Calcul différentiel et équations différentielles

Sylvie Benzoni-Gavage

Calcul différentiel et équations différentielles

Cours et exercices corrigés

2e édition

DUNOD

Illustration de couverture : © Digitalvision

ISBN 978-2-10-070611-2

Table des matières

AVANT-PROPOS

Cette seconde édition, revue et corrigée, se veut plus progressive notamment dans la partie qui traite de calcul différentiel. Idéalement, un étudiant en mathématiques devrait pouvoir suivre l'ouvrage pas à pas, depuis la deuxième année de licence jusqu'en master. Des exercices corrigés l'accompagnent au long de ce chemin. Souvent inspirés d'applications, ils sont l'occasion de mettre en pratique les outils introduits dans le texte, et leurs solutions mettent en évidence des relations entre les différents chapitres.

Si l'apprentissage du calcul différentiel a toujours été ressenti comme difficile, il l'est d'autant plus aujourd'hui que les études secondaires ne confrontent plus du tout les élèves à l'abstraction. Quant aux études supérieures, elles sont devenues très « compartimentées », de sorte que les étudiants ont souvent les plus grandes difficultés à mobiliser simultanément des connaissances venant de différentes parties des mathématiques: analyse, algèbre, géométrie, etc. Le calcul différentiel cristallise dans une certaine mesure ces difficultés. Pour autant, rien n'est inaccessible à qui s'en donne la peine.

En guise de motivation, l'étudiant devrait prendre conscience que la beauté de certains concepts et raisonnements épargne à celui qui les maîtrise de fastidieux calculs. C'est particulièrement vrai pour l'analyse des équations différentielles : il est fini le temps où de grands savants se disputaient la meilleure manière de *résoudre* des équations différentielles au moyen de formules plus ou moins explicites. Les travaux de Poincaré sont passés par là et une véritable révolution s'est opérée. Les apprentis mathématiciens actuels ont cette chance: une fois consolidées leurs connaissances de base en calcul différentiel, ils peuvent accéder à une compréhension profonde du comportement des solutions des équations différentielles sans avoir besoin de les calculer explicitement, ce qui est au mieux pénible, sauf pour des équations modèles très simples, et au pire impossible.

Curieusement, cette révolution n'a pas encore vraiment imprégné les autres disciplines, bien que les physiciens, mécaniciens, chimistes, etc. aient plutôt bien intégré la révolution suivante, celle du calcul numérique. S'il ne leur est pas spécifiquement destiné, ce livre et les références qu'il contient pourraient leur être utiles. C'est en tous cas mon souhait, en cette année de célébration du bicentenaire de Lagrange, dont les travaux ont été plus que largement diffusés dans ces disciplines connexes aux mathématiques que sont la mécanique et la physique.

Même s'ils sont antérieurs d'un siècle à ceux de Poincaré, certains fondements posés par Lagrange transparaissent notablement dans ce qui suit et font partie du bagage qu'on peut attendre d'un étudiant en mathématiques, physique ou mécanique. Puissent ces grands esprits du passé et tous les Euler, Laplace, Lyapunov, Noether, inspirer les nouvelles générations.

Lyon, le 30 septembre 2013.

Préface de la 1re édition

Je fais partie de cette génération nourrie à l'analyse des équations aux dérivées partielles (alors appelée abusivement *analyse numérique*), dans l'idée que l'analyse des équations différentielles ordinaires était « dépassée », au point qu'elle n'était guère plus enseignée qu'aux détours de cours de géométrie ou de mécanique. J'ai néanmoins pu apprécier au fil des années la richesse des interactions entre un domaine traditionnellement réservé aux géomètres, celui des *systèmes dynamiques*, et l'analyse appliquée à des modèles divers, que ce soient des équations différentielles ordinaires, des équations aux dérivées partielles, ou encore des équations différentielles fonctionnelles. C'est ce qui m'a incitée à mettre en place un cours d'équations différentielles ordinaires en première année du master « MAIM » (Mathématiques et Applications, Ingénierie Mathématique) à Lyon 1. Cet ouvrage est issu des notes rédigées pour l'occasion, et auparavant pour la préparation à l'épreuve de modélisation à l'agrégation, ainsi que de mes notes de calcul différentiel en troisième année de licence. La transformation en livre de ces notes éparses fut encouragée, sans qu'ils en aient nécessairement conscience, par plusieurs collègues : Michelle Schatzman et Denis Serre (que je remercie au passage pour leur fidèle et amical soutien), Francis Filbet (qui m'a fait bénéficier de sa toute fraîche expérience d'auteur chez Dunod), ainsi que ceux ayant manifesté leur intérêt pour mon « poly » (en ligne sur ma page personnelle). En espérant que cela en valait la peine, je remercie de tout c\oe ur mes proches pour leur patience (et bien plus) pendant ces mois passés à remettre l'ouvrage sur le métier. Je tiens en outre à remercier Pascal Noble, avec qui j'ai eu plaisir à enseigner cette matière et à qui je dois divers énoncés d'exercices, ainsi que des critiques constructives. Quant aux exercices de calcul différentiel, ils proviennent pour l'essentiel de sujets d'examen posés en L3: je remercie notamment Danièle Tarral, Laurent Pujo-Menjouet et Daniel Sondaz pour leurs contributions. Je remercie enfin Sarah Delcourte pour sa relecture attentive.

Malgré le soin que je me suis efforcée d'apporter à la rédaction, le lecteur[1] trouvera sûrement des imperfections, qu'il voudra bien me pardonner ou me signaler. Il pourra aussi regretter des omissions criantes à ses yeux : sur ce point je ne peux qu'assumer mes choix, dictés par mes goûts et la place allouée par l'éditeur.

Lyon, le 29 mai 2009.

1. Si je ne cède pas aux travers de la féminisation du langage, je n'en espère pas moins avoir autant de lectrices que de lecteurs !

INTRODUCTION

Cet ouvrage se présente en deux parties qu'il est conseillé d'aborder dans l'ordre, ou du moins est-il préférable d'avoir étudié le début de la première avant de s'attaquer à la deuxième.

