Calcul différentiel et équations différentielles

Thibaut Delcroix
bureau 308, Batiment 9
thibaut.delcroix@umontpellier.fr

HAX502X 2023-2024

Il s'agit du poly de cours pour l'UE HAX502X en 2023-2024, qui est au premier semestre de L3 Math à la Faculté des Sciences de Montpellier. L'UE comporte 27h de TD et 27h de CM. Les 24 sections du poly représentent des morceaux de cours d'une durée entre 1h et 1h30 en général, avec quelques compléments non fait en cours, ou, beaucoup plus rarement, des parties développées en cours et non détaillées dans le poly (essentiellement des exercices de TD traités en CM). La plupart des démonstrations des résultats sont données, à l'exception notable de l'existence dans le Théorème de Cauchy-Lipschitz. C'est parce que les étudiants n'ont pas fait d'analyse fonctionnelle et de topologie en dimension infinie à ce stade de leur parcours à Montpellier. La topologie des espaces vectoriels de dimension finie est supposée connue, mais quelques rappels et compléments sont inclus (continuité des applications linéaires, normes triples, Théorème du point fixe dans $\mathbb{R}^n$, connexité).

Attention : le poly n'a pas beaucoup été relu, il peut contenir beaucoup de typos, ainsi que des erreurs ou imprécisions perdues dans les raisonnements. Que le lecteur garde un regard critique !

Table des matières

1 Rappels de calcul différentiel en une variable

1.1 Définition

Soit A un sous-ensemble de $\mathbb{R}$, et $f : A \to \mathbb{R}$ une fonction (d'une variable réelle, à valeurs réelles).

Définition 1.1. La fonction f est *dérivable* en $a \in \mathbb{R}$ si :

1. A est un voisinage de a (c'est-à-dire $\exists \varepsilon > 0,]a - \varepsilon, a + \varepsilon[\subset A$),
2. $\frac{f(a+t)-f(a)}{t}$ a une limite dans $\mathbb{R}$ quand t tend vers 0 (le quotient est bien défini pour t vérifiant $0 \neq |t| < \varepsilon$).

Dans ce cas, on appelle cette limite la *dérivée* de f en a.

Si f est dérivable en a, la fonction

$$x \mapsto f(a) + (x - a)f'(a)$$

est la meilleure approximation affine infinitésimale de f en a. On l'appelle la *fonction tangente* à f en a. Son graphe $\{(x, f(a) + (x-a)f'(a))\}$ est une droite affine, la *tangente* en a au graphe de f, et $f'(a)$ donne la pente de cette droite.

1.2 Résultats fondamentaux en une variable

Puisque la fonction tangente en a est la meilleur approximation locale d'une fonction dérivable f en a, on peut s'attendre à ce qu'elle partage des propriétés de f. Par exemple, si f admet un maximum en a, la fonction tangente aussi. Comme la fonction tangente est affine, ce n'est possible que si sa pente est nulle. Ce résultat ce démontre formellement comme suit, et il est à la base des résultats fondamentaux de calcul différentiel en une variable, que nous rappelons ensuite.

Théorème 1.2. *Si f admet un maximum en a, et f est dérivable en a, alors $f'(a) = 0$.*

Démonstration. Soit $\varepsilon > 0$ tel que $]a-\varepsilon, a+\varepsilon[\subset A$. Comme f admet un maximum en a, on a

$$\frac{f(a+t) - f(a)}{t} \leq 0$$

pour $0 < t < \varepsilon$, et

$$0 \leq \frac{f(a+t) - f(a)}{t}$$

pour $-\varepsilon < t < 0$. Par passage à la limite quand t tend vers 0 dans chacune de ces inégalités, on obtient

$$0 \leq f'(a) \leq 0$$

donc $f'(a) = 0$. □

Corollaire 1.3. *Si f admet un minimum en a, et f est dérivable en a, alors $f'(a) = 0$.*

Démonstration. Adapter la preuve, ou appliquer le résultat précédent à $-f$. □

De ce résultat découle plusieurs théorèmes fondamentaux du calcul différentiel en une variable.

Théorème 1.4 (Théorème des accroissements finis)**.** *Supposons $f : [a,b] \to \mathbb{R}$ continue sur $[a,b]$, et dérivable en tout point de $]a,b[$ (on dit, dérivable sur $]a,b[$). Alors il existe $c \in]a,b[$ tel que*

$$\frac{f(b) - f(a)}{b-a} = f'(c)$$

Démonstration. On considère la fonction auxilliaire $g : [a,b] \to \mathbb{R}$ définie par

$$g(x) = f(x) - \frac{f(b) - f(a)}{b-a}(x-a)$$

Cette fonction est continue sur $[a,b]$, dérivable sur $]a,b[$. Elle vérifie de plus $g(b) = f(a) = g(a)$.

Il suffit de montrer qu'il existe $c \in]a,b[$ tel que $g'(c) = 0$. En effet, pour $x \in]a,b[$, on a

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b-a}$$

donc $g'(c) = 0$ si et seulement si $\frac{f(b)-f(a)}{b-a} = f'(c)$.

Comme g est continue sur $[a, b]$, elle admet un maximum et un minimum sur $[a, b]$. Si ce maximum est atteint en $c \in]a, b[$, on a $g'(c) = 0$, par le Théorème 1.2. Sinon, le maximum est atteint en a et en b. Considérons alors un point $d \in [a, b]$ où le minimum de g est atteint. Si $d = a$ ou $d = b$, alors la fonction g est constante, donc $g'(c) = 0$ pour tout $c \in]a, b[$. Sinon, on prends $c = d$, et on a bien $g'(c) = 0$ par le Corollaire 1.3. □

Corollaire 1.5 (Inégalité des accroissements finis). *Supposons f continue sur $[a, b]$, dérivable sur $]a, b[$, et supposons qu'il existe $M \in \mathbb{R}$ tel que pour tout $x \in]a, b[$, $|f'(x)| \leq M$. Alors f est M-Lipshitzienne sur $[a, b]$, c'est-à-dire pour tout x, $y \in [a, b]$,*

$$|f(y) - f(x)| \leq M|y - x|$$

Remarque 1.6. C'est en fait une équivalence : on vérifie facilement par la définition, que si une fonction M-Lipschitzienne est dérivable en a, alors $|f'(a)| \leq M$.

Corollaire 1.7. *Soit $f :]a, b[\to \mathbb{R}$ une fonction dérivable sur $]a, b[$. Alors*

1. *f est croissante si et seulement si $f' \geq 0$ sur $]a, b[$*
2. *f est décroissante si et seulement si $f' \leq 0$ sur $]a, b[$*
3. *f est constante si et seulement si $f' \equiv 0$ sur $]a, b[$*
4. *si $f' > 0$ sur $]a, b[$, alors f est strictement croissante sur $]a, b[$. En particulier, f est une bijection de $]a, b[$ sur $f(]a, b[)$.*

1.3 Dérivées supérieures

Définition 1.8. Si f est dérivable sur $]a, b[$, et f' est dérivable en $c \in]a, b[$, on note $f''(c) = (f')'(c)$ et on l'appelle la dérivée seconde de f en c. Par récurrence, on définit de la même manière les dérivées supérieures. On note $f^{(n)}$ la dérivée n-ième de f, avec $f^{(0)} = f$, $f^{(1)} = f'$ et $f^{(2)} = f''$.

L'interprétation est donnée par la formule de Taylor-Young : les dérivées supérieures fournissent la meilleure approximation polynomiale infinitésimale de f.

Théorème 1.9 (Formule de Taylor-Young). *Supposons $f : A \to \mathbb{R}$ n fois dérivable en $a \in A$. Alors, infinitésimalement en a, on a*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + o((x-a)^n)$$

La notation $o((x-a)^n)$ signifie : il existe une fonction $r : A \to \mathbb{R}$ telle que

$$\lim_{x \to a} r(x) = 0$$

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + (x-a)^n r(x)$$

La dérivée seconde s'interprète aussi en terme de convexité.

Proposition 1.10. *Soit $f :]a, b[\to \mathbb{R}$ une fonction deux fois dérivable sur $]a, b[$, alors $f'' > 0$ sur $]a, b[$ si et seulement si f est convexe sur $]a, b[$, c'est-à-dire pour tout x, $y \in]a, b[$, pour tout $t \in [0, 1]$,*

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

1.4 Fonctions à valeurs dans un espace de dimension finie

La notion de dérivée s'étend directement aux fonctions $f : A \to F$ où A est toujours un sous-ensemble de $\mathbb{R}$, mais F est un espace vectoriel réel de dimension finie quelconque.

Théorème 1.11. *Sur un espace vectoriel de dimension finie, toutes les normes sont équivalentes. Elles définissent les mêmes limites, fonctions continues, etc (la même* topologie*).*

Définition 1.12. La fonction $f : A \to F$ est dérivable en $a \in \mathbb{R}$ si :

1. A est un voisinage de a,
2. $\frac{f(a+t)-f(a)}{t}$ a une limite dans F quand t tend vers 0.

Dans ce cas, on appelle cette limite la *dérivée* de f en a, notée $f'(a)$. C'est un élément de F.

Par le théorème d'équivalence des normes rappelé ci-dessus, on peux choisir une norme $\|\cdot\|$ sur F pour formaliser la limite :

$$\lim_{t\to 0} \frac{f(a+t)-f(a)}{t} = f'(a) \quad \Leftrightarrow \quad \lim_{t\to 0} \left\| \frac{f(a+t)-f(a)}{t} - f'(a) \right\| = 0$$

Pour interpréter cette notion, on peux penser par exemple pour $f : A \to \mathbb{R}^2$, que $f(t)$ décrit la position d'un point en mouvement dans le plan, en fonction du temps $t \in A$. Alors $f'(t)$ représente le vecteur vitesse, qui indique la direction, le sens, et la rapidité (infinitésimale) du mouvement.

2 Introduction aux équations différentielles

2.1 Définition informelle et motivation

Commençons par une définition informelle : une *équation différentielle* est une équation dont l'inconnue est une fonction y d'une variable réelle, et qui implique des dérivées (première ou supérieures) de la fonction. L'*ordre* d'une équation différentielle est le plus grand $n \in \mathbb{N}$ tel que la dérivée n-ième soit impliquée.

Exemple 2.1.

$$y' = 0 \qquad y'(t) = t^2 \qquad y' = y \qquad y'' = y \qquad \arctan((y')^2 y + y' + 2) = 0 \qquad \cdots$$

La motivation pour étudier des équations différentielles en mathématiques provient souvent d'une autre science. En biologie, on peut par exemple s'intéresser à la dynamique d'une population, où la croissance d'une colonie de bactéries.

On considère, par exemple, la fonction y d'une variable réelle, telle que $y(t)$ soit le nombre d'individus dans une colonie de bactéries *e. coli.* Comment obtient-on une équation différentielle ? On peux mesurer $y(t)$ à des intervalles de temps fixe espacés de h, et $h > 0$ aussi petit que les instruments de mesure le permettent. En étudiant les données, on constate que le taux de croissance relatif $\frac{y(t+h)-y(t)}{hy(t)}$ est presque constant (dans des conditions optimales, pour une plage de temps restreinte). Si on veux un modèle continu, on interpréte la quantité ci-dessus comme sa limite quand $h \to 0$:

$$\frac{y(t+h) - y(t)}{hy(t)} \to \frac{y'(t)}{y(t)}$$

et d'après les mesures, on considère que cette fonction est constante. Donc la fonction y est solution d'une équation différentielle de la forme

$$y' = ry$$

où $r > 0$ est un paramètre. On connaît les solutions à cette équation :

$$y(t) = y(0)e^{rt}$$

ce qui signifie que le nombre de bactéries augmente de manière exponentielle.

Il s'agit d'un modèle simple, pour lequel on peux donner les solutions de manière explicite. Cependant, le modèle est souvent trop simple : la population ne peux pas vraiment tendre vers $+\infty$ comme la solution exacte, car il y a toujours une limite aux ressources disponible. Un modèle simple permettant de prendre ça en compte est l'équation logistique :

$$y' = ry\left(1 - \frac{y}{K}\right)$$

où $K > 0$ est un autre paramètre. Ici, les solutions sont moins évidentes à trouver. On peux résoudre par séparation des variables, mais pour le faire proprement, il faut utiliser le Théorème de Cauchy-Lipschitz. On peux aussi en faire une étude qualitative, toujours en utilisant le Théorème de Cauchy-Lipschitz, pour apprendre, sans résoudre explicitement, qu'en partant d'une population initiale entre 0 et K, la population sera croissante, et convergera vers K. Un objectif du cours est de comprendre comment faire ça pour beaucoup d'équations différentielles.

Une autre source d'équations différentielle est la physique. Par exemple, la deuxième loi du mouvement de Newton qui dit que la masse fois l'accélération (la dérivée de la vitesse, donc la dérivée seconde de la position) est égale à la somme des forces exercées, donne souvent une équation différentielle. La chute libre verticale d'un corps est par exemple régie par l'équation différentielle

$$y'' = -g$$

où y encode l'altitude, et g est la constante de gravitation.

La fonction inconnue prendra en général ses valeurs dans un espace vectoriel de dimension finie, et pas seulement dans $\mathbb{R}$. C'est utile, par exemple, pour modéliser plusieurs populations en interaction (prédation, compétition, aide mutuelle). C'est aussi utile pour ramener une équation d'ordre élevé à une équation d'ordre 1. Pour illustrer dans le cas très simple de la chute libre $y'' = -g$, au lieu de considérer seulement la position y, on considère aussi la vitesse $v = y'$. Alors l'équation $y'' = -g$ peut s'écrire $v' = -g$, mais pour récupérer y il faut se rappeler l'équation $y' = v$. Ce qu'on est en train de considérer, c'est une équation d'ordre 1 sur la fonction inconnue qui a deux composantes (y, v). Elle est ici à valeur dans $\mathbb{R}^2$, mais si on considère que y est à valeur dans $\mathbb{R}^2$ ou $\mathbb{R}^3$, alors (y, v) est à valeurs dans $\mathbb{R}^4$ où $\mathbb{R}^6$. On voit rapidement qu'il peut être utile de travailler dans des espaces de grande dimension.

2.2 Conventions dans ce cours

Pour ce cours, on va travailler avec une forme particulière d'équations différentielles, de la forme

$$y' = f(t, y)$$

où $f : I \times U \to E$ est une fonction, I est un intervalle ouvert de $\mathbb{R}$, E est un espace vectoriel réel de dimension finie, et U est un ouvert de E. La fonction inconnue y est une fonction d'une variable réelle, à valeurs dans E.

On appellera souvent la variable $t \in I$ la *variable de temps*, et la variable $x \in E$ la *variable d'espace.*

Définition 2.2.

1. Une *solution* est une fonction $y : J \to E$, définie sur un intervalle $J \subset I$, dont l'image est incluse dans U, dérivable sur J, et telle que, pour tout $t \in J$, $y'(t) = f(t, y(t))$.
2. Si $y_1 : J_1 \to E$ et $y_2 : J_2 \to E$ sont deux solutions telles que $J_1 \subsetneq J_2$ et $y_2|_{J_1} = y_1$, on dit que y_2 *prolonge* (strictement) y_1.
3. Une solution $y : J \to E$ est *maximale* s'il n'existe aucune solution qui prolonge y.
4. Un *problème de Cauchy* pour l'équation différentielle $y' = f(t, y)$ est la donnée d'une condition initiale $(t_0, y_0) \in I \times U$. On note ça
$$\begin{cases} y' = f(t, y) \\ y(t_0) = x_0 \end{cases}$$
5. Une *solution au problème de Cauchy* ci-dessus est une solution $y : J \to E$ à l'équation différentielle, telle que $t_0 \in J$ et $y(t_0) = y_0$.

3 Théorème de Cauchy-Lipschitz

On se place dans le cadre de la section précédente, où on considère une équation de la forme $y' = f(t, y)$ avec $f : I \times U \subset \mathbb{R} \times E \to E$, et on fixe une norme $\|\cdot\|$ sur E.

3.1 Fonctions localement, Lipschitziennes en espace

Définition 3.1. Une fonction $f : I \times U \to E$ est *localement, Lipschitzienne en espace* si pour tout $(t_0, x_0) \in I \times U$, il existe $\varepsilon > 0$ et $\lambda \geq 0$ tels que si $t \in I$, x_1, $x_2 \in U$ satisfont $\sup\{|t - t_0|, \|x_1 - x_0\|, \|x_2 - x_0\|\} < \varepsilon$, alors

$$\|f(t, x_1) - f(t, x_2)\| \leq \lambda \|x_1 - x_2\|$$

Exemple 3.2. Une fonction M-Lipschitzienne est localement, Lipschitzienne en espace : pour tout (t_0, x_0) on peut prendre n'importe quel $\epsilon > 0$ et $\lambda = M$ dans la définition. Par exemple, $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (t, x) \mapsto x$ est 1-Lipschitzienne, donc vérifie la définition.

Exemple 3.3. La fonction $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (t, x) \mapsto tx$ n'est pas globalement Lipschitzienne, mais localement, Lipschitzienne en espace. En effet, en prenant (t_0, x_0) quelconque, $\epsilon = 1$ et $\lambda = |t_0| + 1$, on a, si $\sup\{|t - t_0|, \|x_1 - x_0\|, \|x_2 - x_0\|\} < 1$,

$$|f(t, x_1) - f(t, x_2)| = |t||x_1 - x_2| \leq \lambda |x_1 - x_2|$$

Pour voir qu'elle n'est pas globalement Lipschitzienne, on observe que, pour $t > 0$,

$$t = \frac{|f(t, 1) - f(t, 0)|}{|1 - 0|} \to_{t\to\infty} +\infty$$

Exemple 3.4. De même, la fonction $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (t, x) \mapsto x^2$ n'est pas globalement Lipschitzienne, mais localement, Lipschitzienne en espace. Même à t fixé, on a, par exemple pour $x > 0$,

$$x = \frac{|f(t, x) - f(t, 0)|}{|x - 0|} \to_{x\to\infty} +\infty$$

Par contre, si on prends $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}$, $\varepsilon = 1$ et $\lambda = 2(|x_0| + 1)$, on a pour $\sup\{|t - t_0|, \|x_1 - x_0\|, \|x_2 - x_0\|\} < 1$,

$$|f(t, x_1) - f(t, x_2)| = |x_1^2 - x_2^2| = |x_1 + x_2||x_1 - x_2| \leq \lambda |x_1 - x_2|$$

Remarque 3.5. Si $E = \mathbb{R}$, pour tout t fixé, la fonction $x \mapsto f(t, x)$ est une fonction réelle d'une variable réelle. Si elle est dérivable, on peux contrôler si elle est Lipschitzienne grâce à l'inégalité des accroissements finis. On appliquera ça dès aujourd'hui en TD, et on reviendra sur la généralisation en plusieurs variables plus tard dans le cours.

