sous-espace dont la définition repose sur le choix d'un système de coordonnées.

Exercice II.1.4. ($\bullet$) Justifier l'appellation *stationnaire* dans la définition précédente.

D'après la proposition II.1.8, si une application est continûment différentiable sur un ouvert U, alors la matrice jacobienne est définie en tout point de cet ouvert, et la correspondance $x \longmapsto J(x)$ est continue. La proposition suivant assure la réciproque de cette propriété.

Proposition II.1.15. ($\bullet\bullet\bullet$) Soit f une application définie d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}^m$. On suppose que la matrice jacobienne $J(x)$ est définie en chaque point x de U, et que l'application $x \mapsto J(x)$ est continue. Alors f est continûment différentiable sur U.

Démonstration. On écrit la démonstration pour le cas $n = 2$, et l'on suppose que $(0, 0) \in U$ pour simplifier les notations. Nous allons montrer la différentiabilité en $(0, 0)$, la démonstration pour les autres points étant essentiellement la même. Pour h_1, h_2 suffisamment petits, on a

$$f(h_1, h_2) = f(h_1, h_2) - f(h_1, 0) + f(h_1, 0) - f(0, 0) + f(0, 0).$$

On écrit

$$f(h_1, 0) - f(0, 0) = h_1 \int_0^1 \partial_1 f(th_1, 0)\, dt.$$

Comme $x \mapsto J(x)$ est continue, tous les coefficients de la matrice J sont des fonctions continues en x. On a donc en particulier $\partial_1 f(th_1, 0) = \partial_1 f(0, 0) + \varepsilon(th_1)$, et ainsi

$$f(h_1, 0) - f(0, 0) = h_1 \partial_1 f(0, 0) + h_1 \varepsilon(h_1).$$

On a de la même manière

$$f(h_1, h_2) - f(h_1, 0) = h_2 \int_0^t \partial_2 f(h_1, th_2)\, dt,$$

avec, par continuité des dérivées partielles,

$$\partial_2 f(h_1, th_2) = \partial_2 f(0, 0) + \varepsilon(h).$$

On a donc

$$f(h_1, h_2) - f(h_1, 0) = h_2 \partial_2 f(0, 0) + h_2 \varepsilon(h).$$

On a donc finalement

$$f(h_1, h_2) = f(0, 0) + h_1\, \partial_1 f(0, 0) + h_2\, \partial_2 f(0, 0) + o(h),$$

qui exprime la différentiabilité de f en $(0, 0)$, et de la même manière en tout point de l'ouvert U. La différentielle peut s'exprimer matriciellement à partir de la jacobienne $J = [\partial_1 f, \partial_2 f]$ (écriture de la matrice en colonnes, chacune des dérivées partielles étant un vecteur de $\mathbb{R}^m$). Les dérivées partielles étant continues, la correspondance $x \mapsto J(x)$ est continue, l'application est donc continûment différentiable sur U.

La réciproque est une conséquence directe de la proposition II.1.8. $\square$

Proposition II.1.16. (Différentielle de la composée de deux applications)
Soient g une application définie d'un ouvert $U \in \mathbb{R}^n$ dans $\mathbb{R}^p$, et f définie d'un ouvert $U \in \mathbb{R}^p$ dans $\mathbb{R}^m$. On suppose que $g(U) \subset V$. Si g est différentiable en $x \in U$ et f est différentiable en $g(x)$, alors $f \circ g$ est différentiable en x, et l'on a

$$d(f \circ g)(x) = df(g(x)) \circ dg(x).$$

Démonstration. On a

$$\begin{aligned} f\circ g(x+h) = f(g(x+h)) &= f\left(g(x) + dg(x)\cdot h + o(h)\right) \\ &= f(g(x)) + df(g(x))\cdot(dg(x)\cdot h + o(h)) + o(dg(x)\cdot h + o(h)) \\ &= f\circ g\,(x) + \left(df(g(x))\circ dg(x)\right)\,\cdot h + o(h), \end{aligned}$$

qui exprime la différentiabilité de $f\circ g$, avec l'expression annoncée de la différentielle. □

Calcul différentiel

Nous regroupons ici quelques considérations sur la pratique effective du calcul différentiel, et en particulier les notations dx_1, $dx_1 + dx_2$, etc ...

Si l'on considère par exemple une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$, définie par $f(x) = f(x_1, x_2) = x_1^2 + x_2^3 + x_1x_2$, on écrira

$$df = d(x_1^2 + x_2^3 + x_1x_2) = 2x_1\,dx_1 + 3x_2\,dx_2 + x_1\,dx_2 + x_2\,dx_1 = (2x_1 + x_2)\,dx_1 + (3x_2 + x_1)\,dx_2.$$

Dans ce qui précède, dx_1 représente par exemple la différentielle de la fonction $(x_1, x_2) \mapsto x_1$, qui est simplement l'application qui à (h_1, h_2) associe h_1, qui peut se représenter matriciellement par $[1\ \ 0]$. La différentielle de f en (x_1, x_2) est donc représentée dans la base canonique par

$$J(x) = [2x_1 + x_2 \quad 3x_2 + x_1].$$

De façon plus générale, on écrira [6]

$$df(x_1, x_2) = \frac{\partial f}{\partial x_1}(x_1, x_2)\,dx_1 + \frac{\partial f}{\partial x_2}(x_1, x_2)\,dx_2.$$

On prendra garde au fait que l'expression dx_1 dépend du contexte. La même expression correspondre à une application de $\mathbb{R}^3$ dans $\mathbb{R}$, auquel cas dx_1 est représentée matriciellement par $[1\ \ 0\ \ 0]$.