La première partie reprend le calcul différentiel à la base, en passant un peu vite sur les fonctions d'une variable réelle (supposées relativement familières au public visé) pour arriver aux fonctions de plusieurs variables réelles et plus généralement aux fonctions définies sur des *ouverts* de $\mathbb{R}$-*espaces vectoriels normés*[1]. Un objectif avoué est d'amener progressivement le lecteur à se libérer des coordonnées et à être capable de mener des calculs les plus intrinsèques possibles, ce qui est évidemment indispensable en dimension infinie. La notion de *différentielle*, dont l'origine remonte à Leibniz, « co-inventeur » avec Newton du calcul différentiel, est centrale dans cette partie. Bien sûr le regard qu'on lui porte aujourd'hui est bien plus net qu'au XVII^e siècle, la notion de *limite* ayant été éclaircie au XIX^e.

Le calcul différentiel est présenté ici dans un cadre qui généralise de façon assez naturelle le calcul dans $\mathbb{R}^n$. La traduction dans $\mathbb{R}^n$ des résultats énoncés dans un espace vectoriel normé général est en effet immédiate : mise à part l'absence de *compacité* locale (lorsque l'espace est de dimension infinie), il n'y a pas de difficulté supplémentaire (on supposera l'espace *complet* chaque fois que nécessaire). On pourrait bien entendu considérer des cadres encore plus généraux, comme les espaces de Fréchet ou les variétés, mais cela aurait pour défaut de noyer l'essentiel derrière des considérations techniques, et n'apporterait pas grand chose en vue de la seconde partie. La notion même de *variété* (différentiable) ne sera abordée qu'au détour du théorème des multplicateurs de Lagrange dans le chapitre sur les *extrema* et à l'occasion de l'*analyse qualitative* des équations différentielles dans le tout dernier chapitre. On présentera donc les grands classiques du calcul différentiel (théorèmes des accroissements finis, d'inversion locale, des fonctions implicites, formules de Taylor) dans les $\mathbb{R}$-espaces vectoriels normés, le plus souvent supposés complets et alors appelés *espaces de Banach*.

Le chapitre consacré aux problèmes d'extremum peut être vu comme une introduction à l'optimisation continue, à l'analyse convexe et au calcul des variations. Ce dernier est d'ailleurs l'une des motivations pour faire du calcul différentiel en dimension infinie. Ce sont de vastes domaines, dont on présentera seulement quelques bases permettant d'aborder la lecture d'ouvrages plus avancés. Ce chapi-

1. Les termes en italiques sont expliqués plus loin, on peut les retrouver grâce à l'index.

tre est de plus l'occasion de présenter une classe importante d'équations différentielles, à savoir les équations d'Euler-Lagrange.

Cette partie s'achève par un chapitre sur la théorie des *formes différentielles*, souvent absente des cursus d'enseignements universitaires, et pourtant cruciale non seulement en mathématiques dites « pures » mais aussi dans les applications des mathématiques (thermodynamique, électromagnétisme, dynamique des fluides, etc.). On définira les notions essentielles que sont le *produit extérieur* et la *différentielle extérieure* de q-formes différentielles sur $\mathbb{R}^n$, et l'on présentera dans ce cadre les théorèmes de Poincaré, Frobenius et Stokes. Ce chapitre, bien qu'en apparence difficile, nécessite assez peu de pré-requis et peut servir de *vade mecum* sur le sujet.

Hormis son dernier chapitre donc, la première partie constitue le bagage que l'on peut attendre en calcul différentiel d'un étudiant en fin de licence. Les outils ainsi introduits permettent d'aborder sereinement la seconde partie.

Elle concerne les équations différentielles dites « ordinaires » (EDO), elles aussi considérées dans des $\mathbb{R}$-espaces de Banach en général. Ce choix est motivé par l'étude de modèles mathématiques pouvant être vus comme des EDO en dimension infinie : par exemple les équations différentielles sur réseaux, issues ou non de la discrétisation en espace d'équations aux dérivées partielles (EDP) d'évolution, ou encore les équations différentielles dans des espaces fonctionnels comme $L^2(\mathbb{R})$ (certaines EDP d'évolution pouvant être vues comme telles) ; c'est expliqué plus détail dans le chapitre intitulé *Équations modèles et outils de base*.

Le cadre est celui des équations différentielles « non pathologiques », au sens où elles sont supposées résolues (en la dérivée d'ordre le plus élevé) et sans problème de régularité (on choisit de ne pas s'aventurer sur le terrain des solutions généralisées, pour des équations dont les données seraient peu régulières). C'est une partie comportant bien sûr des *outils*, sous forme de lemmes, formules, théorèmes, etc. (comme le lemme de Gronwall, la formule de Duhamel, le théorème de Cauchy-Lipschitz pour ne citer que les outils de base), mais elle est aussi l'occasion d'insister sur diverses *méthodes*, et notamment celles de Picard, Lyapunov-Schmidt et Melnikov. Sans négliger les aspects « élémentaires », comme la résolution explicite dans les cas les plus simples et la classification des points fixes dans le plan, elle va (bien) au-delà du théorème d'existence et d'unicité de Cauchy-Lipschitz. Ceci commence par la question de la dépendance des solutions par rapport aux « conditions initiales » (avec le théorème du flot) et aux paramètres, et se poursuit par un approfondissement de la théorie pour les équations linéaires d'une part, et pour les équations non-linéaires autonomes d'autre part. Pour les premières, cela comprend la notion de résolvante, la théorie de Floquet, des éléments d'analyse spectrale, les notions de projecteurs spectraux et de dichotomies exponentielles. Pour les secondes, il s'agit essentiellement de l'étude de l'existence et des propriétés qualitatives (comportement asymptotique, stabilité par rapport aux paramètres) de solutions particulières (stationnaires, périodiques, orbites homo/hétéroclines), sans chercher à les calculer explicitement, avec notamment la théorie de Lyapunov et les théorèmes de Poincaré-Bendixson, de la variété stable et de bifurcation de Hopf. Cette partie *équations différentielles* peut faire l'objet d'un solide cours de première année de master.