3.2 Énoncé du Théorème

Théorème 3.6 (Théorème de Cauchy-Lipschitz)**.** *On suppose que $f : I \times U \to E$ est une fonction continue et localement, Lipschitzienne en espace. Alors pour tout $(t_0, x_0) \in I \times U$, il existe une unique solution maximale au problème de Cauchy*

$$\begin{cases} y' = f(t, y) \\ y(t_0) = x_0 \end{cases}$$

Il y a bien deux hypothèses à vérifier : la continuité, et le caractère localement, Lipschitzienne en espace. Si on vérifie par exemple que la fonction est globalement Lipschitzienne, ça implique directement que les deux hypothèses sont satisfaites.

Il y a également plusieurs conclusions. Il y a bien sûr un résultat théorique très fort d'existence de solution. Nous ne donnerons pas de preuve dans ce cours, car cela fait appel à quelques notions de topologie sur les espaces de fonctions que vous ne connaissez pas.

L'autre conclusion est l'unicité, qui est en fait un peu plus fine que l'énoncé court ci-dessus ne le laisse entendre. Nous préciserons ce résultat d'unicité et sa démonstration dans une section à venir. Cela permettra aussi de clarifier la notion de solution maximale, sous les hypothèses du Théorème de Cauchy-Lipschitz.

3.3 Premier exemple d'application : variable séparable

Malgré ses conclusion principalement théoriques, le Théorème de Cauchy-Lipschitz permet aussi de déterminer proprement les solutions explicites de certaines équations. On a besoin de l'unicité du Théorème pour avoir des informations *a priori* sur les solutions, et mener un raisonnement intuitif jusqu'au bout. On l'illustre sur un exemple, et on le refera en TD.

Exemple 3.7. On considère la fonction $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (t, x) \mapsto tx^3$, et le problème de Cauchy

$$\begin{cases} y' = ty^3 \\ y(0) = 1 \end{cases}$$

La fonction f est localement, Lipschitzienne en espace. En effet, pour $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}$, prenons $\varepsilon = 1$ et $\lambda = 3(|t_0| + 1)(|x_0| + 1)^2$. On a alors, pour $\sup\{|t - t_0|, \|x_1 - x_0\|, \|x_2 - x_0\|\} < 1$,

$$\begin{aligned} |f(t, x_2) - f(t, x_1)| &= |t||x_2^3 - x_1^3| \\ &= |t||x_2^2 + x_2x_1 + x_1^2||x_2 - x_1| \\ &\leq \lambda|x_2 - x_1| \end{aligned}$$

Par le théorème de Cauchy-Lipschitz, on sait déjà qu'il existe une solution maximale à ce problème de Cauchy, notons la $u : J \to \mathbb{R}$.

L'équation $y' = ty^3$ est une équation à variables séparables, c'est-à-dire qu'on pourrait mettre toutes les variables y d'un côté de l'équation, et toutes les variables t de l'autre :

$$\frac{y'}{y^3} = t$$

Cependant, pour pouvoir faire ça (diviser par y^3), il faut savoir à l'avance que y ne s'annule pas. On utilise l'unicité dans le Théorème de Cauchy-Lipschitz pour montrer ça.

Supposons en effet qu'il existe $t_0 \in J$ avec $u(t_0) = 0$. Alors u est une solution (maximale) au problème de Cauchy

$$\begin{cases} y' = ty^3 \\ y(t_0) = 0 \end{cases}$$

Cependant, il est facile de vérifier que la fonction nulle partout est aussi une solution, maximale, à ce problème de Cauchy. Par unicité, on en déduirait $u(t) = 0$ pour tout $t \in J$, ce qui contredit la condition initiale $u(0) = 1$.

Maintenant, on peut écrire, pour tout $t \in J$,

$$\frac{u'(t)}{u(t)^3} = t$$

En intégrant, on peut écrire

$$\begin{aligned}
\frac{t^2}{2} &= \int_0^t s\,ds \\
&= \int_0^t \frac{u'(s)}{u(s)^3}\,ds \\
&= \left[\frac{-1}{2u(s)^2}\right]_0^t \\
&= \frac{-1}{2u(t)^2} + \frac{1}{2}
\end{aligned}$$

d'où

$$\frac{1}{u(t)^2} = 1 - t^2$$

Comme $u(t) \in \mathbb{R}$ est bien défini, on a nécessairement $1 - t^2 > 0$ et

$$u(t) = \frac{\pm 1}{\sqrt{1-t^2}}$$

Après vérification rapide, l'unique solution maximale est donc la fonction définie par

$$u(t) = \frac{1}{\sqrt{1-t^2}}$$

sur l'intervalle $J =]-1, 1[$.

4 Unicité locale

Pour la première étape de la preuve de l'unicité, on a besoin de quelques outils.

4.1 Lemme de Grönwall

Lemme 4.1 (Lemme de Grönwall). *Soient $t_0 < t_1$ des réels, K un réel, $a : [t_0, t_1] \to \mathbb{R}_+$ une fonction continue à valeurs positives, et $v : [t_0, t_1] \to \mathbb{R}$ une fonction continue. On suppose que, pour tout t dans $[t_0, t_1]$,*

$$v(t) \leq K + \int_{t_0}^t a(s)v(s)\,ds$$

Alors pour tout t dans $[t_0, t_1]$,

$$v(t) \leq K \exp\left(\int_{t_0}^t a(s)\,ds\right)$$

Démonstration. Posons $w(t) = K + \int_{t_0}^t a(s)v(s)\,ds$ (le second membre de l'inégalité en hypothèse). Le principe de la preuve est de montrer directement que $w(t)$ est plus petit que le terme de droite de la conclusion, en intégrant une inéquation différentielle linéaire que w satisfait.

Comme a et v sont continues, par le théorème fondamental de l'analyse, w est dérivable et $w'(t) = a(t)v(t)$. De plus, $a(t) \geq 0$ et $v(t) \leq w(t)$ par hypothèses, donc on a

$$w'(t) \leq a(t)w(t)$$

Par analogie avec la résolution des équations différentielles linéaires d'ordre 1 revue en TD, on considère la dérivée de la fonction

$$z(t) = w(t)e^{-\int_{t_0}^t a(s)\,ds}$$

On a, grâce à l'inéquation différentielle,

$$z'(t) = (w'(t) - a(t)w(t))e^{-\int_{t_0}^t a(s)\,ds} \leq 0$$

donc z est décroissante sur $[t_0, t_1]$. En particulier, pour tout t dans $[t_0, t_1]$, on a $z(t) \leq z(t_0) = w(t_0) = K$. En conclusion,

$$v(t) \leq w(t) = z(t)e^{\int_{t_0}^t a(s)\,ds} \leq Ke^{\int_{t_0}^t a(s)\,ds}$$

□

4.2 Formulation intégrale

Pour appliquer le lemme de Grönwall pour l'étude des équations différentielles, on introduit la *formulation intégrale d'une équation différentielle.*

Proposition 4.2. *On suppose que $f : I \times U \to E$ continue et $(t_0, x_0) \in I \times U$. Soit $u : J \to U$ une fonction. Alors u est solution du problème de Cauchy*

$$\begin{cases} y' = f(t, y) \\ y(t_0) = x_0 \end{cases}$$

si et seulement si u et continue et pour tout $t \in J$,

$$u(t) = x_0 + \int_{t_0}^t f(s, u(s))\,ds$$

Démonstration. La preuve est essentiellement immédiate : si u est solution (en particulier dérivable, donc continue), alors on intègre l'équation différentielle et on obtient l'équation intégrale.

Réciproquement, comme u et f sont continues, $s \mapsto f(s, u(s))$ l'est aussi par composition. Donc la fonction $u : t \mapsto x_0 + \int_{t_0}^t f(s, u(s))\,ds$ est dérivable, de dérivée $f(t, u(t))$ par le théorème fondamental de l'analyse. De plus, elle vérifie bien $u(t_0) = x_0$. □

4.3 Unicité locale des solutions

Lemme 4.3 (Unicité locale). *On suppose $f : I \times U \to E$ continue et localement, Lipschitzienne en espace. Si $u_1 : J_1 \to E$ et $u_2 : J_2 \to E$ sont deux solutions d'un même problème de Cauchy*

$$\begin{cases} y' = f(t, y) \\ y(t_0) = x_0 \end{cases}$$

alors il existe un intervalle ouvert J tel que $t_0 \in J \subset J_1 \cap J_2$ et $u_1|_J = u_2|_J$.

Démonstration. On utilise d'abord l'hypothèse que f est localement, Lipschitzienne en espace. Soient $\varepsilon > 0$ et $\lambda \geq 0$ tels que, pour $\sup\{|t - t_0|, \|x_1 - x_0\|, \|x_2 - x_0\|\} < \varepsilon$,

$$\|f(t, x_1) - f(t, x_2)\| \leq \lambda \|x_1 - x_2\|$$

Comme u_1 est continue (car dérivable) et $u_1(t_0) = x_0$, il existe $0 < \varepsilon_1 \leq \varepsilon$ tel que

$$\forall t \in]t_0 - \varepsilon_1, t_0 + \varepsilon_1[, \qquad \|u_1(t) - x_0\| < \varepsilon$$

De même, comme u_2 est continue (car dérivable) et $u_2(t_0) = x_0$, il existe $0 < \varepsilon_2 \leq \varepsilon_1$ tel que

$$\forall t \in]t_0 - \varepsilon_2, t_0 + \varepsilon_2[, \qquad \|u_2(t) - x_0\| < \varepsilon$$

On en déduit, pour $t \in]t_0 - \varepsilon_2, t_0 + \varepsilon_2[$,

$$\|f(t, u_1(t)) - f(t, u_2(t))\| \leq \lambda \|u_1(t) - u_2(t)\|$$

Maintenant, on utilise la formulation intégrale pour chacune des deux solutions pour écrire, pour $t \in]t_0, t_0 + \varepsilon_2[$,

$$\begin{aligned} \|u_2(t) - u_1(t)\| &= \left\| \int_{t_0}^{t} (f(s, u_1(s)) - f(s, u_2(s)))\, ds \right\| \\ &\leq \int_{t_0}^{t} \|f(s, u_1(s)) - f(s, u_2(s))\|\, ds \\ &\leq \int_{t_0}^{t} \|\lambda \|u_1(s) - u_2(s)\|\|\, ds \end{aligned}$$

On reconnait l'hypothèse du lemme de Grönwall, en prenant $t_0 < t_1 < t_0 + \varepsilon_2$, $v(t) = \|u_1(t) - u_2(t)\|$, $a(t) = \lambda$ et $K = 0$. La conclusion du lemme de Grönwall donne dans ce cas

$$\|u_1(t) - u_2(t)\| \leq 0$$

pour $t \in [t_0, t_1]$ pour tout $t_1 \in]t_0, t_0 + \varepsilon_2[$, donc pour tout $t \in [t_0, t_0 + \varepsilon_2[$. Autrement dit, on a montré

$$u_2|_{[t_0, t_0 + \varepsilon_2[} = u_1|_{[t_0, t_0 + \varepsilon_2[}$$

On adapte facilement la preuve ci-dessus pour montrer la même chose sur $]t_0 - \varepsilon_2, t_0]$, donc le résultat avec $J =]t_0 - \varepsilon_2, t_0 + \varepsilon_2[$. □

5 Solutions maximales et unicité

5.1 Preuve de l'unicité dans Cauchy-Lipschitz

Montrons maintenant le résultat d'unicité dans le théorème de Cauchy-Lipschitz, sous une version un peu plus précise, en utilisant le lemme d'unicité locale précédent, et un raisonnement qui utilise de manière fondamentale l'ordre sur $\mathbb{R}$ (qui se traduit en connexité, comme on en reparlera).

Théorème 5.1 (Unicité)**.** *Soit $f : I \times U \to E$ une fonction continue et localement Lipschitzienne en espace. Si $u_1 : J_1 \to E$ et $u_2 : J_2 \to E$ sont deux solutions d'un même problème de Cauchy*

$$\begin{cases} y' = f(t, y) \\ y(t_0) = x_0 \end{cases}$$

alors pour tout $t \in J_1 \cap J_2$, $u_1(t) = u_2(t)$.

Démonstration. On a $J_1 \cap J_2 =]\inf(J_1 \cap J_2), \sup(J_1 \cap J_2)[$. Considérons

$$t_m = \sup\{t \in J_1 \cap J_2; \quad (u_2 - u_1)|_{[t_0,t[} \equiv 0\}$$

et supposons, par l'absurde, que $t_m < \sup(J_1 \cap J_2)$. On a alors $t_m \in J_1 \cap J_2$, et par continuité,

$$(u_2 - u_1)(t_m) = \lim_{\substack{t \to t_m \\ t < t_m}} (u_2 - u_1)(t) = 0$$

Cela signifie que u_1 et u_2 sont solutions du même problème de Cauchy

$$\begin{cases} y' = f(t, y) \\ y(t_m) = u_1(t_m) \end{cases}$$

donc par le lemme d'unicité locale, il existe $\varepsilon > 0$ tel que

$$(u_2 - u_1)|_{]t_m - \varepsilon, t_m + \varepsilon[} = 0$$

Par conséquent, on a

$$(u_2 - u_1)|_{]t_0, t_m + \varepsilon[} = 0$$

ce qui contredit la définition de t_m.

On a montré $t_m = \sup(J_1 \cap J_2)$. On peux adapter la preuve pour les temps $\leq t_0$, ce qui donne le théorème. □

5.2 Quelques conséquences

Une conséquence principale, déjà utilisée plus tôt, est que deux solutions maximales ne peuvent pas se "croiser". On en verra de nombreuses applications dans les exercices d'études qualitatives. Plus théoriquement, le résultat permet de préciser comment construire des solutions maximales, et pourquoi toute solution se prolonge en une solution maximale, sous les hypothèses de Cauchy-Lipschitz.

Développons un peu ce point. On suppose f continue et localement, Lipschitzienne en espace. Soit $u : J \to E$ une solution de $y' = f(t, y)$, a priori non maximale. Pour construire l'unique solution maximale qui la prolonge, on procède de la manière suivante. Soit $\mathcal{S}$ l'ensemble des solutions $v : K \to E$ de $y' = f(t, y)$ qui prolongent (strictement ou non) u, c'est-à-dire que $J \subset K$ et $v|_J = u$. On pose

$$\hat{J} := \bigcup_{(v:K\to E)\in\mathcal{S}} K$$

et $\hat{u} : \hat{J} \to E$ la fonction définie par, si $t \in K$ pour $(v : K \to E) \in \mathcal{S}$, $\hat{u}(t) = v(t)$.

La fonction $\hat{u}$ est bien définie par le théorème d'unicité prouvé plus haut. C'est clairement une solution à l'équation différentielle. De plus, elle est maximale : s'il existait un prolongement strict de $\hat{u}$, alors ce serait un élément de $\mathcal{S}$, et ça contredirait la définition de $\hat{J}$. Donc $\hat{u}$ est l'unique solution maximale qui prolonge u.

6 Solutions maximales vs solutions globales

6.1 Solution globale

On utilise toujours les mêmes notations pour l'équation différentielle.

Définition 6.1. Une solution $u : J \to E$ de $y' = f(t, y)$ avec $f : I \times U \to E$ est *globale* si $J = I$.

Une solution globale est maximale. En effet, si $v : K \to E$ était un prolongement strict de u solution globale, on aurait $I \subsetneq K$, or $K \subset I$ car v est une solution...

Par contre, la réciproque est fausse. On a vu par exemple que la fonction $u :]-1, 1[\to \mathbb{R}, t \mapsto \frac{1}{\sqrt{1-t^2}}$ est une solution maximale de $y' = ty^3$, alors que $I = \mathbb{R}$ dans ce cas.

Plusieurs outils permettent de démontrer a priori qu'une solution maximale est en fait globale.

6.2 Théorème de Cauchy-Lipschitz global

Le premier outil est une variante du Théorème de Cauchy-Lipschitz, il se prouve d'ailleurs de la même manière. Par contre, ses hypothèses sont nettement plus fortes. Commençons par une définition.

Définition 6.2. La fonction f est *localement en temps, Lipschitzienne en espace*, si pour tout $t_0 \in I$, il existe $\varepsilon > 0$ et $\lambda \geq 0$ tels que :

$$\forall t \in]t_0 - \varepsilon, t_0 + \varepsilon[\qquad \text{(localement en temps)}$$

$$\forall x_1, x_2 \in U, \qquad \|f(t, x_1) - f(t, x_2)\| \leq \lambda \|x_1 - x_2\| \qquad \text{(Lipschitzienne en espace)}$$

Théorème 6.3 (Théorème de Cauchy-Lipschitz global)**.** *Si $U = E$ et f est continue, et localement en temps, Lipschitzienne en espace, alors pour toute condition initiale $(t_0, x_0) \in I \times E$, il existe une unique solution globale au problème de Cauchy*

$$\begin{cases} y' = f(t, y) \\ y(t_0) = x_0 \end{cases}$$

On verra une application du Théorème de Cauchy-Lipschitz global pour l'étude des équations différentielles linéaires bientôt.

6.3 Sortie de tout compact

Le théorème suivant est plus général : sous les hypothèses de Cauchy-Lipschitz, il montre qu'une solution maximale non globale a un comportement très particulier. Ceci permet souvent de démontrer que des solutions maximales sont globales, mais aussi, lorsqu'elles ne le sont pas, de préciser leur comportement asymptotique, dans le cadre d'une étude qualitative.

Pour simplifier les notations, on note $I =]a, b[$ dans la suite.

Théorème 6.4. *On suppose $f :]a, b[\times U \to E$ continue et localement, Lipschitzienne en espace. Soit $u :]c, d[\to E$ une solution maximale de $y' = f(t, y)$. Alors :*
- *si $d < b$, alors pour tout sous-ensemble compact K de U, il existe un $\delta \in]c, d[$ tel que pour $t \in]\delta, d[$, $u(t) \in U \setminus K$*
- *si $a < c$, alors pour tout sous-ensemble compact K de U, il existe un $\gamma \in]c, d[$ tel que pour $t \in]a, \gamma[$, $u(t) \in U \setminus K$.*

Les deux conclusions sont très similaires, elles portent sur les deux côtés de l'intervalle. L'interprétation informelle de la conclusion est que les valeurs $u(t)$ restent en dehors de tout compact de U quand t tends vers une des bornes de l'ensemble de définition de la solution.

On rappelle que K est un compact si c'est un sous-ensemble fermé et borné de E. Dans l'énoncé, le fait que K soit un sous-ensemble de U est important, lorsque $U \neq E$.

Voici une manière très directe d'appliquer ce résultat pour montrer qu'une solution est globale.