Si l'application est à valeurs vectorielles, par exemple dans $\mathbb{R}^2$:

$$f(x_1, x_2) = \begin{bmatrix} x_1^2 + x_2^3 + x_1x_2 \\ x_1 \end{bmatrix},$$

on écrira de la même manière

$$df(x_1, x_2) = \begin{bmatrix} (2x_1 + x_2)\,dx_1 + (3x_2 + x_1)\,dx_2 \\ dx_1 \end{bmatrix},$$

6. Nous commettons ici un abus de notation courant, auquel il convient de s'habituer car il est très répandu : le 'x_1' qui apparaît dans ∂_{x_1} et dans dx_1 représente une variable générique vis à vis de laquelle on dérive, alors que le 'x_1' de $df(x_1, x_2)$ est un nombre réel, première coordonnée du point en lequel on dérive la fonction. On devrait en toute rigueur distinguer ces deux acceptions en utilisant des noms différents, par exemple

$$df(a_1, a_2) = \partial_{x_1} f(a_1, a_2)\,dx_1 + \partial_{x_2} f(a_1, a_2)\,dx_2.$$

qui peut se représenter matriciellement par

$$J(x_1,x_2) = \begin{bmatrix} 2x_1+x_2 & 3x_2+x_1 \\ 1 & 0 \end{bmatrix},$$

de telle sorte que, pour tout h dans $\mathbb{R}^2$, on a le développement limité

$$f(x+h) = f(x) + J(x)\cdot h + o(h),$$

où $J(x)\cdot h$ représente le produit matrice vecteur, comme indiqué précédemment.

Exercice II.1.5. Calculer la différentielle de la forme de Minkovski

$$f \ : \ (x,y,z,t) \in \mathbb{R}^4 \longmapsto x^2+y^2+z^2-c^2t^2,$$

avec $c > 0$ (vitesse de la lumière).

Récapitulatif

Les développement ci-dessus décrivent des manières variées d'exprimer qu'une fonction peut être approchée localement par une fonction affine. Nous récapitulons ici ces différentes manières, en rappelant leur cadre d'utilisation et les liens entre elles. Dans ce qui suit f désigne, sauf mention contraire, une application de $\mathbb{R}^n$ dans $\mathbb{R}^m$.

Comme on l'a vu, f est dite différentiable en $x\in\mathbb{R}^n$ s'il existe une application $df(x)$ telle que

$$f(x+h) = f(x) + df(x)\cdot h + o(h).$$

Le terme $df(x)\cdot h$ désigne l'image par $df(x)$ du vecteur h. Cette expression est intrinsèque, au sens où elle ne dépend pas du choix d'une base. En pratique, on assimile souvent une application linéaire et son écriture matricielle dans la base canonique, mais il est important de garder en tête la différence entre les deux. Cette approche permet notamment une extension immédiate de la définition en dimension infinie, dans un contexte où les bases sont inutilisables.

Si f est différentiable en x, alors (proposition II.1.8) la différentielle admet une représentation matricielle dans la base canonique qui est la matrice jacobienne $J = (\partial_{x_j} f_i)$. On a donc

$$f(x+h) = f(x) + J(x)\cdot h + o(h),$$

où $J(x)\cdot h$ est maintenant un produit matrice-vecteur. L'objet J dépend du choix de la base. l'expression ci-dessus peut être détaillée, de différentes manières. On peut l'écrire composante par composante

$$f_i(x+h) = f_i(x) + \sum_{j=1}^m \partial_{x_j} f_i(x)\ h_j + o(h),$$

ou de façon globale, avec e_i le $i-$ème vecteur de la base canonique de $\mathbb{R}^m$:

$$f(x+h) = f(x) + \sum_{i=1}^n \sum_{j=1}^m \partial_{x_j} f_i(x)\ h_j\ e_i + o(h),$$

Comme il a été précisé, l'existence de dérivées partielles en un point ne garantit pas la différentiabilité. En revanche (proposition II.1.15), si les dérivées partielles sont définies et continues sur un ouvert, alors la fonction est continûment différentiable sur cet ouvert.