Même si l'ouvrage se veut « auto-contenu » sur tout ce qui touche au calcul différentiel et aux équations différentielles, il s'appuie nécessairement sur d'autres domaines des mathématiques. Dans la mesure du possible, les notions indispensables de topologie, calcul intégral, analyse fonctionnelle, analyse complexe, algèbre linéaire ou géométrie, sont rappelées au fil du texte et accompagnées de références. Les notions les plus couramment utilisées sont rappelées ci-après.

Définitions, notations et résultats utiles

• On suppose connue la notion d'espace vectoriel. Le corps de base des espaces vectoriels considérés sera $\mathbb{R}$ (ou éventuellement $\mathbb{C}$). Dans un espace vectoriel normé E, on notera en général $\|\cdot\|_E$ la *norme*, ou simplement $\|\cdot\|$ s'il n'y a pas d'ambiguïté possible. Une norme est caractérisée par les trois propriétés suivantes :
1) quels que soient $\lambda \in \mathbb{R}$ et $x \in E$, $\|\lambda x\| = |\lambda|\,\|x\|$;
2) le seul vecteur x tel que $\|x\|_E = 0$ est $x = 0_E$;
3) et l'on a l'*inégalité triangulaire* :

$$\|x+y\| \leqslant \|x\| + \|y\|\,, \text{ quels que soient } x\,,\ y \in E\,.$$

Une partie V d'un espace vectoriel normé E est un *voisinage* de $x \in E$ s'il existe une boule ouverte de centre x et de rayon $R > 0$,

$$B(x;\,R) = \{y \in E\,;\ \|y - x| < R\}\,,$$

incluse dans V. Un *ouvert* est une partie de E qui est un voisinage de tous ses points. Un *fermé* est un sous-ensemble de E dont le complémentaire est ouvert. Un *compact* est un sous-ensemble de E dans lequel toute suite admet une sous-suite convergente (une *sous-suite* d'une suite $(x_n)_{n\in\mathbb{N}}$ étant une suite de la forme $(x_{\varphi(n)})_{n\in\mathbb{N}}$ avec $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante). L'image d'un compact par une application *continue* est compact. Les compacts d'un espace vectoriel normé de dimension finie sont les fermés bornés (théorème de Bolzano-Weierstrass).

• Un $\mathbb{R}$-*espace de Banach* est un $\mathbb{R}$-espace vectoriel normé *complet*, c'est-à-dire où toutes les *suites de Cauchy* sont convergentes (une suite $(x_n)_{n\in\mathbb{N}}$ étant dite de Cauchy si pour tout $\varepsilon > 0$ il existe $N \in \mathbb{N}$ tel que pour $n, p \geqslant N$, $\|x_n - x_p\| \leqslant \varepsilon$).

• Si E et F sont des espaces de Banach, l'espace vectoriel des applications *linéaires continues* de E dans F, muni de la norme

$$|||\ell||| = \sup_{x\in E\setminus\{0\}} \frac{\|\ell(x)\|_F}{\|x\|_E}$$

est un espace de Banach. Il sera noté $\mathcal{L}(E;\,F)$, ou simplement $\mathcal{L}(E)$ dans le cas $E = F$. En outre, le sous-ensemble des *isomorphismes* de E sur F :

$$\mathrm{Isom(E;\,F)} := \{\,u \in \mathcal{L}(E;\,F)\,;\ \exists v \in \mathcal{L}(F;\,E)\,,\ v \circ u = \mathrm{Id}_E \text{ et } u \circ v = \mathrm{Id}_F\,\}$$

est un ouvert de $\mathcal{L}(E;\,F)$. En vertu du théorème d'analyse fonctionnelle suivant, Isom(E; F) coïncide avec l'ensemble des isomorphismes au sens algébrique.

Théorème 0.1 ***Théorème de Banach***

Si E et F sont des espaces de Banach, la réciproque d'une application *linéaire* continue et bijective de E sur F, est continue.

(Voir [2, Cor. II.6 p. 19].) Autrement dit, pour vérifier qu'une application u est un isomorphisme de E sur F, il « suffit » de vérifier que u est *linéaire continue et bijective*. En dimension finie, toutes les applications linéaires sont continues : ce théorème n'a donc d'intérêt qu'en dimension infinie.

• Si $E_1,\ldots,\ E_n$ sont des espaces de Banach, le produit cartésien $E = E_1 \times \cdots \times E_n$, muni de

$$\|(x_1,\ldots,x_n)\|_E = \|x_1\|_{E_1} + \cdots + \|x_n\|_{E_n}\,,$$

est aussi un espace de Banach. Une application $\phi : E \to F$ est n-*linéaire* (et lorsqu'on ne veut pas préciser n on dit *multi-linéaire*) si pour tout $j \in \{1,\ldots,n\}$ et pour tout $(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n) \in E_1 \times \cdots E_{j-1} \times E_{j+1} \cdots \times E_n$ (avec les conventions naturelles lorsque $j = 1$ ou $j = n$), l'*application partielle*

$$x \in E_j \mapsto \phi(x_1,\ldots,x_{j-1},x,x_{j+1},\ldots,x_n)$$

est *linéaire*. Si de plus ϕ est continue, toutes les applications partielles sont continues et il existe $C \in \mathbb{R}^{+*}$ tel que

$$\|\phi(h_1,\ldots,h_n)\|_F \leqslant C \prod_{j=1}^{n} \|h_j\|_{E_j}$$

quels que soient les vecteurs $h_j \in E_j$. L'ensemble des applications n-linéaires continues forme un espace vectoriel que l'on notera $\mathcal{L}(E_1,\ldots,E_n;\,F)$ (à ne pas confondre avec l'espace $\mathcal{L}(E_1 \times \cdots \times E_n;\,F)$ des applications linéaires sur E). Muni de la norme définie par

$$\|\phi\|_{\mathcal{L}(E_1,\ldots,E_n;F)} = \max\{\,\|\phi(h_1,\ldots,h_n)\|_F\,;\ \|h_j\|_{E_j} \leqslant 1\,\}\,,$$

l'espace $\mathcal{L}(E_1,\cdots,E_n;\,F)$ est complet.

Lorsque $E_1 = \cdots = E_n = G$, on notera simplement $\mathcal{L}_n(G;\,F)$ l'espace des applications n-linéaires continues sur G^n, et $\mathcal{L}_n^s(G;\,F)$ le sous-espace des applications n-linéaires *symétriques*, c'est-à-dire invariantes par permutation des vecteurs $x_1,\ldots,x_n$ composant les vecteurs de G^n.