Corollaire 6.5. *S'il existe un sous-ensemble compact K de U tel que, pour tout $t \in]c, d[$, $u(t) \in K$, alors la solution u est globale.*

Dans le cas particulier où $U = E$, le théorème porte un autre nom.

Corollaire 6.6 (Théorème d'explosion en temps fini)**.** *Supposons $f :]a, b[\times E \to E$ continue, localement Lipschitzienne en espace.*

Soit $u :]c, d[\to E$ une solution maximale de $y' = f(t, y)$. Alors :
- *si $d < b$, $\lim_{t\to d}\|u(t)\| = +\infty$*
- *si $a < c$, $\lim_{t\to c}\|u(t)\| = +\infty$*

C'est un corollaire, car les limites peuvent être interprétées comme : $u(t)$ sort de toute boule fermée centrée en zéro lorsque $t \to c$ ou d, et les boules fermées sont des compacts dans l'espace vectoriel réel de dimension finie E.

Exemple 6.7. On a vu ce type de comportement pour l'équation $y' = ty^3$. En effet, la solution maximale $u :]-1, 1[\to \mathbb{R}, t \mapsto \frac{1}{\sqrt{1-t^2}}$ n'est pas globale, et elle vérifie bien

$$\lim_{t\to\pm1} |u(t)| = \lim_{t\to\pm1} \frac{1}{\sqrt{1-t^2}} = +\infty$$

On fera la preuve du Théorème de sortie de tout compact plus tard.

6.4 Exemple d'étude qualitative

Bien souvent, l'étude qualitative des équations différentielles repose sur le cycle suivant :

1. l'équation différentielle donne une information sur y'
2. en intégrant, on obtient une information sur y
3. on réinjecte cette information dans l'équation différentielle, et on reprends.

Nous allons illustrer ceci sur un exemple, pour voir comment utiliser le Théorème d'explosion en temps fini.

Exemple 6.8. On considère

$$f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad (t, x) \mapsto e^{-tx}$$

Cette fonction est localement, Lipschitzienne en espace. Pour le démontrer, utilisons le théorème des accroissements finis en une variable. Pour $t \in \mathbb{R}$ fixé, on a

$$\frac{\mathrm{d}}{\mathrm{d}x} e^{-tx} = -te^{-tx}$$

et donc, si $x_1 < x_2 \in \mathbb{R}$ il existe $x_3 \in]x_1, x_2[$ tel que

$$f(t, x_2) - f(t, x_1) = -te^{-tx_3}(x_2 - x_1)$$

Notons que si $|x_1 - x_0| < \varepsilon$ et $|x_2 - x_0| < \varepsilon$, alors $|x_3 - x_0| < \varepsilon$, donc

$$|f(t, x_2) - f(t, x_1)| = te^{-tx_3}|x_2 - x_1| \leq |t|e^{|t|(|x_0|+\varepsilon)}|x_2 - x_1|$$

On peut donc vérifier que f est localement, Lipschitzienne en espace d'après la définition. Si $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}$, prenons $\varepsilon = 1$ et $\lambda = (|t_0| + 1)e^{(|t_0|+1)(|x_0|+1)}$ Alors si $|t - t_0| < \varepsilon$, $|x_1 - x_0| < \varepsilon$, $|x_2 - x_0| < \varepsilon$, on a

$$|f(t, x_2) - f(t, x_1)| \leq \lambda |x_2 - x_1|$$

On peut donc appliquer le Théorème de Cauchy-Lipschitz, et le Théorème d'explosion en temps fini, et faire un début d'étude qualitative des solutions.

Soit $u :]c, d[\to \mathbb{R}$ une solution maximale du problème de Cauchy

$$\begin{cases} y' = f(t, y) \\ y(t_0) = x_0 \end{cases}$$

L'équation différentielle donne une information sur u' : comme une exponentielle est toujours strictement positive, pour tout $t \in]c, d[$, $u'(t) = e^{-tu(t)} > 0$. On en déduit que u est strictement croissante. D'après les conditions initiales ($u(0) = 0$), on a donc $u(t) > 0$ sur $]0, d[$ et $u(t) < 0$ sur $]c, 0[$. On réinjecte cette information dans l'équation différentielle.

Le signe de u assure que $tu(t) \geq 0$ partout. Donc $0 < u'(t) \leq e^0 = 1$ partout. On intégre à nouveau, pour $t \geq 0$:

$$u(t) = u(0) + \int_0^t u'(s)\, ds \leq 0 + \int_0^t 1\, ds = t$$

Appliquons maintenant le Théorème d'explosion en temps fini. Si $d \neq +\infty$, alors on devrait avoir $\lim_{t \to d} |u(t)| = +\infty$. C'est impossible car $u(t) \leq t \leq d$... Donc $d = +\infty$. Le même type de raisonnement montre que $c = -\infty$, donc la solution maximale est globale.

En conclusion, on a montré, sans connaître explicitement la solution maximale, que c'est une fonction définie sur $\mathbb{R}$, strictement croissante, avec $u(t) \leq t$ pour $t \in]0, +\infty[$.

7 Structure des solutions des EDL

On se place pour les équations différentielles linéaires (EDL) dans le cas où $E = \mathbb{R}^n$ pour simplifier les notations. On fixe toujours une norme $\|\cdot\|$ sur $\mathbb{R}^n$.

7.1 Norme triple de matrices

On note $\mathrm{M}_n(\mathbb{R})$ l'ensemble des matrices carrées de taille n à coefficients réels. On considère la norme triple $|||\cdot|||$ sur $\mathrm{M}_n(\mathbb{R})$. Il s'agit de la norme définie par

$$|||M||| := \sup \left\{ \frac{\|Mx\|}{\|x\|} \mid 0 \neq x \in \mathbb{R}^n \right\}$$

En exercice, vérifier que c'est bien défini, et que ça donne bien une norme ! On s'en resservira plusieurs fois dans le cours. On note que, pour $M \in \mathrm{M}_n(\mathbb{R})$ et $x \in \mathbb{R}^n$, on a toujours

$$\|Mx\| \leq |||M||| \|x\|$$

La norme triple est sous-multiplicative : si M, $N \in \mathrm{M}_n(\mathbb{R})$, alors

$$|||MN||| \leq |||M||| |||N|||$$

En effet, si $x \in \mathbb{R}^n$, on a

$$\|M(Nx)\| \leq |||M||| \|Nx\| \leq |||M||| |||N||| \|x\|$$

donc

$$|||MN||| = \sup \left\{ \frac{\|MNx\|}{\|x\|} \mid 0 \neq x \in \mathbb{R}^n \right\} \leq |||M||| |||N|||$$

7.2 EDL

Définition 7.1. L'équation différentielle $y' = f(t, y)$ est *linéaire* si f s'écrit sous la forme

$$\begin{aligned} f : & I \times \mathbb{R}^n \to \mathbb{R}^n \\ & (t, x) \mapsto A(t)x + b(t) \end{aligned}$$

où $A : I \to \mathrm{M}_n(\mathbb{R})$ et $b : I \to \mathbb{R}$. Elle est de plus *homogène* si $b \equiv 0$, et elle est *à coefficients constants* si la fonction $A : I \to \mathrm{M}_n(\mathbb{R})$ est constante.

Comme c'est indiqué dans la définition, pour les équations différentielles linéaires, on prends toujours $U = \mathbb{R}^n$, parce que, pour $t \in I$, l'expression $A(t)x + b(t)$ est toujours bien définie pour tout $x \in \mathbb{R}^n$.

Théorème 7.2 (Théorème de Cauchy-Lipschitz linéaire)**.** *On suppose que* $A : I \to \mathrm{M}_n(\mathbb{R})$ *et* $b : I \to \mathbb{R}^n$ *sont continues. Alors pour tout* $(t_0, x_0) \in I \times \mathbb{R}^n$*, le problème de Cauchy*

$$\begin{cases} y' = A(t)y + b(t) \\ y(t_0) = x_0 \end{cases}$$

admet une unique solution globale.

Démonstration. Il suffit de vérifier que le Théorème de Cauchy-Lipschitz global s'applique, c'est-à-dire que la fonction $f : (t, x) \mapsto A(t)x + b(t)$ est localement en temps, Lipschitzienne en espace.

Soit $t_0 \in I$, et $\varepsilon > 0$ tel que $[t_0 - \varepsilon, t_0 + \varepsilon] \subset I$. La fonction $I \to \mathbb{R}, t \mapsto |||A(t)|||$ est continue (par composition : A est continue par hypothèse, et $|||\cdot|||$ est continue par "définition" de la topologie de l'espace vectoriel de dimension finie $\mathrm{M}_n(\mathbb{R})$ et par équivalence des normes). Elle admet donc un maximum M sur le segment $[t_0 - \varepsilon, t_0 + \varepsilon]$. On a donc, pour tout $t \in]t_0 - \varepsilon, t_0 + \varepsilon[$, pour tout x_1, $x_2 \in \mathbb{R}^n$,

$$\begin{aligned} \|A(t)x_1 + b(t) - (A(t)x_2 + b(t))\| &= \|A(t)(x_1 - x_2)\| \\ &\leq |||A(t)||| \|x_1 - x_2\| \\ &\leq M\|x_1 - x_2\| \end{aligned}$$

□

7.3 Structure des solutions d'une EDL homogène

On note $C^1(I, \mathbb{R}^n)$ l'espace vectoriel réel des fonctions définies sur I, à valeurs dans $\mathbb{R}^n$, dérivables et dont la dérivée est continue.

Théorème 7.3. *Soit* $A : I \to \mathrm{M}_n(\mathbb{R})$ *une fonction continue, et* $t_0 \in I$*. Alors l'ensemble* $\mathcal{E}$ *des solutions à l'équation différentielle* $y' = A(t)y$ *est un sous-espace vectoriel de dimension* n *de* $C^1(I, \mathbb{R}^n)$*. De plus, si* $(e_1, \ldots, e_n)$ *est une base de* $\mathbb{R}^n$*, et pour tout* i*,* u_i *est la solution globale du problème de Cauchy*

$$\begin{cases} y' = A(t)y \\ y(t_0) = e_i \end{cases}$$

alors $(u_1, \ldots, u_n)$ *est une base de* $\mathcal{E}$*.*

Démonstration. Vérifions d'abord que $\mathcal{E}$ est un sous-espace vectoriel de $C^1(I, \mathbb{R}^n)$· C'en est bien un sous-ensemble, car si u est solution globale, u est dérivable donc continue, et $u'(t) = A(t)u(t)$ est un produit (de matrices certes) de fonctions continues, donc continue. C'est un sous-ensemble non-vide, par exemple car la fonction nulle est toujours solution. Enfin, il est stable par combinaison linéaire car, si $c \in \mathbb{R}$, u, $v \in \mathcal{E}$, on a

$$\begin{aligned} (cu + v)'(t) &= cu'(t) + v'(t) \\ &= cA(t)u(t) + A(t)v(t) \\ &= A(t)(cu(t) + v(t)) \end{aligned}$$

La famille $(u_1, \dots, u_n)$ est évidemment libre : si $\sum_{i=1}^n \lambda_i u_i \equiv 0$, alors en particulier, $\sum_{i=1}^n \lambda_i u_i(t_0) = \sum_{i=1}^n \lambda_i e_i = 0$, donc $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Enfin, montrons que cette famille est génératrice. Soit $u \in \mathcal{E}$, et notons $\lambda_1, \dots, \lambda_n$ les coordonnées de $u(t_0) \in \mathbb{R}^n$ dans la base $(e_1, \dots, e_n)$. Alors u et $\sum_{i=1}^n \lambda_i u_i$ sont solutions du même problème de Cauchy

$$\begin{cases} y' = A(t)y \\ y(t_0) = \sum_{i=1}^n \lambda_i e_i = u(t_0) \end{cases}$$

donc par unicité, $u = \sum_{i=1}^n \lambda_i u_i$. □

7.4 Cas d'une EDL non-homogène

Pour traiter une EDL non-homogène, l'idée est la même que pour une seule variable. L'observaton clé est la suivante. Si u_0 est une solution de $y' = A(t)y + b(t)$ alors les assertions suivantes sont équivalentes :

1. u est solution de $y' = A(t)y + b(t)$
2. $u - u_0$ est solution de l'EDL homogène $y' = A(t)y$.

Corollaire 7.4. *L'ensemble $\mathcal{A}$ des solutions globales de $y' = A(t)y + b(t)$ est un sous-espace affine de dimension n de $C^1(I, \mathbb{R}^n)$, dont la direction est donnée par l'espace vectoriel $\mathcal{E}$ des solutions globales de $y' = A(t)y$.*

Attention, en général on ne sait pas trouver de solutions particulières à une équation différentielle linéaire, mais on ne sait pas non plus résoudre une EDL homogène ! Dans la suite, on va donner une formule dans le cas très particulier des coefficients constants.

8 Exponentielle de matrices

8.1 Définition

On commence par l'outil qui permet de donner la formule des solutions.

Proposition 8.1. *Soit $A \in M_n(\mathbb{R})$. La série de terme général $\frac{A^k}{k!}$ converge dans $M_n(\mathbb{R})$.*

Définition 8.2. La somme de la série de la proposition est appelée l'*exponentielle* de A. On la note $\exp(A)$ ou e^A.

Démonstration. On se ramène à l'exponentielle classique, pour laquelle le résultat est supposé connu ! Notons

$$S_m := \sum_{k=0}^{m} \frac{A^k}{k!}$$

les sommes partielles, pour $m \in \mathbb{N}$. Pour $l \geq m$, on a

$$\begin{aligned}
|||S_l - S_m||| &= \left|\left|\left| \sum_{k=m+1}^{l} \frac{A^k}{k!} \right|\right|\right| \\
&\leq \sum_{k=m+1}^{l} \left|\left|\left| \frac{A^k}{k!} \right|\right|\right| && \text{par l'inégalité triangulaire} \\
&\leq \sum_{k=m+1}^{l} \frac{|||A|||^k}{k!} && \text{par sous-multiplicativité} \\
&\leq s_l - s_m
\end{aligned}$$

où $s_m := \sum_{k=0}^{m} \frac{|||A|||^k}{k!}$ est la somme partielle de la série qui définit $\exp(|||A|||)$ dans $\mathbb{R}$.

Comme cette dernière série converge, la suite des sommes partielles est de Cauchy. Donc $(S_m)_{m\in\mathbb{N}}$ est une suite de Cauchy dans $\mathrm{M}_n(\mathbb{R})$. Comme $\mathrm{M}_n(\mathbb{R})$ est un espace vectoriel de dimension finie, toute suite de Cauchy dans $\mathrm{M}_n(\mathbb{R})$ converge, donc la série converge. □

8.2 Calculs d'exponentielle de matrices

On peux parfois calculer directement l'exponentielle. Par exemple, si la matrice est nilpotente.

Exemple 8.3.

$$\begin{aligned}
\exp\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \cdots \\
&= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

En effet, pour tout $k \geq 2$, on a $A^k = 0$, donc la suite des sommes partielles est stationnaire.

Attention, il ne faut pas oublier le premier terme de la série, qui est toujours l'identité.

Si la matrice est diagonale, on peut calculer l'exponentielle aussi. En effet, si $A = \mathrm{diag}(a_1, \ldots, a_n)$, on a $A^k = \mathrm{diag}(a_1^k, \ldots, a_n^k)$ pour tout k, donc

$$\begin{aligned}
\exp(A) &= \sum_{k=0}^{+\infty} \mathrm{diag}(a_1^k/k!, \ldots, a_n^k/k!) \\
&= \mathrm{diag}(\sum_{k=0}^{+\infty} a_1^k/k!, \ldots, \sum_{k=0}^{+\infty} a_n^k/k!) \\
&= \mathrm{diag}(e^{a_1}, \ldots, e^{a_n})
\end{aligned}$$

Plus généralement, si une matrice est diagonalisable, on peut calculer son exponentielle, grâce à la proposition suivante.

Proposition 8.4. *Si $A \in M_n(\mathbb{R})$ et $P \in \mathrm{GL}_n(\mathbb{R})$, alors* $\exp(PAP^{-1}) = P\exp(A)P^{-1}$.

Démonstration. On vérifie d'abord, sur les sommes partielles,

$$\sum_{k=0}^{m} \frac{(PAP^{-1})^k}{k!} = \sum_{k=0}^{m} \frac{PA^kP^{-1}}{k!} = P\left(\sum_{k=0}^{m} \frac{A^k}{k!}\right)P^{-1}$$

Puis on passe à la limite sur les extrémités de l'égalité quand $m \to +\infty$ pour obtenir

$$\exp(PAP^{-1}) = P\exp(A)P^{-1}$$

Du côté gauche, c'est la définition, mais pour le côté droit, on utilise que la multiplication par une matrice fixée (ici, P ou P^{-1}) est une opération continue dans $\mathrm{M}_n(\mathbb{R})$. $\square$

Remarque 8.5. On peut aussi appliquer les résultats précédents aux matrices à coefficients complexes. Par conséquent, si une matrice réelle est diagonalisable dans $\mathbb{C}$ mais pas dans $\mathbb{R}$, on peut quand même calculer son exponentielle, qui sera une matrice réelle, en passant par une diagonalisation complexe.

8.3 Dérivée de l'exponentielle le long des droites

Proposition 8.6. *Soit $A \in M_n(\mathbb{R})$. L'application $\Phi : \mathbb{R} \to M_n(\mathbb{R}), t \mapsto e^{tA}$ est dérivable, et pour tout $t \in \mathbb{R}$, on a $\Phi'(t) = Ae^{tA}$.*

Démonstration. Par définition, Φ est la somme de la série de fonctions $\mathbb{R} \to \mathrm{M}_n(\mathbb{R})$ de terme général $\Phi_k : t \mapsto t^kA^k/k!$. Les fonctions Φ_k sont dérivables sur $\mathbb{R}$ (car polynomiales), avec pour tout $t \in \mathbb{R}$,

$$\Phi_k'(t) = \frac{t^{k-1}A^k}{(k-1)!}$$

si $k > 0$, et $\Phi_0' \equiv 0$.

On a

$$\begin{aligned}
\left|\left|\left|A\exp(tA) - \sum_{k=0}^{m} \Phi_k'(t)\right|\right|\right| &\leq \left|\left|\left|A\exp(tA) - A\sum_{k=0}^{m-1} \frac{t^kA^k}{k!}\right|\right|\right| + \left|\left|\left|A\sum_{k=0}^{m-1} \frac{t^kA^k}{k!} - \sum_{k=1}^{m} \frac{t^{k-1}A^k}{(k-1)!}\right|\right|\right| \\
&\leq \left|\left|\left|A\exp(tA) - A\sum_{k=0}^{m-1} \frac{t^kA^k}{k!}\right|\right|\right| \\
&\leq |||A||| \left|\left|\left|\sum_{k=m}^{+\infty} \frac{t^kA^k}{k!}\right|\right|\right| \\
&\leq |||A||| \sum_{k=m}^{+\infty} \frac{|t|^k|||A|||^k}{k!}
\end{aligned}$$

Pour $T > 0$ fixé, et $t \in [-T, T]$, c'est

$$\leq |||A||| \sum_{k=m}^{+\infty} \frac{T^k|||A|||^k}{k!}$$

qui est $|||A|||$ fois le reste de la série définissant $\exp(T|||A|||)$, qui tend vers 0 quand $m \to 0$.