Lorsque la fonction est à valeurs dans $\mathbb{R}$ (cas $m = 1$), la matrice jacobienne est une matrice ligne, et l'application différentielle $df(x)$ est une forme linéaire. On peut alors écrire $df(x) \cdot h$ sous la forme d'un produit scalaire $\langle g \,|\, h \rangle$, où g est appelé gradient de f au point x, et noté $\nabla f(x)$. On a alors le développement

$$f(x+h) = f(x) + \langle \nabla f(x) \,|\, h \rangle + o(h).$$

Le vecteur $\nabla f(x)$ dépend du produit scalaire choisi. Lorsque ce choix n'est pas précisé, il s'agit du gradient associé au produit scalaire canonique sur l'espace euclidien $\mathbb{R}^n$. Lorsque l'on se place dans la base canonique de $\mathbb{R}^n$, que l'on considère muni du produit scalaire canonique, le vecteur $\nabla f(x)$ est représenté dans la base canonique par la matrice-ligne $J(x)$:

$$\nabla f(x) = (\partial_{x_1} f, \partial_{x_2} f, \dots, \partial_{x_n} f).$$

II.1.2 Théorème fondamental de l'analyse

Le théorème fondamental de l'analyse peut prendre plusieurs formes selon le sens que l'on donne à la notion d'intégrale. Le chapitre sur l'intégrale de Lebesgue montre que l'on peut définir cette intégrale pour des classes très générales de fonctions. Nous nous en limiterons ici à une définition plus classique de l'intégrale, en nous limitant à des fonctions continûment différentiables, de telle sorte que l'on n'aura besoin d'intégrer que des fonctions continues. On pourra donc s'en tenir à la notion d'intégrale de Riemann. L'objet de cette section est de généraliser au cas vectoriel la propriété portant sur les fonctions de $\mathbb{R}$ dans $\mathbb{R}^m$: pour toute fonction f continûment dérivable sur $]a,b[$, à valeurs dans $\mathbb{R}^m$, pour tout x dans $]a,b[$, tout h tel que $x+h \in]a,b[$, on a

$$f(x+h) = f(x) + \int_x^{x+h} f'(s)\, ds.$$

Cette intégrale peut s'écrire différemment en introduisant la fonction $t \in [0,1] \mapsto f(x+th)$, donc la dérivée en t est $f'(x+th)\, h$. On a

$$f(x+h) = f(x) + \int_0^1 f'(x+th) h \, dt.$$

Le théorème suivant généralise cette propriété aux fonctions de plusieurs variables.

Théorème II.1.17. Soit f une application définie d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}^m$, continûment différentiable sur U, et h tel que le segment

$$[x, x+h] = \{x + \theta h\,,\ h \in [0,1]\}$$

soit inclus dans U. On a alors

$$f(x+h) = f(x) + \int_0^1 df(x+th) \cdot h \, dt.$$

Démonstration. On introduit l'application Φ de $[0,1]$ dans $\mathbb{R}^n$, définie par

$$\Phi \ : \ t \in [0,1] \longmapsto \Phi(t) = f(x+th).$$

D'après la proposition II.1.16, cette application est continûment différentiable (on dira plus simplement *dérivable*, puisqu'il s'agit d'une fonction de $\mathbb{R}$ dans $\mathbb{R}^m$), de dérivée

$$\Phi'(t) = df(x+th)\cdot h.$$

On a donc

$$f(x+h) - f(x) = \Phi(1) - \Phi(0) = \int_0^1 \Phi'(t)\,dt = \int_0^1 df(x+th)\cdot h\,dt,$$

qui est l'identité annoncée. □

Ce théorème nous conduit naturellement au théorème des accroissements finis pour les fonctions de plusieurs variables, qui exprime un principe simple que l'on retrouve dans différents contextes [7] : si l'on contrôle les variations d'une certaine quantité le long d'un chemin de longueur finie qui va de x vers $x+h$, alors on peut contrôler la différence des valeurs entre $x+h$ et x.

Théorème II.1.18. (des accroissements finis)
Soit f une application définie d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}^m$, continûment différentiable sur U, et h tel que le segment

$$[x, x+h] = \{x + \theta h\,,\ h \in [0,1]\}$$

soit inclus dans U. Alors

$$\|f(x+h) - f(x)\| \leq \max_{t\in[0,1]} \|df(x+th)\|\,\|h\|\,.$$

Démonstration. Notons en premier lieu que, la différentielle df étant continue sur le compact $[x, x{+}h]$, elle est bien bornée et atteint ses bornes, en particulier le max ci-dessus est bien défini comme un réel positif. On prend la norme de l'identité établie dans le théorème précédent : On a alors

$$\|f(x+h) - f(x)\| = \left\|\int_0^1 df(x+th)\cdot h\,dt\right\| \leq \int_0^1 \|df(x+th)\cdot h\|\;dt$$

$$\leq \int_0^1 \|df(x+th)\|\,\|h\|\;dt \leq \max_{t\in[0,1]} \|df(x+th)\|\,\|h\|\,,$$

qui est bien l'inégalité annoncée. □

Exercice II.1.6. L'inégalité établie précédemment est-elle valide si l'on munit $\mathbb{R}^n$ et $\mathbb{R}^m$ d'autres normes que la norme euclidienne ?

7. On pourra penser à une version *Tour de France* de cette propriété très générale : si un coureur cycliste effectue un parcours de 10 km sur une route dont la pente n'excède pas 7 %, il sait qu'il n'aura pas monté en altitude de plus de 10 km $\times$ 0.07 = 700 m.