• Dans les espaces vectoriels normés de dimension finie, toutes les normes sont équivalentes, et tous ces espaces sont complets (indépendamment de la norme choisie). C'est le cas en particulier de $\mathbb{R}^n$, pour lequel des normes classiques sont

$$\|x\|_k = \left(\sum_{j=1}^{n} |x_j|^k\right)^{1/k} \quad \text{si } k < \infty \quad \text{et } \|x\|_\infty = \sup_{1\leqslant j\leqslant n} |x_j|\,,$$

où $x_1,\ldots,x_n$ désignent les composantes de $x \in \mathbb{R}^n$. De même, l'espace des matrices $\mathcal{M}_{p,q}(\mathbb{R})$ à p lignes et q colonnes, est complet quelle que soit la norme choisie. On peut notamment considérer les normes subordonnées aux normes $\|\cdot\|_k$ dans $\mathbb{R}^q$ et $\mathbb{R}^p$, définies par

$$|||M|||_k \;=\; \sup_{x\in\mathbb{R}^q, x\neq 0} \frac{\|Mx\|_k}{\|x\|_k}\,.$$

Si M a pour coefficients $m_{i,j}$ on a (voir par exemple [34, p. 43])

$$|||M|||_1 \;=\; \max_j \sum_i |m_{i,j}|\,,$$
$$|||M|||_\infty \;=\; \max_i \sum_j |m_{i,j}|\,,$$
$$|||M|||_2 \;=\; \sqrt{\varrho(M^{\mathrm{t}}M)}\,,$$

où $\varrho(M^{\mathrm{t}}M)$ désigne le *rayon spectral*, c'est-à-dire la plus grande valeur propre (en valeur absolue) de la matrice symétrique positive $M^{\mathrm{t}}M$, la matrice M^{t} étant par définition la *transposée* de M, de coefficients $(m_{j,i})_{j\leqslant q, i\leqslant p}$.

• Si $E_1,\ldots,E_n$ sont des espaces de Banach, le produit cartésien $E = E_1\times\cdots\times E_n$ peut être muni de l'une quelconque des normes, équivalentes entre elles,

$$\|(x_1,\ldots,x_n)\|_{k,E} \;=\; \|(\|x_1\|_{E_1},\ldots,\|x_n\|_{E_n})\|_k\,.$$

• Des exemples classiques d'espaces de Banach de dimension infinie sont les espaces de suites

$$\ell^p(\mathbb{N}) \;:=\; \{u = (u_n)_{n\in\mathbb{N}}\,;\; \sum_n |u_n|^p < +\infty\}\,,$$

munis des normes définies par

$$\|u\|_p \;=\; \left(\sum_n |u_n|^p\right)^{1/p}\,,$$

pour $p < +\infty$, et

$$\ell^\infty(\mathbb{N}) \;:=\; \{u = (u_n)_{n\in\mathbb{N}}\,;\; \|u\|_\infty := \sup_n |u_n| < +\infty\}\,,$$

ainsi que divers espaces de fonctions, comme l'espace $\mathcal{C}(K;E)$ des fonctions continues sur un compact K et à valeurs dans un espace de Banach E, muni de la « norme sup », définie par

$$\|f\| = \max_{x\in K} \|f(x)\|_E\,.$$

Moyennant l'identification des fonctions coïncidant presque partout, les *espaces de Lebesgue* (voir par exemple [32, chap. 3])

$$L^p(\Omega) = \left\{ f : \Omega \to \mathbb{R} \text{ mesurable} \,;\ \int_\Omega |f(x)|^p \,\mathrm{d}x < +\infty \right\} ,$$

$$L^\infty(\Omega) = \{ f : \Omega \to \mathbb{R} \text{ mesurable essentiellement bornée} \} ,$$

où Ω est un ouvert de $\mathbb{R}^n$, munis des normes

$$\|f\|_p := \left(\int_\Omega |f(x)|^p \,\mathrm{d}x \right)^{1/p} , \quad \|f\|_\infty := \sup_{x \in \Omega} \mathrm{ess} |f(x)| ,$$

sont également des espaces de Banach de dimension infinie. Plus généralement, il est possible definir des espaces $L^p(\Omega; E)$ pour des fonctions à valeurs dans un espace de Banach E (par exemple un autre espace L^p) : cependant cela soulève des questions délicates de mesurabilité (on renvoie à [11] pour de plus amples détails).

• Les *espaces de Hilbert* sur $\mathbb{R}$ sont des cas particuliers de $\mathbb{R}$-espaces de Banach, dans lesquels la norme est définie par un *produit scalaire*, forme bilinéaire symétrique définie positive $\langle \cdot , \cdot \rangle$:

$$\|u\| = \langle u , u \rangle^{1/2} .$$

Les espaces de Hilbert sur $\mathbb{C}$, où la norme est associée à un *produit hermitien*, forme sesquilinéaire hermitienne définie positive, sont quant à eux des $\mathbb{C}$-espaces de Banach. Une propriété fondamentale des espaces de Hilbert est l'*inégalité de Cauchy-Schwarz* : quels que soient u et v dans un espace de Hilbert,

$$|\langle u , v \rangle| \leqslant \|u\| \, \|v\| .$$

Partie I

Calcul différentiel

DIFFÉRENTIABILITÉ 1

Ainsi il ne faut point s'étonner, si notre nouveau calcul des différences et des sommes, qui enveloppe la considération de l'infini et s'éloigne par conséquent de ce que l'imagination peut atteindre, n'est pas venu d'abord à sa perfection. […] Il faut rendre justice à M. Newton (à qui la Géométrie, l'Optique, et l'Astronomie ont de grandes obligations) qu'encore en ceci il a eu quelque chose de semblable de son chef, suivant ce qu'on en a sçu depuis.

Gottfried Wilhelm Leibniz,
Considérations sur la différence qu'il y a entre l'analyse ordinaire et le nouveau calcul des transcendantes,
Journal des Sçavans de l'an 1694.