On a ainsi montré que la série de terme général Φ'_k est uniformément convergente sur tout segment $[-T, T]$. Par théorème de dérivation terme à terme, la somme Φ de la série des Φ_k est dérivable, de dérivée

$$\Phi'(t) = \sum_{k=0}^{+\infty} \Phi'_k(t) = A \sum_{k=1}^{+\infty} \frac{t^{k-1} A^{k-1}}{(k-1)!} = Ae^{tA}$$

□

9 Solutions des EDL à coefficients constants

9.1 Solution d'une EDL homogène à coefficients constants

Le résultat suivant ramène la résolution d'une EDL à coefficients constants homogène au calcul d'une famille d'exponentielles de matrices. On a déjà vu quelques méthodes de calcul. On va voir ensuite comment appliquer ce théorème pour avoir encore plus de méthodes de calcul.

Théorème 9.1. *Soit $A \in M_n(\mathbb{R})$ et $x_0 \in \mathbb{R}^n$. L'unique solution globale au problème de Cauchy*

$$\begin{cases} y' = Ay \\ y(0) = x_0 \end{cases}$$

est la fonction $u : \mathbb{R} \to \mathbb{R}^n, t \mapsto e^{tA} x_0$.

Démonstration. Par le Théorème de Cauchy-Lipschitz linéaire, on sait qu'il existe une unique solution globale. Il suffit de vérifier que la fonction donnée dans l'énoncé est bien solution.

Pour la condition initiale, c'est immédiat, car l'exponentielle de la matrice nulle est l'identité, donc $u(0) = e^{0_n} x_0 = I_n x_0 = x_0$.

Pour l'équation différentielle, on vérifie très rapidement que la fonction est différentielle et on calcule sa dérivée de la manière suivante. On a

$$\begin{aligned} u'(t) &= \lim_{\varepsilon \to 0} \frac{u(t+\varepsilon) - u(t)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{e^{(t+\varepsilon)A} - e^{tA}}{\varepsilon} x_0 \end{aligned}$$

Ici, on a utilisé le fait que la multiplication d'une matrice par un vecteur est une opération continue (car linéaire) donc passe à la limite.

$$= \Phi'(t) x_0$$

où on note $\Phi(t) = e^{tA}$ comme dans la section précédente.

$$\begin{aligned} &= Ae^{tA} x_0 \\ &= Au(t) \end{aligned}$$

□

On note que dans le théorème précédent, on considère toujours des conditions initiales en $t_0 = 0$. En exercice, déterminer l'expression pour un t_0 quelconque, si besoin en examinant la preuve du corollaire suivant, ou la formule de Duhamel à venir.

9.2 Quelques applications théoriques

Corollaire 9.2. *Pour t, $s \in \mathbb{R}$, on a $e^{(t+s)A} = e^{tA}e^{sA}$.*

Bien qu'on puisse prouver ce résultat (et la généralisation quelques paragraphes plus loin) directement en termes de séries, on le démontre en appliquant l'unicité dans Cauchy-Lipschitz, pour illustrer les possibles applications théoriques.

Démonstration. On fixe $s \in \mathbb{R}$. Soit $x \in \mathbb{R}^n$. La fonction $u : \mathbb{R} \to \mathbb{R}^n$ définie par $u(t) = e^{(t+s)A}x$ vérifie $u'(t) = Au(t)$ et $u(0) = e^{sA}x$ (par la même preuve que dans la preuve précédente). Donc par unicité dans le Théorème précédent avec $x_0 = e^{sA}x$, on a, pour tout $t \in \mathbb{R}$, $u(t) = e^{tA}e^{sA}x$. C'est vrai pour tout $x \in \mathbb{R}^n$, donc c'est vrai pour les matrices : $e^{(t+s)A} = e^{tA}e^{sA}$. □

En particulier, si on applique cette formule avec $s = -t$, on obtient :

Corollaire 9.3. *Pour tout $A \in M_n(\mathbb{R})$, $\exp(A) \in \mathrm{GL}_n(\mathbb{R})$, et $(\exp(A))^{-1} = \exp(-A)$.*

Passons à la généralisation. Le résultat suivant est la généralisation naturelle, pour les matrices, de la propriété fondamentale $e^{a+b} = e^a e^b$ de l'exponentielle réelle.

Corollaire 9.4. *Soient A et $B \in M_n(\mathbb{R})$ deux matrices telles que $AB = BA$. Alors $\exp(A + B) = \exp(A)\exp(B) = \exp(B)\exp(A$.*

L'hypothèse de commutativité est la plus pratique, mais pas la plus générales : des matrices A et B qui ne commutent pas peuvent satisfaire $e^{A+B} = e^A e^B$. Cependant, la formule n'est pas vraie pour toute paire de matrices.

Démonstration. On va encore utiliser l'unicité dans Cauchy-Lipshitz. Soit $x \in \mathbb{R}^n$, et $u : \mathbb{R} \to \mathbb{R}^n$ la fonction définie par $u(t) = e^{tA}e^{tB}x$. En imitant la dérivée d'un produit de fonctions réelles (si vous n'avez pas déjà vu ce genre de résultat), on vérifie facilement que

$$u't) = Ae^{tA}e^{tB}x + e^{tA}Be^{tB}x$$

Comme le produit de matrice n'est pas commutatif, il faut vérifier qu'on a le droit de faire passer le facteur B à gauche dans le dernier terme. Pour ça, on observe que c'est immédiat sur les sommes partielles définissant e^{tA}, car A et B commutent :

$$\left(\sum_{k=0}^{m} \frac{t^k A^k}{k!}\right) B = B \left(\sum_{k=0}^{m} \frac{t^k A^k}{k!}\right)$$

En passant à la limite quand $m \to \infty$, on a bien

$$e^{tA}B = Be^{tA}$$

d'où

$$u'(t) = (A + B)e^{tA}e^{tB}x = (A + B)u(t)$$

De plus, $u(0) = I_n I_n x = x$, donc u est solution globale du problème de Cauchy

$$\begin{cases} y' = (A + B)y \\ y(0) = x \end{cases}$$

Par unicité,

$$e^{tA}e^{tB}x = u(t) = e^{t(A+B)}x$$

Comme c'est vrai pour tout $x \in \mathbb{R}^n$, on a bien

$$e^{tA}e^{tB} = e^{t(A+B)}$$

donc le résultat, en prenant $t = 1$. □

Exemple 9.5. Calculons l'exponentielle de tA avec $A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$ pour a et $b \in \mathbb{R}$. Pour cela, on remarque que

$$tA = \begin{bmatrix} ta & 0 \\ 0 & ta \end{bmatrix} + \begin{bmatrix} 0 & tb \\ 0 & 0 \end{bmatrix}$$

et que ces deux matrices commutent (en fait taI_2 commute avec n'importe quelle matrice). Alors

$$\begin{aligned}
\exp(tA) &= \exp\left(\begin{bmatrix} ta & 0 \\ 0 & ta \end{bmatrix} + \begin{bmatrix} 0 & tb \\ 0 & 0 \end{bmatrix}\right) \\
&= \exp\left(\begin{bmatrix} ta & 0 \\ 0 & ta \end{bmatrix}\right) \exp\left(\begin{bmatrix} 0 & tb \\ 0 & 0 \end{bmatrix}\right) \\
&= \begin{bmatrix} e^{ta} & 0 \\ 0 & e^{ta} \end{bmatrix} \begin{bmatrix} 1 & tb \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} e^{ta} & tbe^{ta} \\ 0 & e^{ta} \end{bmatrix}
\end{aligned}$$

Remarque 9.6. Si A est une matrice 2×2 non diagonalisable sur $\mathbb{C}$, alors A est conjuguée à une matrice comme ci-dessus. Donc on a vu dans ce cours comment calculer l'exponentielle de n'importe quelle matrice 2×2.

Plus généralement, par réduction de Jordan complexe, et en utilisant les propriétés de l'exponentielle prouvées ici, on peux théoriquement calculer l'exponentielle de n'importe quelle matrice réelle.

9.3 Cas des EDL non-homogènes

En maintenant l'hypothèse des coefficients constants (c'est-à-dire que la fonction à valeurs matricielles est constantes), mais en considérant un second membre qui peux être non nul et non constant, on a toujours une formule pour les solutions. Cette formule est essentiellement la même que dans le cas des équations différentielles linéaires d'ordre un en une variable, elle fait apparaitre une intégrale (ou une primitive) qu'on ne peux en général pas exprimer de manière simple.

Théorème 9.7 (Formule de Duhamel)**.** *Soit $A \in M_n(\mathbb{R})$, et $b : I \to \mathbb{R}^n$ une fonction continue. Soit $t_0 \in I$ et $x_0 \in \mathbb{R}^n$. Alors l'unique solution globale au problème de Cauchy*

$$\begin{cases} y' = Ay + b(t) \\ y(t_0) = x_0 \end{cases}$$

est la fonction définie par

$$u(t) = e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-s)A}b(s)\,ds\,.$$

Démonstration. Par le Théorème de Cauchy-Lipschitz linéaire, on sait qu'il existe une unique solution globale. Il suffit de vérifier que la fonction de l'énoncé convient. La condition initiale est immédiate : $u(t_0) = I_n x_0 + 0 = x_0$. Pour calculer la dérivée, on écrit

$$u(t) = e^{tA}\left(e^{-t_0A}x_0 + \int_{t_0}^{t} e^{-sA}b(s)\,ds\right)$$

puis on dérive ce produit :

$$\begin{aligned} u'(t) &= Ae^{tA}\left(e^{-t_0A}x_0 + \int_{t_0}^{t} e^{-sA}b(s)\,ds\right) + e^{tA}(0 + e^{-tA}b(t)) \\ &= Au(t) + b(t) \end{aligned}$$

□

10 Retour sur la sortie de tout compact

10.1 IAF pour les fonctions à valeurs vectorielles

On repasse au cadre où E est un espace vectoriel réel de dimension finie, toujours avec une norme fixée $\|\cdot\|$. On va revenir à la preuve du (corollaire du) Théorème de sortie de tout compact. Pour ça, on a besoin d'un outil : l'inégalité des accroissements finis pour les fonctions d'une variable réelle, à valeurs vectorielles.

Théorème 10.1 (IAF $2/n$)**.** *Soit $u : J \to E$ une fonction dérivable sur un inervalle ouvert J de $\mathbb{R}$. On suppose qu'il existe un $M \geq 0$ tel que, pour tout $t \in J$, $\|u'(t)\| \leq M$. Alors*

$$\forall x, y \in J, \qquad \|u(x) - u(y)\| \leq M|x - y|$$

Remarque 10.2. La meilleure valeur possible de M est $\sup_{t\in J}\|u'(t)\|$. Si on change de norme sur E, les valeurs possibles de M changent, mais pas l'existence d'un tel M.

Démonstration. On suppose $x < y$ (car les rôles sont symétriques, et l'inégalité est triviale si $x = y$). Soit $\varepsilon > 0$. On considère l'ensemble

$$F = \{t \in [x, y] \mid \forall s \in [x, t], \|u(s) - u(x)\| \leq (M + \varepsilon)(s - x)\}$$

C'est une partie de $\mathbb{R}$ qui est non-vide (car $x \in F$), et majorée (par y). Notons $c = \sup F$.

On montre d'abord que $F = [x, c]$. Il suffit de montrer que $c \in F$, étant donné la définition de F. D'abord, par caractérisation séquentielle de la borne sup, il existe une suite (t_n) à valeurs dans F, qui converge vers c. En particulier, pour tout n, comme $t_n \in F$, on a

$$\|u(t_n) - u(x)\| \leq (M + \varepsilon)(t_n - x)$$

et en passant à la limite, par continuité,

$$\|u(c) - u(x)\| \leq (M + \varepsilon)(c - x)$$

Ça ne suffit pas à affirmer que $c \in F$ puisqu'il faut montrer cette inégalité pour tout $t \in [x, c]$. Pour $t \in [x, c[$, comme $t_n \to c$, il existe un $N \in \mathbb{N}$ tel que $t \in [x, t_n]$. Or $t_n \in F$, donc on a bien

$$\|u(t) - u(x)\| \leq (M + \varepsilon)(t - x)$$

Supposons maintenant par l'absurde que $c < y$. Par définition de la dérivée, on a

$$\lim_{h \to 0} \frac{u(c+h) - u(c)}{h} = u'(c)$$

En particulier, en passant à la norme et en considérant uniquement les $h > 0$, on a

$$\lim_{h \to 0} \frac{\|u(c+h) - u(c)\|}{h} = \|u'(c)\| \leq M$$

Donc, il existe $\delta > 0$ tel que pour $h \in]0, \delta[$,

$$\frac{\|u(c+h) - u(c)\|}{h} \leq M + \varepsilon$$

autrement dit,

$$\|u(c+h) - u(c)\| \leq (M + \varepsilon)h$$

Par inégalité triangulaire, on a

$$\begin{aligned} \|u(c+h) - u(x)\| &\leq \|u(c+h) - u(c)\| + \|u(c) - u(x)\| \\ &\leq (M + \varepsilon)h + (M + \varepsilon)(c - x) \\ &\leq (M + \varepsilon)(c + h - x) \end{aligned}$$

On a alors par exemple $c + \frac{\delta}{2} \in F$, ce qui contredit $c = \sup F$.

On a montré par l'absurde que $c = y$, et en particulier, que

$$\|u(y) - u(x)\| \leq (M + \varepsilon)(y - x)$$

On l'a montré pour un $\varepsilon > 0$ quelconque, ce qui permet de passer à la limite quand $\varepsilon \to 0$, et on obtient bien

$$\|u(y) - u(x)\| \leq M(y - x)$$

□

Remarque 10.3. Le principe de la preuve est très similaire à celui de la preuve de l'unicité globale dans Cauchy-Lipschitz.

10.2 Retour sur la sortie des compacts

On rappelle maintenant l'énoncé du corollaire du Théorème de sortie de tout compact, que l'on souhaite démontrer indépendamment de ce théorème.

Proposition 10.4. *On suppose que la fonction $f :]a, b[\times U \subset \mathbb{R} \times E \to E$ est continue, et localement, Lipschitzienne en espace. Soit $u :]c, d[\to E$ une solution maximale de l'équation différentielle $y' = f(t, y)$. S'il existe un compact K de E, inclus dans U, tel que pour tout $t \in]c, d[$, $u(t) \in K$, alors $]c, d[=]a, b[$.*

Démonstration. On raisonne par l'absurde : on suppose par exemple $d < b$, et qu'il existe un compact K tel que pour tout $t \in]c, d[$, $u(t) \in K$. Soit (t_n) une suite strictement croissante dans $]c, d[$, qui converge vers d. Par compacité, quitte à remplacer (t_n) par une sous-suite, on peut supposer que la suite $(u(t_n))$ converge vers un $x_0 \in K$. Montrons qu'on a, en fait, $\lim_{t \to d} u(t) = x_0$.

Comme $[t_0, d] \times K$ est un compact, et f est continue sur ce compact, elle est bornée. Notons $M \geq 0$ une borne : pour tout $(t, x) \in [t_0, d] \times K$, $\|f(t, x)\| \leq M$. Comme $u'(t) = f(t, u(t))$ et $u(t) \in K$, on en déduit que, pour $t \in [t_0, d[$, $\|u'(t)\| \leq M$. Donc, par inégalité des accroissements finis, u est M-Lipschitzienne.

Soit $\varepsilon > 0$. Par convergence des suites, il existe un $N \in \mathbb{N}$ tel que $|t_N - d| < \frac{\varepsilon}{2M}$ et $\|u(t_N) - x_0\| < \frac{\varepsilon}{2}$. Alors pour $t \in [t_N, d]$, on a

$$\begin{aligned} \|u(t) - x_0\| &\leq \|u(t) - u(t_N)\| + \|u(t_N) - x_0\| \\ &\leq M|t - t_N| + \frac{\varepsilon}{2} \\ &\leq \varepsilon \end{aligned}$$

On a bien montré que $\lim_{t \to d} u(t) = x_0$.

Pour prolonger u, on considère la solution maximale $v :]\alpha, \beta[\to E$ au problème de Cauchy

$$\begin{cases} y' = f(t, y) \\ y(d) = x_0 \end{cases}$$

et la fonction $w :]c, \beta[\to E$ définie par

$$w(t) = \begin{cases} u(t) \text{ si } t < d \\ v(t) \text{ si } t \geq d \end{cases}$$

D'après l'étude précédente, w est continue en d. De plus, w est dérivable sur $]c, d[\cup]d, \beta[$, et

$$\lim_{\substack{t \to d \\ t < d}} w'(t) = \lim_{\substack{t \to d \\ t < d}} u'(t) = f(d, x_0) = \lim_{\substack{t \to d \\ t > d}} v'(t) = \lim_{\substack{t \to d \\ t > d}} w'(t)$$

donc par prolongement C^1, w est dérivable sur $]c, \beta[$ et satisfait l'équation différentielle partout. C'est un prolongement strict de la solution maximale u, une contradiction. □

10.3 Preuve du Théorème de sortie de tout compact (bonus)

Dans cette partie, comme ailleurs dans le cours, on note $B(x, r)$ la boule ouverte de centre x et de rayon r dans E pour la norme $\|\cdot\|$.

On va maintenant démontrer le Théorème de sortie de tout compact, en utilisant la Proposition précédente, qu'on avait initialement présenté comme un corollaire. Le début de la preuve est similaire : on suppose par l'absurde que $d \in]a, b[$, et qu'il existe un compact K tel que, pour tout $\delta \in]c, d[$, il existe $t \in]\delta, d[$ avec $u(t) \in K$.

De ceci on construit d'abord par récurrence une suite strictement croissante (t_n), à valeurs dans $]c, d[$, qui converge vers d, telle que pour tout n, $u(t_n) \in K$. Pour t_0, on choisit un élément quelconque de $]c, d[$ tel que $u(t_0) \in K$. Si t_n est construit, on choisit un t_{n+1} donné par l'hypothèse ci-dessus, avec $\delta = \max(t_n, d - 1/(n+1))$, de sorte qu'on ait bien $u(t_{n+1}) \in K$. Alors on vérifie facilement que le premier terme du max assure que (t_n) est strictement croissante, et que le second terme du max assure que (t_n) converge vers d.

Quitte à prendre une sous-suite, on peux supposer, par compacité de K, que $u(t_n)$ converge vers un $x_0 \in K$.

Comme f est localement Lipschitzienne, il existe $\varepsilon > 0$ et $\lambda \geq 0$ tels que, si t, x_1, x_2, vérifient $|t - d| < \varepsilon$, $\|x_i - x_0\| < \varepsilon$, alors

$$\|f(t, x_1) - f(t, x_2)\| \leq \lambda \|x_1 - x_2\|$$

Posons

$$\eta = \min\left(\varepsilon, \frac{\varepsilon}{2(\|f(d, x_0)\| + \lambda\varepsilon)}\right)$$

Considérons la restriction $\tilde{f} :]d - \eta, d + \eta[\times B(x_0, \varepsilon) \to E$ de f, qui est (globalement) λ-Lipschitzienne en espace. Montrons que pour tout $(\tilde{t}, \tilde{x}) \in]d - \eta, d + \eta[\times B(x_0, \frac{\varepsilon}{2})$, la solution maximale au problème de Cauchy

$$\begin{cases} y' = \tilde{f}(t, y) \\ y(\tilde{t}) = \tilde{x} \end{cases}$$

est globale, c'est-à-dire définie sur $]d - \eta, d + \eta[$.

Considérons pour cela la solution maximale $\tilde{u} :]\tilde{c}, \tilde{d}[\to E$ à ce problème de Cauchy. On a, pour tout t,

$$\begin{aligned} \|\tilde{u}'(t)\| &= \|f(t, \tilde{u}(t))\| \\ &= \|f(t, \tilde{u}(t)) - f(d, x_0) + f(d, x_0)\| \\ &\leq \|f(t, \tilde{u}(t)) - f(d, x_0)\| + \|f(d, x_0)\| \\ &\leq \lambda \|\tilde{u}(t) - x_0\| + \|f(d, x_0)\| \end{aligned}$$

comme $\tilde{u}(t) \in B(x_0, \varepsilon)$, c'est

$$\leq \lambda\varepsilon + \|f(d, x_0)\|$$

Appliquons maintenant l'inégalité des accroissements finis à $\tilde{u}$, pour obtenir, pour tout $t \in]\tilde{c}, \tilde{d}[$,

$$\begin{aligned}
\|\tilde{u}(t) - \tilde{u}(\tilde{t})\| &\leq (\lambda\varepsilon + \|f(d, x_0)\|)\,|t - \tilde{t}| \\
&\leq (\lambda\varepsilon + \|f(d, x_0)\|)\,\eta \\
&\leq (\lambda\varepsilon + \|f(d, x_0)\|)\,\frac{\varepsilon}{2(\|f(d, x_0)\| + \lambda\varepsilon)} \\
&\leq \frac{\varepsilon}{2}
\end{aligned}$$

Par conséquent, $\tilde{u}$ prends ses valeurs dans la boule fermée de centre $\tilde{x}$ et de rayon $\frac{\varepsilon}{2}$, qui est un compact de E, inclus dans l'ouvert $B(x_0, \varepsilon)$ car $\tilde{x} \in B(x_0, \frac{\varepsilon}{2})$. Par la proposition prouvée à la sous-section précédente, $\tilde{u}$ est bien globale, définie sur $]d - \eta, d + \eta[$.

On applique maintenant ceci à u, restreinte à un voisinage de $\tilde{t} = t_N$ pour N assez grand, et $\tilde{x} = u(t_N)$. Par unicité du prolongement en une solution maximale, on en déduit que u se prolonge sur $]c, d + \eta[$, une contradiction.

11 Dérivées directionnelles

11.1 Définition

On a beaucoup utilisé l'inégalité des accroissements finis depuis le début du cours et en TD, par exemple pour montrer qu'une fonction est localement Lipschitzienne. On vient de généraliser ce résultat, initialement valable pour les fonctions à valeurs réelles d'une variable réelle, aux fonctions à valeurs vectorielles d'une variable réelle. Mais on aimerait mieux : un tel résultat pour les fonctions de plusieurs variables, comme par exemple la fonction $f : I \times U \to E$ qui définit une équation différentielle. La question maintenant est de savoir quelles information infinitésimale va permettre de faire ça.

Vous avez déjà vu dans d'autres cours (HAX404X pour les étudiants de l'UM) la notion de dérivée directionnelle. Nous allons revenir dessus pour voir comment cette notion permet de généraliser l'inégalité des accroissements finis, mais aussi relever les difficultés liées à cette notion qui rendent beaucoup plus pratique la notion que nous introduiront ensuite : la différentielle.

Notre cadre évolue un peu, par rapport aux sections sur les équations différentielles. On considère deux espaces vectoriels réels E et F de dimension finie, U un ouvert de E, et $f : U \to F$ une fonction.

Définition 11.1. Pour $a \in U$ et $v \in E$ un vecteur, la *dérivée directionnelle de f en a selon v* est la limite

$$\lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} \in F$$

si elle existe. Dans ce cas, on la note $\partial_v f(a)$.

Si la limite n'existe pas, on dit simplement que f n'admet pas de dérivée directionnelle en a selon le vecteur v.

Dans le cas où $E = \mathbb{R}^p$, si on note $(e_1, \ldots, e_p)$ la base canonique, et $(x_1, \ldots, x_p)$ les coordonnées dans cette base, on utilise souvent la notation

$$\frac{\partial f}{\partial x_i}(a) = \partial_{e_i} f(a)$$

et on appelle ces p dérivées directionnelles particulières les *dérivées partielles* de f.

Si $E = \mathbb{R}^p$ et $F = \mathbb{R}$, on appelle *gradient* de f en a, et on note $\nabla f(a)$ le vecteur

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_i}(a), \dots, \frac{\partial f}{\partial x_i}(a) \right)$$

(lorsque toutes les dérivées partielles existent).

11.2 Lien entre dérivée directionnelle et dérivée

Étant donné $a \in U$ et $v \in E$, on considère la fonction $g : I \to F$ définie par $I = \{t \in \mathbb{R} \mid a + tv \in U\}$ et $g(t) = f(a + tv)$.

Proposition 11.2. *L'ensemble I est un ouvert de $\mathbb{R}$, g est dérivable en t si et seulement si $\partial_v f(a + tv)$ existe, et dans ce cas, $g'(t) = \partial_v f(a + tv)$.*

Démonstration. Pour montrer que I est ouvert, on note que $I = \ell^{-1}(U)$ où $\ell : \mathbb{R} \to E, t \mapsto a + tv$. Cette application est une application affine entre espaces de dimension finie, donc elle est continue. Plus explicitement, si $t_0 \in \mathbb{R}$, on a $\|L(t) - L(t_0)\|_F = \|(t - t_0)v\|_F = |t - t_0|\|v\|_F$ qui tend vers 0 lorsque t tend vers t_0. Comme l'image réciproque d'un ouvert par une application continue est un ouvert, on obtient en effet que I est ouvert.

Pour $t \in I$, g est dérivable en t si et seulement si la limite suivante existe :

$$\lim_{\varepsilon \to 0} \frac{g(t + \varepsilon) - g(t)}{\varepsilon}$$

or, on a

$$\frac{g(t + \varepsilon) - g(t)}{\varepsilon} = \frac{f(a + tv + \varepsilon v) - f(a + tv)}{\varepsilon}$$

et f admet une dérivée directionnelle en $a + tv$ selon le vecteur v si et seulement si la même limite existe :

$$\lim_{\varepsilon \to 0} \frac{f(a + tv + \varepsilon v) - f(a + tv)}{\varepsilon}$$

Enfin, comme les limites sont les mêmes si elles existe, on a bien $g'(t) = \partial_v f(a + tv)$. □

11.3 IAF par les dérivées directionnelles

On rappelle que, dans un espace vectoriel E, si a et $b \in E$, le segment $[a, b]$ désigne l'ensemble

$$[a, b] = \{(1 - t)a + tb \mid t \in [0, 1]\}$$

Théorème 11.3 (IAF par les dérivées directionnelles)**.** *Soient $a \neq b \in U$ tels que $[a, b] \subset U$. On suppose que, pour tout $c \in [a, b]$, la dérivée directionnelle de f en c selon $b - a$ existe. On suppose de plus qu'il existe $M \in \mathbb{R}$ tel que, pour tout $c \in [a, b]$, $\|\partial_{b-a} f(c)\|_F \leq M$. Alors*

$$\|f(b) - f(a)\|_F \leq M$$

Démonstration. On considère la fonction $g : I \to F, t \mapsto f(a + t(b - a))$ où $I = \{t \mid a + tv \in U\}$, comme dans la proposition précédente. L'hypothèse montre que g est dérivable en tout point de $[0, 1]$, et que sa dérivée vérifie $\|g'(t)\|_F \leq M$ pour $t \in [0, 1]$. En appliquant l'IAF pour les fonctions à valeurs vectorielles, on obtient

$$f(b) - f(a) = g(1) - g(0) \leq M(1 - 0) = M$$

□

Cet énoncé n'est pas très pratique pour plusieurs raisons :

1. Pourquoi est-ce que $\|b - a\|_E$ n'apparaît pas ?
2. Si on voulait montrer, par exemple, que f est Lipschitzienne sur U, il faudrait pouvoir considérer chacune des restrictions possibles à une droite (il y en a beaucoup...), et vérifier à chaque fois que la restriction est dérivable, de dérivée bornée.

Le deuxième point sera en partie résolu en considérant la notion plus forte de différentiabilité, qui permettra, par exemple, de vérifier qu'une fonction est Lipschitzienne en contrôlant simplement ses dérivées partielles (seulement si elle est différentiable !). Illustrons d'abord à quel point les dérivées directionnelles peuvent dépendre du vecteur v considéré.

Exemple 11.4. On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de l'Exercice 34 du TD. Elle est définie par $f(x, y) = \frac{y^2}{x}$ si $x \neq 0$ et $f(0, y) = 0$. Par un calcul direct, on vérifie que $\partial_{(v_1,v_2)} f(0, 0) = \frac{v_2^2}{v_1}$ si $v_1 \neq 0$, et $\partial_{(v_1,v_2)} f(0, 0) = 0$ sinon. On constate que la fonction $v \mapsto \partial_v f(a)$ peut être très compliquée, et on imagine sans mal comment construire d'autres exemples arbitrairement compliqués.

Pour revenir au premier point, il est facile à éclaircir grâce à la proposition suivante.

Proposition 11.5. *Si $\partial_v f(a)$ existe, et $\lambda \in \mathbb{R}$, alors*

$$\partial_{\lambda v} f(a) = \lambda \partial_v f(a)$$

Démonstration. Il suffit de remarquer l'égalité

$$\frac{f(a + t\lambda v) - f(a)}{t} = \lambda \frac{f(a + t\lambda v) - f(a)}{t\lambda}$$

et de passer à la limite quand $t \to 0$ des deux côtés. □

Enfin, un autre défaut des dérivées directionnelles tient dans l'observation suivante : il existe des fonctions qui admettent des dérivées directionnelles selon tous vecteurs en un point, mais qui ne sont pas continues en ce point. L'exemple de l'exercice 34 illustre ce phénomène. En effet, on a par exemple $\lim_{x \to 0} f(x^2, x) = 1 \neq 0 = f(0, 0)$. C'est en contraste avec le cas des fonctions dérivables : une fonction dérivable en a est nécessairement continue a. Là encore, on va voir que la notion de différentiabilité se comporte mieux.

12 Différentiabilité

12.1 Définition

Plutôt que d'imiter la définition de dérivée comme limite d'un taux d'accroissement (qui n'a aucun sens sans se restreindre à une droite), on imite la caractérisation par développement limité pour trouver la notion de différentiabilité. Pour une variable réelle, $f : I \subset \mathbb{R} \to F$ est dérivable en $a \in I$ si et seulement si il existe un $p \in F$ tel que, près de a,

$$f(x) = f(a) + (x - a)p + o(x - a)$$

et dans ce cas, $p = f'(a)$.

Définition 12.1. La fonction f est *différentiable en a* s'il existe une *application linéaire* $df(a) : E \to F$ telle que

$$f(x) = f(a) + df(a)(x - a) + o(x - a)$$

quand x tends vers a. Cette application linéaire $df(a)$ est appelée la *différentielle de f en a*. La fonction $E \to F, x \mapsto f(a) + df(a)(x - a)$ est appelée la *fonction tangente à f en a*.

On dit que f est *différentiable sur U* si pour tout $a \in U$, f différentiable en a.

Remarque 12.2. Il est important de noter dans quel espace vit $df(a)$. On note, pour tout le cours, $L(E, F)$ l'espace vectoriel des applications linéaires de E dans F. On a donc $df(a) \in L(E, F)$.

Remarque 12.3. Que signifie le $o(x - a)$? Cela signifie *négligeable devant $x - a$ quand x tend vers a*. On peut le caractériser de différentes manières en utilisant des normes (mais la notion est indépendante de la norme utilisée, puisque toutes les normes sont équivalentes en dimension finie). Premièrement,

$$f(x) - f(a) - df(a)(x - a) = o(x - a)$$

si et seulement si il existe une fonction $\varepsilon : U \to F$ telle que

$$f(x) - f(a) - df(a)(x - a) = \|x - a\|\varepsilon(x)$$

et

$$\lim_{x \to a} \varepsilon(x) = 0$$

C'est une manière sûre d'écrire les choses, dès lors qu'on doit faire des manipulations sur les développements. Si on veut simplement montrer la différentiabilité, on peut souvent utiliser l'équivalence avec :

$$\lim_{x \to a} \frac{f(x) - f(a) - df(a)(x - a)}{\|x - a\|_E} = 0 \qquad \text{dans } F$$

qui, souvent, se démontre en passant par

$$\lim_{x \to a} \frac{\|f(x) - f(a) - df(a)(x - a)\|_F}{\|x - a\|_E} = 0 \qquad \text{dans } \mathbb{R}$$

Enfin, on notera souvent $x - a = h$ pour écrire des développements. Si on écrit alors le reste sous la forme $\|h\|_E\varepsilon(h)$, il faut être conscient que la fonction ε n'est plus définie sur U mais sur le translaté $-a + U$. Ce n'est généralement pas un problème.

Remarque 12.4. On fait un abus de notation ici : pour l'instant, on n'a pas démontré que, si elle existe, la différentielle de f en a est uniquement déterminée.

12.2 Différentielle et dérivées directionnelles

Proposition 12.5. *Soit f une fonction différentiable en a. Pour tout vecteur v, f admet une dérivée directionnelle en a selon v, et*

$$\partial_v f(a) = df(a)(v)$$

Démonstration. On suppose f différentiable en a. Soit $\varepsilon : U \to F$ la fonction telle que

$$f(x) = f(a) + df(a)(x-a) + \|x-a\|\varepsilon(x) \qquad \text{et} \qquad \lim_{x\to a} \varepsilon(x) = 0$$

Soit $v \in E$. On a

$$\begin{aligned} \frac{f(a+tv)-f(a)}{t} &= \frac{df(a)(tv) + \|tv\|_E \varepsilon(a+tv)}{t} \\ &= df(a)(v) + \frac{|t|}{t}\|v\|_E \varepsilon(a+tv) \end{aligned}$$

Comme $\frac{|t|}{t}\|v\|_E = \pm\|v\|_E$ est borné, et $\lim_{t\to 0}\varepsilon(a+tv) = 0$, on a bien $\partial_v f(a)$ existe, et

$$\partial_v f(a) = \lim_{t\to 0} \frac{f(a+tv)-f(a)}{t} = df(a)(v)$$

□

Définition 12.6. Soit $f : U \subset \mathbb{R}^p \to \mathbb{R}^q, (x_1,\ldots,x_p) \mapsto (f_1(x_1,\ldots,x_p),\ldots,f_q(x_1,\ldots,x_p))$ une fonction différentiable en a. La matrice *Jacobienne* de f en a, notée $J_f(a)$, est la matrice de taille $q \times p$ dont le coefficient à la ligne k et colonne l est le nombre $\frac{\partial f_k}{\partial x_l}(a) := df_k(a)(0,\ldots,0,1,0\ldots,0) = \partial_{(0,\ldots,0,1,0\ldots,0)} f_k(a)$ (où le 1 est en l-ème position). C'est la matrice qui représente l'application linéaire $df(a)$ dans les bases canoniques.

12.3 Le point clef : la linéarité

Comme $df(a)$ est linéaire, on peux retrouver toutes les valeurs de $df(a)$ (donc aussi les dérivées directionnelles par la proposition précédente) à partir de ses valeurs sur une base de E. Ceci résout le problème des dérivées directionnelles mentionné plus tôt.

Plus précisément, si $(e_1,\ldots,e_p)$ est une base de E, et $v = \sum_{i=1}^p v_i e_i$, alors

$$\partial_v f(a) = df(a)(v) = df(a)\left(\sum_{i=1}^p v_i e_i\right) = \sum_{i=1}^p v_i df(a)(e_i) = \sum_{i=1}^p v_i \partial_{e_i} f(a)$$

Si $E = \mathbb{R}^p$ avec sa base canonique, et qu'on note les coordonnées par $(x_1,\ldots,x_p)$, alors on peut exprimer la différentielle avec les dérivées partielles :

$$df(a)(v_1,\ldots,v_p) = \sum_{i=1}^p v_i \frac{\partial f}{\partial x_i}(a)$$

13 Différentiabilité, en pratique

13.1 Comment montrer qu'une fonction n'est pas différentiable ?

On rappelle que, d'après la section précédente, si f est différentiable en a, alors $df(a)(v) = \partial_v f(a)$. En particulier, s'il existe un vecteur v tel que $\partial_v f(a)$ n'existe pas, alors f n'est **pas différentiable** en a.

Exemple 13.1. Une fonction $f : U \subset \mathbb{R} \to F$ qui n'est pas dérivable en $a \in U$ n'est pas différentiable en a.

En général, si $E = \mathbb{R}^p$, une stratégie raisonnable est de commencer par vérifier si f admet des dérivées partielles selon toutes les variables.

De plus, l'application $df(a)$ est par définition linéaire. Par conséquent, si pour tout v, $\partial_v f(a)$ existe, mais l'application $v \mapsto \partial_v f(a)$ n'est pas linéaire, alors f n'est **pas différentiable** en a.

Exemple 13.2. On reprend l'exemple $f : \mathbb{R}^2 \to \mathbb{R}$ de l'Exercice 34 du TD, déjà utilisée plus tôt. On a vu que $\partial_{(v_1,v_2)} f(0,0) = \frac{v_2^2}{v_1}$ si $v_1 \neq 0$, et $\partial_{(v_1,v_2)} f(0,0) = 0$ sinon. Ce n'est pas linéaire en $v = (v_1, v_2)$. En effet, on a

$$\partial_{(1,1)} f(0,0) = 1 \neq 0 + 0 = \partial_{(1,0)} f(0,0) + \partial_{(0,1)} f(0,0)$$

Comme on l'a vu, la fonction de l'exemple précédent n'est même pas continue en $(0,0)$. C'est en fait la première chose à vérifier pour qu'une fonction puisse être différentiable.