1.1 NOTIONS DE BASE

Ce chapitre est consacré à la notion de différentiabilité et aux théorèmes fondamentaux qui lui sont attachés. Avant de parler de différentiabilité, il n'est peut-être pas superflu de rappeler ce qu'est la *dérivabilité* d'une fonction d'une variable réelle. Prenons le cas le plus simple d'une fonction *scalaire* $g : I \subset \mathbb{R} \to \mathbb{R}$ sur un intervalle ouvert I. On dit qu'elle est *dérivable* en un point a de I si la fonction *taux d'accroissement*

$$t \in I\backslash\{a\} \mapsto \frac{g(t) - g(a)}{t - a}$$

a une limite $\ell \in \mathbb{R}$ au point a. Si c'est le cas, on définit la dérivée de g au point a comme le nombre réel ℓ. Ce nombre est en général noté $g'(a)$, de sorte que, par définition,

$$g'(a) = \lim_{\substack{\neq \\ t \to a}} \frac{g(t) - g(a)}{t - a}\,.$$

Il s'interprète graphiquement comme le *coefficient directeur* de la *tangente* à la courbe de g au point a, tandis que chaque taux d'accroissement correspond au *coef-*

ficient directeur d'une *corde* entre le point de coordonnées $(a, g(a))$ et un autre point de la courbe.

La notion de dérivée s'étend sans problème aux fonctions à valeurs vectorielles, pourvu que l'on dispose d'une notion de limite dans l'espace d'arrivée. Ainsi, une fonction $\mathbf{g} : I \subset \mathbb{R} \to \mathbb{R}^p$ est dérivable en $a \in I$ si la fonction vectorielle

$$t \in I \setminus \{a\} \mapsto \frac{1}{t-a}(\mathbf{g}(t) - \mathbf{g}(a))$$

a une limite, qui est alors forcément un élément de $\mathbb{R}^p$ ($p \geqslant 2$). Si c'est le cas, on définit

$$\mathbf{g}'(a) = \lim_{\substack{\neq \\ t \to a}} \frac{1}{t-a}(\mathbf{g}(t) - \mathbf{g}(a)) \,. \tag{1.1}$$

Si la fonction $\mathbf{g}$ a pour *composantes* g_1,..., g_p, le vecteur $\mathbf{g}'(a)$ a pour composantes $g_1'(a)$,..., $g_p'(a)$.

Plus généralement, si E est un $\mathbb{R}$-espace vectoriel normé quelconque, une fonction $\mathbf{g} : I \subset \mathbb{R} \to E$ est dérivable en $a \in I$ si la fonction

$$\begin{array}{lcl} I \setminus \{a\} & \to & E \\ t & \mapsto & \frac{1}{t-a}(\mathbf{g}(t) - \mathbf{g}(a)) \end{array}$$

a une limite (dans E). Si c'est le cas on définit encore $\mathbf{g}'(a) \in E$ par la formule (1.1) ci-dessus.

Lorsqu'une fonction $\mathbf{g} : I \subset \mathbb{R} \to E$ est dérivable en tout point de I, on lui associe une autre fonction, qu'on appelle sa *dérivée* et qui est définie par

$$\begin{array}{rcl} \mathbf{g}' : I & \to & E \\ t & \mapsto & \mathbf{g}'(t) = \lim\limits_{\substack{\neq \\ k \to 0}} \frac{1}{k}(\mathbf{g}(t+k) - \mathbf{g}(t)) \,. \end{array} \tag{1.2}$$

(Attention, la lettre t ci-dessus joue le rôle qu'avait a dans (1.1). Noter également que t est une « variable muette » dans (1.1), tout comme k ci-dessus.)

Il n'y a pas beaucoup plus à dire sur les dérivées de fonctions d'une variable à valeurs vectorielles. En revanche, pour les fonctions de plusieurs variables, qu'elles soient scalaires ou vectorielles, l'extension de la notion de dérivée est beaucoup plus délicate. Rappelons en effet qu'il est formellement interdit de diviser par un vecteur, et qu'il n'y a donc pas de notion naturelle de taux d'accroissement pour une fonction de plusieurs variables. C'est pourquoi d'ailleurs il est préférable de voir la définition de $\mathbf{g}'(t)$ dans (1.2) comme la relation d'*approximation*

$$\mathbf{g}(t+k) \,-\, \mathbf{g}(t) \underset{k \to 0}{=} k\,\mathbf{g}'(t) \,+\, o(k) \,, \tag{1.3}$$

où o est la notation de *Landau* indiquant qu'une fonction (en l'occurrence ici la fonction $k \mapsto \mathbf{g}(t+k) - \mathbf{g}(t) \,-\, k\,\mathbf{g}'(t)$, à t *fixé*) est *négligeable* devant une autre (ici l'identité, $k \mapsto k$). La notion de différentielle fournira un substitut au terme principal $k\,\mathbf{g}'(t)$ dans cette relation.

1.1.1 Dérivées partielles

Une façon élémentaire de faire du calcul différentiel avec une fonction de plusieurs variables consiste à considérer les variables séparément. Par exemple, pour une fonction de deux variables $f : (x,t) \mapsto f(x,t)$, on peut fixer x et considérer $t \mapsto f(x,t)$. Si cette dernière est dérivable, on dira que f admet une *dérivée partielle* par rapport à t,

$$\frac{\partial f}{\partial t}(x,t) = \lim_{\substack{\neq \\ k\to 0}} \frac{f(x,t+k) - f(x,t)}{k}\,.$$

Cette notation de la dérivée partielle recèle des pièges. D'abord le trait de fraction ne signifie pas que l'on puisse faire une quelconque simplification entre ce qui semble être un numérateur (∂f) et ce qui semble être un dénominateur (∂t). Cela n'aurait pas plus de sens que dans l'expression $\frac{\mathrm{d}g}{\mathrm{d}t}$, *notation de Leibniz* pour la dérivée d'une fonction d'une variable $g : t \mapsto g(t)$. En outre, la lettre t joue deux rôles différents : elle est « muette » dans ∂t, qui signifie juste ici que l'on dérive f par rapport à sa « deuxième » variable, tandis qu'elle sert à préciser le point où on évalue cette dérivée partielle dans la parenthèse (x,t). Ainsi, on pourrait très bien choisir un autre point, (x,s) par exemple, et calculer

$$\frac{\partial f}{\partial t}(x,s) = \lim_{\substack{\neq \\ k\to 0}} \frac{f(x,s+k) - f(x,s)}{k}\,.$$