Proposition 13.3. *Soit f une fonction différentiable en a, alors f est continue en a.*

Attention, la réciproque est fausse ! La fonction $f : x \mapsto |x|$ est continue en 0, mais pas dérivable en 0. D'après le premier exemple de cette section, elle n'est pas différentiable en 0.

Démonstration. Le principe est très simple : un développement limité à l'ordre 1 implique un développement limité à l'ordre 0, qui est équivalent à la continuité. Plus précisément, soit $\varepsilon : U \to F$ la fonction telle que

$$f(x) = f(a) + df(a)(x-a) + \|x-a\|\varepsilon(x) \qquad \text{et} \qquad \lim_{x \to a} \varepsilon(x) = 0$$

On a aussi $\lim_{x\to a} \|x-a\| = 0$, et $\lim_{x\to a} df(a)(x-a) = 0$ par continuité des applications linéaires (en dimension finie). En passant à la limite, on a donc bien $\lim_{x\to a} f(x) = f(a)$. □

On a utilisé dans la preuve la continuité des applications linéaires en dimension finie (et on l'a déjà utilisé plusieurs fois précédemment). Rappelons comment on peux démontrer ce résultat, puisqu'on va faire une preuve similaire bientôt. Soit $\ell : E \to F$ une application linéaire. On choisit une base $(e_1, \ldots, e_p)$ de E, et on considère la norme ∞ associée :

$$\|\sum_{i=1}^{p} h_i e_i\|_E := \sup\{|h_i| \mid 1 \leq i \leq p\}$$

On a, pour n'importe quelle norme sur F, et $a \in E$,

$$\begin{aligned}
\|l(x) - l(a)\|_F &= \|l(x-a)\|_F \\
&= \left\| l\left(\sum_{i=1}^{p} h_i e_i\right)\right\|_F \qquad \text{en notant } x - a = \sum_{i=1}^{p} h_i e_i \\
&= \left\| \sum_{i=1}^{p} h_i l(e_i) \right\|_F \\
&\leq \sum_{i=1}^{p} |h_i| \|l(e_i)\|_F \\
&\leq \sum_{i=1}^{p} \|x-a\|_E \|l(e_i)\|_F \\
&\leq \|x-a\|_E \sum_{i=1}^{p} \|l(e_i)\|_F
\end{aligned}$$

Comme $\sum_{i=1}^{p} \|l(e_i)\|_F$ est une constante, on a bien $\lim_{x \to a} f(x) = f(a)$.

13.2 Comment montrer qu'une fonction est différentiable, et calculer $df(a)$?

La méthode principale, avec les opérations qui viendront plus tard, est d'appliquer directement la définition. Pour cela, il faut *deviner* la différentielle, puis montrer que, en notant $df(a)$ pour le candidat, par abus de notation, que le reste

$$f(x) - f(a) - df(a)(x-a)$$

est négligeable devant $x - a$ directement. D'ailleurs, lorsque les méthodes données ci-dessus pour montrer qu'une fonction n'est pas différentiable ne suffisent pas, la dernière solution est habituellement de justement montrer que le reste n'est pas un $o(x-a)$.

Pour trouver le candidat, on peux d'abord essayer d'écrire directement un développement faisant apparaitre un premier terme constant (égal à $f(a)$), puis un terme linéaire en $(x-a)$, puis un reste sans partie affine.

Exemple 13.4. Soit $f : E \to F, x \mapsto y + l(x)$ une fonction affine (c'est-à-dire que $y \in F$ est une constante, et $l : E \to F$ est une application linéaire). Alors f est différentiable sur E, et pour tout $a \in E$, $df(a) = l$.

En effet, si on essaie de faire un développement, on commence par faire apparaître le $f(a)$:

$$\begin{aligned}
f(x) &= f(a) + f(x) - f(a) \\
&= f(a) + l(x) - l(a) \\
&= f(a) + l(x-a) + 0
\end{aligned}$$

Comme l est linéaire, et $0 = o(x-a)$, on a bien montré que f est différentiable en a et $df(a) = l$.

Pour deviner la différentielle, on peux passer par les dérivées partielles.

Proposition 13.5 (différentiable implique dérivable, **en une variable**). *Soit $f : U \subset \mathbb{R} \to F$, et $a \in U$. Alors f différentiable en a si et seulement si f est dérivable en a. Dans ce cas, on a $df(a)(1) = f'(a)$.*

Démonstration. On a déjà montré que différentiable implique dérivable dans ce cadre (au début de la première section), on veut montrer la réciproque. Supposons f dérivable en a. Si f est différentiable en a, on sait que $df(a)(1) = \partial_1 f(a) = f'(a)$. La seule application linéaire de $\mathbb{R}$ dans F qui vérifie ça est l'application $t \mapsto tf'(a)$. On a donc notre candidat pour la différentielle. Il reste à vérifier que $f(x) - f(a) - (x-a)f'(a) = o(x-a)$. Pour ça, on écrit

$$\frac{f(x) - f(a) - (x-a)f'(a)}{|x-a|} = \frac{x-a}{|x-a|}\left(\frac{f(a+(x-a)) - f(a)}{x-a} - f'(a)\right)$$

Comme $\frac{x-a}{|x-a|} = \pm 1$ est borné, par définition de la dérivabilité on a bien que le quotient ci-dessus tend vers 0 quand x tend vers a. □

Lorsque $E = \mathbb{R}^p$ et $F = \mathbb{R}^q$, munis des bases canoniques, le lien entre dérivées partielles et différentielles est rendu plus clair par la notion de matrice jacobienne. On note $(x_1, \dots, x_p)$ les coordonnées sur $\mathbb{R}^p$, et $f = (f_1, \dots, f_q)$ pour les composantes de f.

Définition 13.6. Si f est différentiable en a, la *matrice jacobienne* de f en a, notée $J_f(a)$, est la matrice qui représente l'application linéaire $df(a)$ dans les bases canoniques.

Remarque 13.7. C'est une matrice de taille $q \times p$, dont les coefficients sont les

$$\left(\frac{\partial f_i}{\partial x_j}(a)\right)_{1 \leq i \leq q,\ 1 \leq j \leq p}$$

Il faut se rappeler que les lignes correspondent aux composantes de f (il y en a autant que la dimension de l'espace d'arrivée $\mathbb{R}^q$), et les colonnes aux variables de f (il y en a autant que la dimension de l'espace de départ $\mathbb{R}^p$).

On peut utiliser la jacobienne pour deviner le candidat pour la différentielle : on calcule d'abord les dérivées partielles de f, donc le candidat pour $J_f(a)$, et on utilise la relation

$$df(a)(v) = J_f(a)v$$

où le terme de droite est un produit matriciel, où v est interprété comme un vecteur colonne.

14 Opérations sur les différentielles

14.1 Opérations élémentaires

On énonce d'abord deux propriétés simples, dont on écrit pas la démonstration qui est directe, mais qu'on utilisera régulièrement dans la suite.

Proposition 14.1 (Différentiabilité par composantes)**.** *Une fonction* $f : U \to \mathbb{R}^q, x \mapsto (f_1(x), \ldots, f_q(x))$ *est différentiable en* $a \in U$ *si et seulement si chacune des fonctions* f_i *l'est. Dans ce cas,* $df(a)(v) = (df_1(a)(v), \ldots, df_q(q)(v))$.

Proposition 14.2 (Linéarité)**.** *Soit* $\lambda \in \mathbb{R}$, $f : U \subset E \to F$ *et* $g : U \subset E \to F$ *deux fonctions différentiables en* $a \in U$*. Alors* $\lambda f + g$ *est différentiable en* a*, et* $d(\lambda f + g)(a) = \lambda df(a) + dg(a)$.

14.2 Différentielle d'une composée

Théorème 14.3. *Soit* $f : U \subset E \to F$ *une fonction différentiable en* a*,* $g : V \subset F \to G$ *(où* G *est un espace vectoriel réel de dimension finie, et* V *un ouvert de* F*), avec* $f(U) \subset V$ *et* g *différentiable en* $f(a)$*. Alors* $g \circ f$ *est différentiable en* a*, et* $d(g \circ f)(a) = dg(f(a)) \circ df(a)$.

Il faut garder en tête les espaces dans lesquels vivent chaque objet pour se rappeler l'ordre des ingrédients. On a $df(a) \in L(E, F)$, $dg(f(a)) \in L(F, G)$, et $d(g \circ f)(a) \in L(E, G)$. Lorsque $E = \mathbb{R}^p$, $F = \mathbb{R}^q$ et $G = \mathbb{R}^n$, le résultat se traduit agréablement en termes des jacobiennes :

$$J_{g \circ f}(a) = J_g(f(a)) J_f(a)$$

où le terme de droite est un produit de matrices. En appliquant la formule pour les coefficients d'un produit de matrices, on en déduit comment calculer les dérivées partielles d'une fonction composée :

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^{q} \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a)$$

où $(x_1, \ldots, x_p)$ sont les coordonnées sur $\mathbb{R}^p$, et $(y_1, \ldots, y_q)$ sont les coordonnées sur $\mathbb{R}^q$.

Pour la preuve, on utilisera une norme triple sur $L(E, F)$, qui se définit comme pour les matrices (d'ailleurs en choisissant des bases sur E et F, c'est la norme triple pour la matrice représentant l'application linéaire dans ces bases). On rappelle la définition : étant fixée des normes $\|\cdot\|_E$ et $\|\cdot\|_F$ sur E et F, la norme triple de $\ell \in L(E, F)$ est définie par

$$|||\ell||| = \sup\{\frac{\|\ell(x)\|_F}{\|x\|_E} \mid 0 \neq x \in E\}$$

Démonstration. On écrit les hypothèses explicitement, pour manipuler les développements (et on note plutôt $x = a + h$).

$$f(a+h) = f(a) + df(a)(h) + \|h\|_E \varepsilon(h) \qquad \text{avec} \qquad \lim_{h \to 0} \varepsilon(h) = 0$$

$$g(f(a) + k) = g(f(a)) + dg(f(a))(k) + \|k\|_F \eta(k) \qquad \text{avec} \qquad \lim_{k \to 0} \eta(k) = 0$$

On peut alors écrire

$$\begin{aligned}
(g \circ f)(a+h) &= g(f(a+h)) \\
&= g(f(a) + df(a)(h) + \|h\|_E \varepsilon(h)) \\
&= g(f(a)) + dg(f(a))(k(h)) + \|k(h)\|_F \eta(k(h))
\end{aligned}$$

où $k(h) = df(a)(h) + \|h\|_E\varepsilon(h)$

$$= g \circ f(a) + dg(f(a))(df(a)(h)) + \{dg(f(a))(\|h\|_E\varepsilon(h)) + \|k(h)\|_F\eta(k(h))\}$$

Dans ce développement, on a fait apparaître le terme constant, et le terme linéaire qui est le candidat pour la différentielle d'après l'énoncé. Il reste à démontrer que le reste (entre accolades) est un $o(h)$. Pour cela, on écrit

$$\frac{dg(f(a))(\|h\|_E\varepsilon(h)) + \|k(h)\|_F\eta(k(h))}{\|h\|_E} = dg(f(a))(\varepsilon(h)) + \|df(a)(\frac{h}{\|h\|}) + \varepsilon(h)\|_F\eta(k(h))$$

et on contrôle chaque terme. Par définition, on a $\lim_{h\to 0}\varepsilon(h) = 0$. Par continuité des applications linéaires, on a donc $\lim_{h\to 0} dg(f(a))(\varepsilon(h)) = 0$. De même, il est facile de vérifier que $\lim_{h\to 0} k(h) = 0$. Comme $\eta(k(h)) \to 0$, il reste à vérifier, par exemple, que $df(a)(\frac{h}{\|h\|})$ est borné. Ça découle de la linéarité en utilisant la norme triple : $\|df(a)(\frac{h}{\|h\|})\|_F \leq |||df(a)|||$. □

Ce résultat est un résultat clef, et permet de montrer la différentiabilité de fonctions construites par de nombreuses autres opérations naturelles. Pour ça, on peut avoir besoin de l'outil suivant.

14.3 Différentiabilité des applications bilinéaires et application

Proposition 14.4 (Différentielle d'une application bilinéaire). *Soit $B : E_1 \times E_2 \to F$ une application bilinéaire, et $a = (a_1, a_2) \in E_1 \times E_2$. Alors B est différentiable en a, et pour tout $h = (h_1, h_2) \in E_1 \times E_2$, on a*

$$dB(a)(h) = B(a_1, h_2) + B(h_1, a_2).$$

Démonstration. On écrit, pour utiliser la bilinéarité,

$$B(a + h) = B(a_1 + h_1, a_2 + h_2) = B(a_1, a_2) + B(a_1, h_2) + B(h_1, a_2) + B(h_1, h_2)$$

Le premier terme est la constante $B(a)$ alors que les deuxième et troisième termes forment une expression linéaire en $h = (h_1, h_2)$, qui coïncide avec le candidat pour la différentielle donné par l'énoncé. Il suffit donc de montrer que le reste est un $o(h)$. Pour ça, on va s'inspirer de la preuve de la continuité des applications linéaires qu'on a rappelé plus tôt.

Soit $(e_1, \dots, e_p)$ et $(f_1, \dots, f_q)$ des bases de E_1 et E_2. On considère les normes ∞ associées sur E_1 et E_2 :

$$\|\sum_{i=1}^{p} x_i e_i\|_{E_1} := \sup_i |x_i| \quad \text{et} \quad \|\sum_{i=1}^{p} y_j f_j\|_{E_2} := \sup_j |y_j|$$

On considère de plus la norme sur $E_1 \times E_2$ définie par

$$\|(h_1, h_2)\|_{E_1\times E_2} = \sup\{\|h_1\|_{E_1}, \|h_2\|_{E_2}\}$$

En notant $h_1 = \sum_{i=1}^p x_i e_i$ et $h_2 = \sum_{i=1}^p y_j f_j$, on a, par bilinéarité,

$$B(h_1, h_2) = \sum_{i,j} x_i y_j B(e_i, f_j)$$

Donc

$$\begin{aligned}
\frac{\|B(h_1,h_2)\|_F}{\|(h_1,h_2)\|_{E_1\times E_2}} &\leq \frac{\sum_{i,j}|x_i||y_j|\|B(e_i,f_j)\|_F}{\|(h_1,h_2)\|_{E_1\times E_2}} \\
&\leq \frac{\sum_{i,j}\|h_1\|_{E_1}\|h_2\|_{E_2}\|B(e_i,f_j)\|_F}{\|(h_1,h_2)\|_{E_1\times E_2}} \\
&\leq \frac{\sum_{i,j}\|(h_1,h_2)\|_{E_1\times E_2}^2\|B(e_i,f_j)\|_F}{\|(h_1,h_2)\|_{E_1\times E_2}} \\
&\leq \|(h_1,h_2)\|_{E_1\times E_2}\sum_{i,j}\|B(e_i,f_j)\|_F
\end{aligned}$$

qui tend vers 0 quand $h = (h_1, h_2)$ tend vers 0, car $\sum_{i,j}\|B(e_i, f_j)\|_F$ est une constante. □

Voici comment appliquer la différentiabilité d'une composée pour montrer qu'une opération préserve la différentiabilité.

Proposition 14.5 (Différentiabilité d'un produit). *Soient f et g deux fonctions de $U \subset E$ dans $\mathbb{R}$, différentiables en a. Alors la fonction $f \times g : x \mapsto f(x)g(x)$ est différentiable en a, et $d(f \times g)(a) = f(a)dg(a) + g(a)df(a)$.*

Démonstration. La fonction $C : U \to \mathbb{R}^2, x \mapsto (f(x), g(x))$ est différentiable en a (par composantes), et $dC(a) = (df(a), dg(a))$. La fonction $P : \mathbb{R}^2 \to \mathbb{R}, (s,t) \mapsto st$ est différentiable partout (car bilinéaire), et $dP(s,t)(u,v) = sv + tu$. Par composition, la fonction $f \times g = P \circ C$ est différentiable en a, et pour $h \in E$,

$$\begin{aligned}
d(f\times g)(a)(h) &= d(P\circ C)(a)(h) \\
&= (dP(C(a))\circ dC(a))(h) \\
&= dP(f(a),g(a))(df(a)(h), dg(a)(h)) \\
&= f(a)dg(a)(h) + g(a)df(a)(h)
\end{aligned}$$

□

15 Retour sur l'IAF et fonctions de classes C^1

15.1 Digression : retour sur $df(a)(v) = \partial_v f(a)$

On se rappelle qu'on a démontré $\partial_v f(a) = df(a)(v)$ si f est différentiable, mais aussi (avant) que $\partial_v f(a) = g'(0)$ où $g : I \to F, t \mapsto f(a + tv)$ et $I = \{t \in \mathbb{R} \mid a + tv \in U\}$. Montrons, par composition, l'égalité $g'(0) = df(a)(v)$ directement.

Soit $\ell : I \to E, t \mapsto a + tv$. C'est la restriction d'une fonction affine à I, donc elle est différentiable en 0, et $d\ell(0)$ est la multiplication par v. Si f est différentiable en a alors $g = f \circ \ell$ est différentiable en 0 par composition, et on a :

$$\begin{aligned} g'(0) &= dg(0)(1) \\ &= df(a)(d\ell(0)(1)) \\ &= df(a)(v) \end{aligned}$$

Il faut retenir, comme illustré dans cet exemple, que la différentiabilité d'une composée peut servir à dériver une fonction d'une variable réelle, en passant par des différentielles de fonctions de plusieurs variables.

15.2 L'énoncé pratique de l'IAF

Maintenant qu'on a rappelé la relation entre $df(a)$ et les dérivées directionnelles, on peux traduire l'IAF prouvée plus tôt en terme de différentielle, pour obtenir un énoncé facilement utilisable. Le résultat précédent reste la version plus générale de l'IAF.

On fixe des normes $\|\cdot\|_E$ et $\|\cdot\|_F$ sur E et F, et on considère la norme triple correspondante $|||\cdot|||$ sur $L(E, F)$. Dans l'énoncé qui suit, le choix des normes est important : l'énoncé est précis en termes de ces normes compatibles, et il serait moins précis en prenant une norme arbitraire sur $L(E, F)$.

Pour simplifier les hypothèses, on va supposer que l'ouvert U est *convexe*, c'est-à-dire que pour n'importe quels a et b dans U, le segment $[a, b] = \{(1-t)a + tb \mid t \in [0, 1]\}$ est inclus dans U. Par exemple, un intervalle dans $\mathbb{R}$ est convexe, ou une boule définie par une norme est convexe (par l'inégalité triangulaire).