Bien entendu, on peut aussi fixer t et considérer

$$\frac{\partial f}{\partial x}(x,t) = \lim_{\substack{\neq \\ h\to 0}} \frac{f(x+h,t) - f(x,t)}{h}$$

pourvu que cette limite existe. Là encore, on peut très bien changer de point et calculer

$$\frac{\partial f}{\partial x}(y,t) = \lim_{\substack{\neq \\ h\to 0}} \frac{f(y+h,t) - f(y,t)}{h}$$

si elle existe. Les notations $\frac{\partial f}{\partial t}$ et $\frac{\partial f}{\partial x}$ font référence aux notations « génériques » pour les variables t et x dont dépend f : il est formellement déconseillé d'y toucher en cours de calcul !

Plus généralement, soit une fonction

$$\begin{array}{rcl} \mathbf{f} : U & \to & \mathbb{R}^p \\ \mathbf{x} = (x_1,\ldots,x_d) & \mapsto & \mathbf{f}(x) = \mathbf{f}(x_1,\ldots,x_d)\,, \end{array}$$

où U est un ouvert de $\mathbb{R}^d$, $d \geqslant 2$. Pour tout $\mathbf{x} = (x_1, \ldots, x_d) \in U$, l'ensemble

$$V_i(\mathbf{x}) := \{t \in \mathbb{R}\,;\ (x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_d) \in U\}$$

est un voisinage ouvert de x_i dans $\mathbb{R}$. Considérons alors l'*application partielle*

$$\begin{array}{rccl} \mathbf{g}_i : & V_i(\mathbf{x}) & \to & \mathbb{R}^p \\ & t & \mapsto & \mathbf{f}(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_d)\,. \end{array}$$

Si elle est dérivable au point x_i, on dit que $\mathbf{f}$ admet une *dérivée partielle* par rapport à x_i au point $\mathbf{x}$ et on la note

$$\frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x})\,, \quad \text{ou } \partial_{x_i}\mathbf{f}(\mathbf{x})\,, \quad \text{voire simplement } \partial_i \mathbf{f}(\mathbf{x})\,.$$

(Noter que seule cette dernière notation permet d'éviter les ambigüités signalées ci-dessus, tant qu'on ne change pas l'ordre des variables bien sûr.) Par définition,

$$\frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x}) = \lim_{\substack{\neq \\ h \to 0}} \frac{\mathbf{f}(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_d) - \mathbf{f}(x_1, \ldots, x_d)}{h}$$

pourvu que cette limite existe.

Matrice jacobienne

Si une fonction $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$, de composantes $(f_1, \ldots, f_p)$, admet une dérivée partielle par rapport à chacune de ses variables au point x, on définit sa *matrice jacobienne* au point x par

$$\mathrm{D}f(x) = \begin{pmatrix} \partial_1 f_1(x) & \cdots & \partial_d f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_p(x) & \cdots & \partial_d f_p(x) \end{pmatrix} \in \mathcal{M}_{p,d}(\mathbb{R})\,.$$

Autrement dit, le coefficient de la matrice jacobienne de f d'indice $i \in \{1, \cdots, p\}$ en ligne et $j \in \{1, \cdots, d\}$ en colonne est

$$(\mathrm{D}f(x))_{i,j} = \partial_j f_i(x)\,.$$

En particulier pour une fonction scalaire, c'est-à-dire si $p = 1$, $\mathrm{D}f(x)$ est une matrice ligne.

Opérateur « nabla » On note $\nabla f(x)$ (qui se lit « nabla f de x ») la matrice transposée de $\mathrm{D}f(x)$, de sorte que :

$$(\nabla f(x))_{i,j} = \partial_i f_j(x)\,.$$

Cette notation est souvent utilisée lorsque $p = 1$, auquel cas $\nabla f(x)$ est une matrice colonne, que l'on identifie à un vecteur de $\mathbb{R}^d$ appelé *gradient* de f au point x.

Opérateurs différentiels classiques

Comme on vient de le voir, pour une fonction à *valeurs scalaires* $\varphi : U \subset \mathbb{R}^d \to \mathbb{R}$ admettant des dérivées partielles en tout point de U on définit son *gradient* par :

$$\begin{array}{rcl} \operatorname{grad}\varphi \;:\; U \subset \mathbb{R}^d & \to & \mathbb{R}^d \\ x & \mapsto & (\operatorname{grad}\varphi)(x) := \big(\partial_1\varphi(x), \ldots, \partial_d\varphi(x)\big)^{\mathrm{t}}. \end{array}$$

On écrit indifféremment $\operatorname{grad}\varphi$ ou $\nabla\varphi$.

Par ailleurs, pour une fonction $f : U \subset \mathbb{R}^d \to \mathbb{R}^d$ (noter l'égalité des dimensions au départ et à l'arrivée), de composantes $(f_1, \ldots, f_d)$, admettant des dérivées partielles, on définit la *divergence* par

$$\begin{array}{rcl} \operatorname{div} f \;:\; U \subset \mathbb{R}^d & \to & \mathbb{R} \\ x & \mapsto & (\operatorname{div} f)(x) := \operatorname{tr}(\mathrm{D}f(x)) = \displaystyle\sum_{i=1}^{d} \partial_i f_i(x). \end{array}$$

Il est parfois commode d'écrire $\operatorname{div} f = \nabla \cdot f$, où ∇ est l'opérateur nabla défini ci-dessus, dont les « composantes » sont les dérivées partielles ∂_i, la notation $\cdot$ se rapportant ici au produit scalaire canonique sur $\mathbb{R}^d$, défini par

$$x \cdot y = \sum_{i=1}^{d} x_i y_i .$$

Lorsque $d = 3$, on définit aussi le *rotationnel* de f

$$\begin{array}{rcl} \operatorname{rot} f \;:\; U \subset \mathbb{R}^3 & \to & \mathbb{R}^3 \\ x & \mapsto & (\operatorname{rot} f)(x) \end{array}$$

par

$$(\operatorname{rot} f)(x) := \big(\partial_2 f_3(x) - \partial_3 f_2(x), \partial_3 f_1(x) - \partial_1 f_3(x), \partial_1 f_2(x) - \partial_2 f_1(x)\big)^{\mathrm{t}}.$$

Si $\times$ désigne le *produit vectoriel* dans $\mathbb{R}^3$, il est parfois commode de voir le rotationnel comme $\operatorname{rot} f = \nabla \times f$. Il faut toutefois garder à l'esprit que cette notation est formelle, et qu'elle peut conduire à des erreurs, notamment lorsqu'on cherche à calculer $\operatorname{rot}(f \times g)$ (voir l'exercice 4.2).