Théorème 15.1. *Soit U un ouvert convexe de E, et $f : U \to F$ une fonction différentiable en tout point de U. On suppose qu'il existe $M \in \mathbb{R}$ tel que, pour tout $a \in U$, $|||df(a)||| \leq M$. Alors f est M-Lipschitzienne : pour tous x, y dans U, on a*

$$\|f(y) - f(x)\|_F \leq M\|y - x\|_E$$

Démonstration. Il suffit d'appliquer la version précédente de l'IAF à la fonction f sur le segment $[x, y]$, pour x et y fixés. En effet, on a pour $c \in [a, b] \subset U$,

$$\partial_{y-x} f(c) = df(c)(y - x)$$

donc

$$\|\partial_{y-x} f(c)\|_F \leq |||df(c)||| \|y - x\|_E \leq M\|y - x\|_E$$

□

15.3 Fonctions de classe C^1

Définition 15.2. Une fonction $f : U \to F$ est de classe C^1 sur U si elle est différentiable sur U et si l'application différentielle $df : U \to L(E, F), a \mapsto df(a)$ est continue.

Vous connaissez, pour la plupart, une autre définition des fonctions de classe C^1, en terme des dérivées partielles. Ces deux définitions coïncident par un résultat essentiellement prouvé dans le cours d'HAX404X l'an dernier.

Théorème 15.3. *Soit $f : U \subset \mathbb{R}^p \to F$ et $a \in U$. On suppose que f admet des dérivées partielles selon chacune des variables sur U, et que les fonctions $U \to F$, $x \mapsto \frac{\partial f}{\partial x_i}(x)$ sont continues en a pour tout $1 \leq i \leq p$. Alors f est différentiable en a.*

Pour être complets, nous inclurons une preuve de ce résultat dans la sous-section suivante.

Corollaire 15.4. *Si f admet des dérivées partielles continues sur U selon chaque variable, alors f est de classe C^1.*

L'application directe du Théorème précédent montre seulement que f est différentiable sur U, mais comme les $df(a) = \frac{\partial f}{\partial x_i}(a)e_i^*$ où $(e_1^*, \ldots, e_p^*)$ est la base duale de la base canonique $(e_1, \ldots, e_p)$ de $\mathbb{R}^p$, l'hypothèse du corollaire assure aussi que $a \mapsto df(a)$ est continue.

Passons à un corollaire de l'IAF pour les fonctions C^1, qui est particulièrement utile pour vérifier les hypothèses de Cauchy-Lipschitz.

Corollaire 15.5. *Soit $f : U \subset E \to F$ une fonction de classe C^1. Alors f est localement, Lipschitzienne. Autrement dit, pour tout $x_0 \in U$, il existe $\varepsilon > 0$ et $\lambda \geq 0$ tel que si x_1, $x_2 \in U$ avec $\|x_1 - x_0\|_E < \varepsilon$ et $\|x_2 - x_0\|_E < \varepsilon$, on a*

$$\|f(x_2) - f(x_1)\|_F \leq \lambda \|x_2 - x_1\|_E$$

Démonstration. Soit $x_0 \in U$, et soit $\varepsilon > 0$ tel que la boule fermée $\overline{B(x_0, \varepsilon)}$ de centre x_0 et de rayon ε (pour la norme $\|\cdot\|_E$) soit contenue dans U (il existe un tel ε car U est ouvert). La fonction $df|_{\overline{B(x_0,\varepsilon)}}$ est une fonction continue sur un compact, donc elle est bornée : il existe $\lambda \geq 0$ tel que, pour tout $a \in \overline{B(x_0, \varepsilon)}$, $|||df(a)||| \leq \lambda$. C'est vrai en particulier sur la boule ouverte $B(x_0, \varepsilon)$, qui est un ouvert convexe, donc on peut appliquer l'IAF et obtenir : pour tout x_1, $x_2 \in B(x_0, \varepsilon)$, $\|f(x_2) - f(x_1)\|_F \leq \lambda \|x_2 - x_1\|_E$. □

15.4 Preuve de la caractérisation des fonctions C^1 (bonus)

Démonstration. On prouve le résultat par récurrence sur le nombre de variables p. Pour $p = 1$, le résultat a déjà été vu : une fonction d'une variable qui est dérivable est différentiable. Faisons l'hypothèse de récurrence que le résultat est vrai pour toute fonction de p variables. Soit $f : U \subset \mathbb{R}^{p+1} \to F$ une fonction qui satisfait les hypothèses du théorème, on souhaite montrer qu'elle est différentiable en a.

Notons $a = (a_1, \ldots, a_{p+1})$ et $h = (h_1, \ldots, h_{p+1}) \in \mathbb{R}^{p+1}$. On a

$$f(a+h) - f(a) - \sum_{i=1}^{p+1} h_i \frac{\partial f}{\partial x_i}(a) = E_1(h) + E_2(h_1, \ldots, h_p)$$

où

$$E_1(h) = f(a+h) - f(a_1 + h_1, \ldots, a_p + h_p, a_{p+1}) - h_{p+1} \frac{\partial f}{\partial x_{p+1}}(a)$$

et

$$E_2(h_1, \ldots, h_p) = f(a_1 + h_1, \ldots, a_p + h_p, a_{p+1}) - f(a) - \sum_{i=1}^{p} h_i \frac{\partial f}{\partial x_i}(a)$$

Par hypothèse de récurrence appliquée à la fonction $g : V \subset \mathbb{R}^p \to F$ définie par $V = \{(x_1, \ldots, x_p) \in \mathbb{R}^p \mid (x_1, \ldots, x_p, a_{p+1}) \in U\}$ et $g(x_1, \ldots, x_p) = f(x_1, \ldots, x_p, a_{p+1})$, on sait que g est différentiable en $(a_1, \ldots, a_p)$, et donc que

$$E_2(h_1, \ldots, h_p) = o((h_1, \ldots, h_p)) = o(h)$$

Pour voir qu'un $o((h_1, \ldots, h_p))$ est un $o(h)$, on peut utiliser la norme infinie $\|\cdot\|_\infty$ sur $\mathbb{R}^p$ et $\mathbb{R}^{p+1}$ et remarquer que :

$$\begin{aligned}\|(h_1, \ldots, h_p)\|_\infty &= \sup_{1 \le i \le p} |h_i| \\ &\le \sup_{1 \le i \le p+1} |h_i| \\ &\le \|h\|_\infty\end{aligned}$$

Pour le terme $E_1(h)$, on doit utiliser la continuité de $\frac{\partial f}{\partial x_{p+1}}$ en a. Ceci se traduit en

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \|h\|_\infty < \eta \Rightarrow \|\frac{\partial f}{\partial x_{p+1}}(a+h) - \frac{\partial f}{\partial x_{p+1}}(a)\|_F < \varepsilon$$

Fixons pour le moment $(h_1, \ldots, h_p)$, et considérons la fonction d'une variable réelle $\varphi : I \subset \mathbb{R} \to F$ définie par $I = \{t \in \mathbb{R} \mid (a_1 + h_1, \ldots, a_p + h_p, t) \in U\}$ et

$$\varphi(t) = f(a_1 + h_1, \ldots, a_p + h_p, t) - (t - a_{p+1}) \frac{\partial f}{\partial x_{p+1}}(a)$$

C'est une fonction dérivable (c'est une somme de fonctions dérivables : la composée de f avec une fonction affine, plus une autre fonction affine), et on a

$$\varphi'(t) = \frac{\partial f}{\partial x_{p+1}}(a_1 + h_1, \ldots, a_p + h_p, t) - \frac{\partial f}{\partial x_{p+1}}(a)$$

Par l'IAF appliqué à cette fonction φ, et la continuité, on a, pour tout $\varepsilon > 0$, il existe un $\eta > 0$ tel que

$$|h_{p+1}| < \eta \quad \Rightarrow \quad \|E_1(h)\|_F = \|\varphi(h_{p+1}) - \varphi(0)\|_F \le \varepsilon |h_{p+1}|$$

Cela revient à dire que $E_1(h) = o(|h_{p+1}|) = o(h)$. Finalement, on a bien

$$f(a+h) - f(a) - \sum_{i=1}^{p+1} h_i \frac{\partial f}{\partial x_i}(a) = E_1(h) + E_2(h_1, \ldots, h_p) = o(h)$$

ce qui montre la différentiabilité de f en a. □

16 D'autres applications de l'IAF

16.1 Fonctions de différentielle nulle

On a un premier corollaire immédiat de l'IAF :

Corollaire 16.1. *Soit $f : U \subset E \to F$ une fonction différentiable sur U. On suppose que U est connexe, et que pour tout $a \in U$, $df(a) = O_{L(E,F)}$. Alors f est constante sur U.*

Démonstration. Comme $|||0_{L(E,F)}||| = 0$, on a pour tout $a \in U$, $|||df(a)||| \leq 0$. Par l'IAF, on a, pour tout x, $y \in U$,

$$\|f(y) - f(x)\|_F \leq 0 \times \|y - x\|_E = 0$$

donc $f(y) = f(x)$. □

C'est bien sûr faux sur un ouvert général. Par exemple la fonction $f :]0,1[\cup]2,3[\to \mathbb{R}$, définie par $f(x) = 0$ si $x \in]0,1[$ et $f(x) = 1$ si $x \in]2,3[$ vérifie $f'(x) = 0$ partout, mais n'est pas constante.

Le résultat est optimal en une variable réelle, mais pas en plusieurs variables, car dans ce dernier cas, l'hypothèse convexe est trop restrictive. On améliore ça en considérant la notion de connexité.

Définition 16.2. Une partie A de E est *connexe* si, pour tout $B \subset A$ tel que B est non-vide, ouvert et fermé pour la topologie induite sur A, on a $B = A$.

On rappelle qu'un ouvert de A pour la topologie induite est un sous-ensemble de A de la forme $V \cap A$, où V est un ouvert de E. De même, un fermé de A pour la topologie induite est un sous-ensemble de A de la forme $K \cap A$ où K est un fermé de E. Ce sont les notions appropriées pour les images inverses par les applications continues : si $f : A \subset E \to F$ est une application continue, et W est un ouvert de F, alors $f^{-1}(W)$ est un ouvert de A pour la topologie induite. De même, si Q est un fermé de F, alors $f^{-1}(Q)$ est un fermé de A pour la topologie induite.

Nous ne passerons pas beaucoup de temps sur les espaces connexes, contentons-nous de donner un exemple.

Proposition 16.3. *L'espace total E est connexe.*

Démonstration. Soit B une partie non-vide, ouverte et fermée de E. Par l'absurde, soit $a \in E \setminus B$. Fixons $b \in B$, et considérons l'ensemble

$$I = \{s \in \mathbb{R}_+ \mid [b, b + s(a - b)] \subset B\}$$

On a $0 \in I$, et $I \subset [0, 1[$ car $a = b + 1 \times (a - b) \notin B$. Soit $t := \sup I \in [0, 1]$, et (t_n) une suite d'éléments de I telle que $\lim_{n\to\infty} t_n = t$.

On a $b + t_n(a - b) \in B$ pour tout n, et B est fermée, donc $b + t(a - b) \in B$. De plus, si $0 \leq s < t$, alors $s < t_n$ pour n assez grand, donc $s \in I$. Ainsi, on a montré $t \in I$.

Comme B est ouvert, il existe $\varepsilon > 0$ tel que, si $|s| < \varepsilon$, $b + t(a - b) + s(a - b) \in B$. On a donc $t + \varepsilon/2 \in I$, ce qui contredit $t = \sup I$. □

La même preuve peut essentiellement être adaptée pour montrer que tout ensemble convexe est connexe, mais il y a bien plus de parties connexes que de parties convexes. En tout cas, on peut étendre le corollaire précédent aux ouverts connexes.

Corollaire 16.4. *Soit $f : U \to F$ différentiable sur U, avec U ouvert connexe et pour tout $a \in U$, $df(a) = 0_{L(E,F)}$. Alors f est constante sur U.*

Démonstration. Soit $z \in U$ fixé. On veut montrer que f est constante, donc que $f(x) = f(z)$ pour tout $x \in U$. Autrement dit, on veut montrer que $f^{-1}(f(z)) = U$. Comme U est connexe, il suffit de montrer que $f^{-1}(f(z))$ est non-vide, ouvert et fermé dans U.

Comme $z \in f^{-1}(f(z))$, cet ensemble est non-vide.

Comme f est différentiable, f est continue, donc $f^{-1}(f(z)) = f^{-1}(\{f(z)\})$ est fermé dans U, comme image inverse du fermé $\{f(z)\}$ de F.

Montrons finalement que $f^{-1}(f(z))$ est ouvert. Soit $x \in f^{-1}(f(z))$. Soit $\varepsilon > 0$ tel que la boule ouverte $B(x, \varepsilon)$ de centre x et de rayon ε soit incluse dans U (un tel ε existe car U est ouvert donc voisinage de x). Une telle boule est convexe, donc on peut appliquer le corollaire précédent : pour tout $y \in B(x, \varepsilon)$, $f(y) = f(x) = f(z)$. Donc $B(x, \varepsilon) \subset B$, ce qui montre que B est un voisinage de x. C'est valable pour tout x dans B, donc B est ouvert. $\square$

16.2 IAF et Théorème du point fixe

Le théorème du point fixe est une motivation fréquente pour montrer qu'une fonction est Lipschitzienne, en contrôlant la constante de Lipschitz. Rappelons qu'une fonction f admet x pour point fixe si $f(x) = x$.

Définition 16.5. Une fonction $f : A \subset E \to F$ est *contractante* s'il existe $\lambda \in]0, 1[$ tel que, pour tout x, $y \in A$,

$$\|f(y) - f(x)\|_F \leq \lambda \|y - x\|_E$$

Autrement dit, f est Lispchitzienne avec une constante de Lipschitz dans $]0, 1[$. L'IAF permet de vérifier cette propriété pour une fonction $f : U \subset E \to F$ différentiable sur U, avec U ouvert convexe et $|||df(a)||| \leq \lambda < 1$ pour tout $a \in U$.

Attention, cette définition dépend fortement des normes choisies : f peut être contractante pour certains choix de $\|\cdot\|_E$ et $\|\cdot\|_F$, mais pas pour d'autres.

Théorème 16.6 (Théorème du point fixe)**.** *Soit C un fermé non-vide de E, et $f : C \to E$ une application contractante, avec $f(C) \subset C$. Alors il existe un unique point fixe $x_0 \in C$ pour f.*

La preuve du Théorème du point fixe procure, en même temps, une méthode d'approximation de cet unique point fixe.

Démonstration. On considère la suite (u_n) à valeurs dans C définie par $u_0 \in C$ quelconque et $u_{n+1} = f(u_n)$. Comme $f(C) \subset C$, cette suite est bien définie. Comme l'application est contractante, on a

$$\begin{aligned} \|u_{n+2} - u_{n+1}\|_E &= \|f(u_{n+1}) - f(u_n)\|_E \\ &\leq \lambda \|u_{n+1} - u_n\|_E \end{aligned}$$

Par récurrence immédiate, $\|u_{n+1} - u_n\|_E \leq \lambda^n \|u_1 - u_0\|_E$. On a alors

$$\begin{aligned} \|u_{n+m} - u_n\|_E &\leq \left(\sum_{k=n}^{n+m-1} \lambda^k \right) \|u_1 - u_0\| \\ &\leq \left(\sum_{k=n}^{+\infty} \lambda^k \right) \|u_1 - u_0\| \end{aligned}$$

C'est un majorant indépendant de m, et la somme en facteur est le reste d'une série géométrique convergente (car $\lambda \in]0,1[$), donc le tout vers zéro uniformément en m quand $n \to +\infty$. Autrement dit, la suite (u_n) est de Cauchy.

Comme un espace vectoriel réel de dimension finie est complet, la suite (u_n) converge vers une limite $x_0 \in E$. Comme C est fermé, $x_0 \in C$. Finalement, en passant à la limite dans l'égalité $u_{n+1} = f(u_n)$, on obtient $x_0 = f(x_0)$, donc x_0 est un point fixe de f. Notons qu'on a le droit de passer à la limite car une fonction contractante (comme une fonction Lipschitzienne en général) est continue.

Pour conclure, il reste à montrer que ce point fixe est unique. Si x_0 et y_0 sont deux points fixes de f distincts, alors on a

$$\begin{aligned} \|y_0 - x_0\|_E &= \|f(y_0) - f(x_0)\|_E \\ &\le \lambda \|y_0 - x_0\| \end{aligned}$$

En divisant par $\|y_0 - x_0\|_E$, on obtient $\lambda \geq 1$, une contradiction. □

On utilisera le Théorème du point fixe dans la preuve du Théorème d'inversion locale, qui est l'un des résultats majeurs du cours. En fait, ce Théorème du point fixe dans une version un peu plus générale sert aussi à démontrer l'existence dans le Théorème de Cauchy-Lipschitz.

16.3 Commentaires sur l'existence dans Cauchy-Lipschitz

Pour pouvoir appliquer le Théorème du point fixe dans des contextes différents, il faut d'abord formuler le problème qui nous intéresse comme un problème de point fixe. Dans le cas du Théorème de Cauchy-Lipschitz, on l'a en fait déjà fait sans relever cet aspect, avec la formulation intégrale d'une équation différentielle. En effet, l'équation

$$y(t) = x_0 + \int_{t_0}^{t} f(s, y(s))\, ds$$

est une équation de point fixe pour l'application $\mathcal{E} : C^0(J, F) \to C^0(J, F)$, qui a la fonction $y \in C^0(J, F)$ associe la fonction définie par

$$\mathcal{E}(y) : t \mapsto x_0 + \int_{t_0}^{t} f(s, y(s))\, ds$$

Il y a cependant plusieurs subtilités à pendre en compte :

1. Il faut préciser J. Par les résultats d'unicité que nous avons démontré, il suffit de montrer l'existence pour un J de la forme $]x_0 - \varepsilon, x_0 + \varepsilon[$ avec ε très petit, mais il faut quand même s'assurer que la fonction $\mathcal{E}$ est bien définie.
2. Supposons que J soit bien choisi, et que $\mathcal{E}$ soit bien définie. Il faut encore montrer qu'elle est contractante, et donc, aussi choisir une norme adaptée pour montrer ça. En utilisant l'hypothèse que f est localement Lipschitzienne en espace, on remarque que si y_1 et y_2 prennent des valeurs proches, alors pour la norme sup

$\|\cdot\|_0$ sur les fonctions, on a

$$\begin{aligned}
\|\mathcal{E}(y_2) - \mathcal{E}(y_1)\|_0 &= \sup_t \|\mathcal{E}(y_2)(t) - \mathcal{E}(y_1)(t)\|_F \\
&= \sup_t \left\| \int_{t_0}^t (f(s, y_2(s)) - f(s, y_1(s)))\, ds \right\|_F \\
&\leq \sup_t \lambda \int_{t_0}^t \|y_2(s) - y_1(s)\|_F\, ds \\
&\leq \lambda\varepsilon \|y_2 - y_1\|_0
\end{aligned}$$

C'est ce type d'argument qui permet de vérifier l'hypothèse contractante.