On peut aussi définir le rotationnel d'une fonction de deux variables. C'est alors une fonction scalaire, parfois appelée *tourbillon* : si $f : U \subset \mathbb{R}^2 \to \mathbb{R}^2$ admet des dérivées partielles, on note

$$\begin{array}{rcl} \operatorname{rot} f \;:\; U \subset \mathbb{R}^2 & \to & \mathbb{R}^2 \\ x & \mapsto & (\operatorname{rot} f)(x) := \partial_1 f_2(x) - \partial_2 f_1(x). \end{array}$$

Ce n'est rien d'autre que la troisième composante du rotationnel de f vue comme une fonction de (x, y, z) ne dépendant en fait pas de z.

Ce sont en tous cas des définitions purement techniques, dont il est difficile à première vue de comprendre la signification. On renvoie aux exercices 1.2, 1.3, 1.4, 1.5, 1.6 et 1.8 pour une mise en perspective de ces opérateurs différentiels.

1.1.2 Dérivées directionnelles

Les dérivées partielles sont en fait des cas particuliers de dérivées directionnelles, ou dérivées suivant un vecteur. Si E et F sont deux $\mathbb{R}$-espaces vectoriels normés et U est un ouvert de E, on dit qu'une fonction $f : U \to F$ admet une *dérivée suivant un vecteur* $h \in E$ en $x \in U$ si la fonction d'une variable $g : t \mapsto g(t) := f(x + th)$ est dérivable en zéro. (Noter qu'il existe un voisinage I de zéro dans $\mathbb{R}$ tel que $x + th \in U$ pour tout $t \in I$.) Si c'est le cas, on notera

$$D_h f(x) = \lim_{\substack{\neq \\ t \to 0}} \frac{f(x + th) - f(x)}{t}$$

cette dérivée (attention, c'est une notation loin d'être universelle). Si $E = \mathbb{R}^d$ et $h = \mathrm{e}_i$, le i-ème vecteur de la *base canonique* de $\mathbb{R}^d$, dont toutes les composantes sont nulles sauf la i-ème qui vaut 1, on constate que $f(x + th) = f(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_d)$, d'où

$$D_{\mathrm{e}_i} f(x) = \frac{\partial f}{\partial x_i}(x) \, .$$

Une dérivée partielle est donc une dérivée suivant un vecteur de base.

Les dérivées directionnelles jouent un rôle très important du point de vue pratique, car elles constituent un pont entre les dérivées de fonctions d'une variable et les différentielles de fonctions de plusieurs variables.

1.1.3 Fonctions différentiables, fonctions de classe $\mathcal{C}^1$

Nous arrivons maintenant à la notion centrale du chapitre. On suppose que U est un ouvert non vide d'un $\mathbb{R}$-espace vectoriel normé E, et l'on considère une fonction

$$f : U \to F \, ,$$

où F est un autre espace vectoriel normé. Le fait que U soit ouvert permettra de considérer les valeurs de f au voisinage de tout point[1] de U. Si cela n'était pas le

1. Même si les éléments de E sont des vecteurs, on parlera souvent de points de E, par abus de langage.

cas, il faudrait restreindre la définition qui suit aux points *intérieurs*[2] à U. On note $\mathcal{L}(E;\,F)$ désigne l'espace des applications *linéaires continues* de E dans F.

Définition 1.1

La fonction f est *différentiable* en un point $x \in U$ s'il existe $\ell \in \mathcal{L}(E;\,F)$ tel que

$$\lim_{h \overset{\neq}{\to} 0_E} \frac{\|f(x+h) - f(x) - \ell(h)\|_F}{\|h\|_E} = 0\,. \tag{1.4}$$

Autrement dit, f est différentiable en x s'il existe une application linéaire continue $\ell \in \mathcal{L}(E;\,F)$, $R > 0$ et une application $\tau : B(0_E;\,R) \to \mathbb{R}^+$ tendant vers 0 en 0_E telles que pour tout $h \in B(0_E;\,R)$,

$$\|f(x+h) - f(x) - \ell(h)\|_F = \tau(h)\,\|h\|_E\,.$$

Il suffit en effet de poser $\tau(0_E) = 0$ et, si $R > 0$ est choisi de sorte que la boule $B(0_E;\,R)$ (de centre 0_E de rayon R) soit incluse dans U, $\tau(h) := \|f(x+h) - f(x) - \ell(h)\|_F / \|h\|_E$ pour $0 < \|h\|_E < R$. Cela signifie que $\ell(h)$ « approche » la différence $f(x+h) - f(x)$ lorsque h tend vers 0_E, au sens où

$$f(x+h) \underset{h \to 0_E}{=} f(x) + \ell(h) + o(h)\,,$$

en utilisant à nouveau la notation de Landau o pour « négligeable ».

Proposition 1.1

Une fonction $f : U \to F$ différentiable en un point $x \in U$ (au sens de la définition 1.1) est nécessairement continue au point x.

DÉMONSTRATION. En introduisant la fonction τ comme ci-dessus, on a par l'inégalité triangulaire

$$\|f(x+h) - f(x)\|_F \leqslant \tau(h)\,\|h\|_E + \|\ell(h)\|_F\,,$$

où le membre de droite tend vers 0 lorsque h tend vers 0_E grâce à (1.4) et à la continuité de ℓ en 0_E. □

Remarque 1.1

Par un argument analogue, on obtient la continuité de ℓ si l'on a celle de f. En effet, si f est continue en x et s'il existe ℓ (*a priori* seulement) linéaire telle que l'on ait (1.4), alors cette application est nécessairement continue, car par l'inégalité triangulaire

2. C'est-à-dire les $x \in U$ pour lesquels il existe $R > 0$ tel que $B(x;\,R) \subset U$.

$$\|\ell(h)\|_F \leqslant \tau(h)\,\|h\|_E + \|f(x+h) - f(x)\|$$

tend vers 0 lorsque h tend vers 0_E. (On rappelle que la continuité d'une application *linéaire* équivaut à sa continuité en 0.) On montre de plus que ℓ est unique.

Notation. Bien sûr, l'application ℓ dans la définition 1.1 dépend du point x. Il faut donc adopter une notation faisant apparaître x. Afin de bien voir la différentielle d'une fonction (dans la définition ci-après) comme une nouvelle fonction, nous choisissons de noter $\ell = \mathrm{d}f(x)$. La difficulté tient ici au fait que $\mathrm{d}f(x)$ est elle-même une application. Cependant, comme elle est linéaire, nous noterons simplement (pour éviter la juxtaposition de parenthèses) $\mathrm{d}f(x)\cdot h$ la valeur prise par cette application au vecteur $h \in E$. Ainsi, (1.4) se réécrit

$$\lim_{\substack{\neq \\ h\to 0_E}} \frac{\|f(x+h) - f(x) - \mathrm{d}f(x)\cdot h\|_F}{\|h\|_E} = 0\,,$$

que l'on peut aussi voir comme la généralisation de (1.3),

$$f(x+h) - f(x) \underset{h\to 0_E}{=} \mathrm{d}f(x)\cdot h + o(h)\,. \tag{1.5}$$

L'application linéaire $\mathrm{d}f(x)$ est appelée *différentielle* de f au point x.

Proposition 1.2

Une fonction $f : U \to F$ différentiable en un point $x \in U$ admet une dérivée suivant n'importe quel vecteur $h \in E$ en x et

$$D_h f(x) = \mathrm{d}f(x)\cdot h\,.$$

DÉMONSTRATION. D'après (1.5), la différentiabilité de f en x implique, à $h \in E$ fixé,

$$f(x+th) - f(x) \underset{t\to 0}{=} \mathrm{d}f(x)\cdot(th) + o(th)\,.$$

Or $\mathrm{d}f(x)\cdot(th) = t\,\mathrm{d}f(x)\cdot h$ par linéarité de $\mathrm{d}f(x)$. On en déduit que la fonction $t \neq 0 \mapsto (f(x+th) - f(x))/t$ a pour limite $\mathrm{d}f(x)\cdot h$ quand t tend vers zéro. □

Cet énoncé fournit une interprétation très importante de la différentielle. D'un point de vue pratique, pour calculer $\mathrm{d}f(x)\cdot h$ il « suffit » de dériver la fonction d'une variable $t \mapsto f(x+th)$ en 0. D'un point de vue plus théorique, une condition nécessaire pour qu'une fonction soit différentiable est qu'elle soit dérivable dans toutes les directions. Cependant cette condition n'est *pas suffisante*, elle correspond « seulement » à ce que l'on appelle la *Gâteaux-différentiabilité*. Si l'on considère par exemple la fonction

$$\begin{array}{rcl} f : \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & f(x,y) := \begin{cases} x^3/(x^2+y^2) & \text{si } (x,y) \neq (0,0)\,, \\ 0 & \text{si } (x,y) = (0,0)\,, \end{cases} \end{array}$$

elle est dérivable suivant tout vecteur $(h,k) \in \mathbb{R}^2$, avec en particulier

$$D_{(h,k)} f(0,0) = h^3/(h^2+k^2)$$

si $(h,k) \neq (0,0)$. Cependant, l'application $(h,k) \mapsto D_{(h,k)} f(0,0)$ n'étant pas linéaire, la fonction f ne peut pas être différentiable en $(0,0)$, précisément à cause de la proposition 1.2.

Définition 1.2

On dit qu'une fonction $f : U \to F$ est *différentiable* sur U si elle est différentiable en tout point $x \in U$. Dans ce cas, on appelle *différentielle* de f la fonction

$$\begin{array}{rccl} \mathrm{d}f : & U & \to & \mathcal{L}(E;F) \\ & x & \mapsto & \mathrm{d}f(x) . \end{array}$$

Si de plus $\mathrm{d}f$ est continue, on dit que f est *continûment différentiable*, ou de façon équivalente que f est de *classe* $\mathcal{C}^1$.

Remarque 1.2

Une erreur fréquente est de confondre $\mathrm{d}f$ et sa valeur en x, $\mathrm{d}f(x)$. En particulier, la continuité de $\mathrm{d}f$ n'a rien à voir avec la continuité de $\mathrm{d}f(x)$, cette dernière étant *par définition* linéaire continue. (Notons au passage que $\mathrm{d}f$ n'a aucune raison d'être linéaire en général.) Ainsi, une fonction différentiable n'est pas nécessairement de classe $\mathcal{C}^1$, pas plus qu'une fonction dérivable n'est nécessairement de classe $\mathcal{C}^1$: un contre-exemple classique est la fonction

$$\begin{array}{rcl} f : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & f(x) := \begin{cases} x^2 \sin(1/x) & \text{si } x \neq 0 , \\ 0 & \text{si } x = 0 , \end{cases} \end{array}$$

qui est dérivable y compris en 0 (de dérivée nulle) mais dont la dérivée n'a pas de limite en 0. Or on va voir ci-après que les notions de dérivabilité et de différentiabilité coïncident pour les fonctions d'une variable réelle.

1.1.4 Cas particuliers importants de fonctions différentiables

• Une fonction d'une variable réelle est *différentiable* si et seulement si elle est *dérivable* : la différentielle d'une fonction dérivable $g : U \subset \mathbb{R} \to F$ est donnée par

$$\mathrm{d}g(t) \cdot k = k\, g'(t) \quad \text{quels que soient } t \in U \text{ et } k \in \mathbb{R} .$$

(Notons que $k \in \mathbb{R}$ est un scalaire et $g'(t) \in F$ est un vecteur en général, c'est pourquoi on les a écrits dans cet ordre.) En effet, si g est dérivable en t alors