3. Enfin, l'espace $C^0(J, F)$ est de dimension infinie ! En fixant une norme (qui ne sont dans ce cas pas toutes équivalentes), on peut essayer de suivre la preuve du Théorème du point fixe donnée ici. Le seul point qui pose problème est pour passer d'une suite de Cauchy à une suite convergente. Ceci est vrai exactement pour les espaces vectoriels normés complets. Il faut donc démontrer cette propriété pour la norme choisie. C'est cette partie du raisonnement qui sort complètement du cadre du cours.

17 Difféomorphismes

Nous allons passer à l'un des résultats majeurs du cours, le Théorème d'inversion locale. Ce résultat permet de caractériser les difféomorphismes (locaux) qui sont (localement), des bijections qui préservent toutes les informations différentielles. Les difféomorphismes sont utiles, très directement, dans le cadre du cours de mesure et intégration, pour pouvoir faire des changements de variables dans des intégrales en plusieurs variables. Plus tard, en Master par exemple, vous pourriez apprendre que ce résultat est aussi le fondement de la géométrie différentielle, qui permet d'appliquer les outils de calcul différentiels pour des fonctions définies non pas sur des ouverts d'un espace vectoriel réel, mais sur des variétés différentiables, qui y ressemblent localement.

17.1 Définition

Définition 17.1. Soit U un ouvert de E et W un ouvert de F. Une fonction $f : U \to F$ est un *difféomorphisme (de classe C^1)* de U sur W si :

1. f est une bijection de U sur W,
2. f est de classe C^1,
3. la fonction réciproque $f^{-1} : W \to E$ est de classe C^1.

Remarque 17.2. Attention, le troisième point n'est pas une conséquence des deux premiers.

Exemple 17.3. Donnons quelques exemples de difféomorphismes (ou pas) parmi des fonctions bien connues.

1. La fonction $\mathbb{R} \to \mathbb{R}, x \mapsto x^3$ est de classe C^1, c'est une bijection, mais ce n'est *pas* un difféomorphisme de $\mathbb{R}$ sur $\mathbb{R}$ car sa fonction réciproque n'est même pas dérivable en 0.
2. La fonction $\exp : \mathbb{R} \to \mathbb{R}$ est un difféomorphisme de $\mathbb{R}$ sur $\mathbb{R}_+^*$, et sa fonction réciproque est ln.
3. La fonction $\tan :]-\pi/2, \pi/2[\to \mathbb{R}$ est un difféomorphisme de $]-\pi/2, \pi/2[$, de fonction réciproque arctan.
4. Une fonction linéaire inversible $\ell : E \to F$ est un difféomorphisme : une fonction linéaire est de classe C^1, et l'application réciproque d'une application linéaire (inversible) est aussi linéaire.
5. Plus généralement, une fonction affine est un difféomorphisme si et seulement si l'application linéaire associée est inversible.

Le Théorème d'inversion locale à venir va montrer que le calcul différentiel permet de se "ramener" essentiellement à ce cas là (localement), en comparant la fonction f à sa fonction tangente en un point $x \mapsto f(a) + df(a)(x-a)$. Voyons déjà la condition nécessaire avec la proposition suivante, qui identifie également la différentielle de la fonction réciproque.

Proposition 17.4. *Si $f : U \to V$ est un difféomorphisme, alors pour tout $x \in U$, $df(x)$ est inversible, et*

$$d(f^{-1})(f(x)) = (df(x))^{-1}.$$

Remarque 17.5.

1. Comme conséquence de cette proposition, f ne peux être un difféomorphisme que si E et F ont même dimension (car ils doivent être isomorphes comme espaces vectoriels réels de dimensions finies).
2. Si $E = F = \mathbb{R}$, $df(x)$ est inversible si et seulement si $f'(x) \neq 0$. Si f est un difféomorphisme, et U est un intervalle ouvert, alors on en déduit que f' est soit strictement positif, soit strictement négatif sur U.

Démonstration. Notons $g = f^{-1}$. Par définition, on a, pour tout $y \in V$, $f \circ g(y) = y$, c'est-à-dire que, sur V, $f \circ g$ est égal à la restriction de l'identité Id_F à V. Comme cette application est linéaire, elle est différentiable, et la différentielle de Id_F en n'importe quel point est égale à Id_F. D'autre part, en différentiant $f \circ g$ comme fonction composée, on a pour tout $y \in V$,

$$\mathrm{Id}_F = d(f \circ g)(y) = df(g(y)) \circ dg(y)$$

d'où le résultat. $\square$

Pour être plus précis sur le terme *local* inséré depuis le début de cette partie, introduisons la définition de difféomorphisme local.

Définition 17.6. Une fonction $f : U \subset E \to F$ est un *difféomorphisme local* autour de $a \in U$ (de classe C^1) s'il existe un voisinage ouvert V de a dans U et un voisinage ouvert W de $f(a)$ dans F tels que la restriction $f|_V$ de f à V soit un difféomorphisme (de classe C^1) de V sur W.

Attention, une fonction peut très bien être un difféomorphisme local autour de tout point, mais ne pas être un difféomorphisme de U sur $f(U)$.

17.2 Théorèmes d'inversions

Théorème 17.7 (Théorème d'inversion locale)**.** *Soit $f : U \subset E \to F$ une fonction de classe C^1, et $a \in U$. Alors f est un difféomorphisme local autour de a si et seulement si $df(a)$ est inversible.*

Remarque 17.8.
1. Comme tous les théorèmes, il faut faire attention à toutes les hypothèses. En particulier, il faut vérifier que la fonction est C^1 avant d'espérer y appliquer le TIL.
2. Lorsque $E = F = \mathbb{R}^n$, on a $df(a)$ inversible si et seulement si $J_f(a) \in \mathrm{GL}_n(\mathbb{R}) item$, si et seulement si $\det(J_f(a)) \neq 0$.
3. La preuve est longue et plutôt difficile, on la fera plus tard.

Corollaire 17.9 (Théorème d'inversion globale)**.** *Soit $f : U \subset E \to F$ une fonction de classe C^1. La fonction f est un difféomorphisme de classe C^1 de U sur $f(U)$ si et seulement si f est injective et pour tout $a \in U$, $df(a)$ est inversible.*

Démonstration. On a déjà vu que, si $f : U \to V$ est un difféomorphisme, alors pour tout $a \in U$, $df(a)$ est inversible. De plus f est bijective, donc, en particulier, injective.

Montrons maintenant l'autre sens de l'équivalence : on suppose que f est une fonction de classe C^1, injective, et que pour tout $a \in U$, $df(a)$ est inversible. Comme f est injective, c'est une bijection de U sur $f(U)$. Il faut montrer que $f(U)$ est ouvert et que $f^{-1} : f(U) \to E$ est de classe C^1. Il suffit de montrer que, pour tout $a \in U$, $f(U)$ contient un ouvert de F qui contient $f(a)$, et f^{-1} est de classe C^1 sur cet ouvert. On applique le TIL à f en a, qui vérifient les hypothèses. On obtient bien l'existence d'un voisinage ouvert V de a dans U, d'un voisinage ouvert W de $f(a)$ dans F, tels que $f|_V : V \to W$ est un difféomorphisme. En particulier, $(f|_V)^{-1} = f^{-1}|_W$ est de classe C^1. On a bien démontré ce qu'il fallait. □

Exemple 17.10. On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(u, v) = (uv, u + v)$. C'est une fonction de classe C^1 car elle est polynomiale. On calcule facilement sa jacobienne en tout point :

$$J_f(u, v) = \begin{bmatrix} v & u \\ 1 & 1 \end{bmatrix}$$

On a $\det(J_f(u, v)) = v - u$ donc $df(u, v)$ est inversible si et seulement si $v \neq u$.

Par le Théorème d'inversion locale, on en déduit que f est un difféomorphisme local autour de tout point (u, v) avec $v \neq u$. On peut interpréter ce résultat de la manière suivante. Si un polynôme du second degré unitaire à pour racines u et $v \in \mathbb{R}$, on peut l'écrire sous la forme

$$(X - u)(X - v) = X^2 - (u + v)X + uv$$

donc uv et $u + v$ correspondent aux coefficients du polynôme. Le Théorème d'inversion locale montre que les racines sont une fonction C^1 des coefficients du polynôme, au voisinage d'un polynôme avec des racines distinctes.

Avec le Théorème d'inversion globale, on peut être plus précis. Notons d'abord que f n'est pas injective : $f(u, v) = f(v, u)$, donc f ne peut pas être un difféomorphisme sur son image. Par contre, sur l'ouvert $U = \{(u, v) \in \mathbb{R}^2 \mid u < v\}$, f est injective. Donc

f est un difféomorphisme de U sur $f(U)$ par le TIG. On peut vouloir décrire l'image plus précisément. La définition donne une paramétrisation $f(U) = \{(uv, u+v) \mid u < v\}$. On peut aussi le décrire par des conditions : (a, b) est dans $f(U)$ si et seulement si $X^2 - bX + a$ est un polynôme avec deux racines distinctes, donc si et seulement si le discriminant est strictement positif. Finalement,

$$f(U) = \{(a, b) \in \mathbb{R}^2 \mid b^2 - 4a > 0\}$$

18 Preuve du TIL

18.1 L'inversion des applications linéaires

Avant de commencer la preuve, on aura besoin de quelques résultats sur l'inversion d'applications linéaires, résumés dans la proposition suivante. On note $GL(E)$ le sous-ensemble de $L(E, E)$ formé des applications linéaires inversibles, et $\mathrm{Inv} : \mathrm{GL}(E) \to L(E, E), u \mapsto u^{-1}$ l'application qui, à une application linéaire inversible associe son inverse.

Proposition 18.1. *Le sous-ensemble* $\mathrm{GL}(E)$ *est ouvert dans* $L(E, E)$*, et l'application* Inv *est différentiable sur* $\mathrm{GL}(E)$*. De plus, sa différentielle est donnée par*

$$d\mathrm{Inv}(u)(h) = -u^{-1} \circ h \circ u$$

pour $u \in \mathrm{GL}(E)$ *et* $h \in L(E, E)$*.*

Démonstration. Vérifions d'abord que $\mathrm{GL}(E)$ est ouvert. On fixe une base $(e_1, \ldots, e_p)$ de E, qui définit un isomorphisme linéaire $\mathcal{B}$ entre $L(E, E)$ et l'espace $\mathrm{M}_p(\mathbb{R})$ des matrices carrées de taille p (on associe à un endomorphisme de E la matrice qui le représente dans cette base). Cette application est linéaire en dimension finie, donc continue. L'application déterminant $\det : \mathrm{M}_p(\mathbb{R}) \to \mathbb{R}$ est polynomiale, donc continue, donc $\det \circ \mathcal{B}$ aussi. Finalement, par les cours d'algèbre linéaire, on sait que $\mathrm{GL}(E) = (\det \circ \mathcal{B})^{-1}(\mathbb{R}^*)$, donc $\mathrm{GL}(E)$ est l'image inverse d'un ouvert par une application continue, c'est un ouvert de $L(E, E)$.

Vérifions que Inv est différentiable en utilisant le candidat pour la différentielle donné dans l'énoncé (en exercice : trouver comment deviner ce candidat, en différentiant l'égalité $u^{-1} \circ u = \mathrm{Id}_E$).

$$\begin{aligned} f(u+h) - f(u) - (-u^{-1} \circ h \circ u^{-1}) &= (u+h)^{-1} - u^{-1} + u^{-1} \circ h \circ u^{-1} \\ &= u^{-1} \circ (u \circ (u+h)^{-1} - \mathrm{Id}_E + h \circ u^{-1}) \\ &= u^{-1} \circ (((u+h) \circ u^{-1})^{-1} - \mathrm{Id}_E + h \circ u^{-1}) \\ &= u^{-1} \circ ((\mathrm{Id}_E + h \circ u^{-1})^{-1} - \mathrm{Id}_E + h \circ u^{-1}) \end{aligned}$$

Pour conclure, on utilise le résultat, admis ici, que si $\|h \circ u^{-1}\| < 1$, alors $\mathrm{Id}_E + h \circ u^{-1} \in \mathrm{GL}(E)$, et

$$\left(\mathrm{Id}_E + h \circ u^{-1}\right)^{-1} = \mathrm{Id}_E - h \circ u^{-1} + (h \circ u^{-1})^2 \sum_{k=0}^{\infty} (-1)^k (h \circ u^{-1})^k$$

où la dernière somme est la somme d'une série absolument convergente, pour la norme triple sur $L(E,E)$, et uniformément convergente, en h, pour $\|h\circ u^{-1}\| \leq c < 1$. (En exercice, vous pouvez démontrer ce résultat : faites comme pour l'exponentielle de matrice, comparez via la norme triple avec une série convergente d'une variable réelle.)

On a donc

$$\begin{aligned}
\left|\!\left|\!\left| f(u+h) - f(u) - (-u^{-1}\circ h\circ u^{-1}) \right|\!\right|\!\right| &= \left|\!\left|\!\left| u^{-1}\circ (h\circ u^{-1})^2 \sum_{k=0}^{\infty} (-1)^k (h\circ u^{-1})^k \right|\!\right|\!\right| \\
&\leq \left|\!\left|\!\left| u^{-1} \right|\!\right|\!\right| \left|\!\left|\!\left| h\circ u^{-1} \right|\!\right|\!\right|^2 \sum_{k=0}^{\infty} \left|\!\left|\!\left| h\circ u^{-1} \right|\!\right|\!\right|^k \\
&\leq |\!|\!| h |\!|\!|^2 \left|\!\left|\!\left| u^{-1} \right|\!\right|\!\right|^3 \sum_{k=0}^{\infty} \left|\!\left|\!\left| h\circ u^{-1} \right|\!\right|\!\right|^k \\
&= o(|\!|\!| h |\!|\!|)
\end{aligned}$$

Puisque la somme est uniformément convergente. □

18.2 Un lemme pour simplifier les notations

Lemme 18.2. *Soit $f : U \subset E \to F$ une fonction de classe C^1, et $a \in U$ tel que $df(a)$ soit inversible. Soit $\tilde{f} : \tilde{U} \subset E \to E$ la fonction définie par*

$$\tilde{U} := \{x \in E \mid a + x \in U\}$$

et

$$\tilde{f}(x) = (df(a))^{-1} (f(a+x) - f(a))$$

Alors, $\tilde{U}$ est ouvert, $\tilde{f}$ est de classe C^1, $0 \in \tilde{U}$, $d\tilde{f}(0) = Id_E$, $\tilde{f}(0) = 0$. De plus, f est un difféomorphisme local autour de a si et seulement si $\tilde{f}$ est un difféomorphisme local autour de 0.

Démonstration. Soit $\phi : E \to E, x \mapsto a + x$, et $\psi : F \to E, z \mapsto (df(a))^{-1}(z - f(a))$. Ces deux applications sont affines, en particulier de classe C^1. De plus, elle sont inversibles.

L'ensemble $\tilde{U}$ est l'image inverse de l'ouvert U par *phi*. Donc $\tilde{U}$ est ouvert. De plus, $a = \phi(0) \in U$ donc $0 \in \tilde{U}$. On a $\tilde{f}(0) = (df(a))^{-1}(f(a) - f(a)) = (df(a))^{-1}(0) = 0$ par linéarité.

Comme $f = \psi \circ f \circ \phi$, par composition de fonctions de classe C^1, la fonction f est de classe C^1. De plus, d'après la formule pour la différentielle d'une composée, on a $d\tilde{f}(0) = (df(a))^{-1} \circ df(a) = \mathrm{Id}_E$.

Enfin, comme ϕ et ψ sont inversible, g est un inverse local C^1 de f près de a si et seulement si $\tilde{g} = \phi^{-1} \circ g \circ \psi^{-1}$ est un inverse local de $\tilde{f}$ près de 0. □

En appliquant ce lemme, on constate que quitte à remplacer f par $\tilde{f}$, il suffit de démontrer le TIL en supposant que $f : U \subset E \to E$, $0 \in U$, $f(0) = 0$, $df(0) = \mathrm{Id}_E$.

18.3 Étape 1 : Construction de l'inverse local

On va traduire ça en un problème de point fixe, et utiliser le théorème du point fixe.

Pour $y \in E$ fixé, on considère la fonction $h_y : U \to E, x \mapsto y + x - f(x)$. Alors on $y = f(x)$ si et seulement si $h_y(x) = x$. Autrement dit, x est un antécédent de y par f si et seulement si x est un point fixe de h_y. On veut montrer que, pour tout y dans un ouvert assez petit, il existe un unique antécédent de y par f qui soit dans un ouvert assez petit. Donc on veut montrer, pour tout y dans un ouvert assez petit, que la restriction de h_y à un petit ouvert a un unique point fixe. Il suffit essentiellement de montrer qu'elle est contractante.

Considérons d'abord la partie de h_y qui ne dépend pas de y : on pose $g : U \to E, x \mapsto x - f(x)$.

Lemme 18.3. *Pour $r > 0$ assez petit, la restriction $g|_{\overline{B(0,r)}}$, de g à la boule fermée $\overline{B(0,r)}$ de centre 0 et de rayon r, est Lispchitzienne de constante de Lipschitz $\frac{1}{2}$.*

Démonstration. La fonction g est de classe C^1 par opérations. Elle satisfait $g(0) = 0 - f(0) = 0$ et $dg(0) = \mathrm{Id}_E - \mathrm{Id}_E = 0$. En particulier, $|||dg(0)||| = 0$ et, par continuité, il existe $r > 0$ tel que, si $\|x\| < 2r$, alors $|||dg(x)||| < \frac{1}{2}$.

En appliquant l'IAF à $g|_{B(0,r)}$, on obtient que cette fonction est $\frac{1}{2}$-Lipschitzienne. Par continuité, la restriction de g à la boule fermée est toujours $\frac{1}{2}$-Lipschitzienne. □

Montrons maintenant que, pour $r > 0$ donné par le lemme précédent, et $y \in B(0,r)$, la fonction $h_y : \overline{B(0,2r)} \to E, x \mapsto y + g(x)$ satisfait les hypothèses du théorème du point fixe. D'abord, $\overline{B(0,2r)}$ est clairement fermé. Il est stable par h_y : si $x \in \overline{B(0,2r)}$, alors on a

$$\|h(x)\| = \|y + g(x)\| \leq \|y\| + \|g(x)\|$$

Par le lemme précédent,

$$\|g(x)\| = \|g(x) - g(0)\| \leq \frac{1}{2}\|x\| \leq r$$

donc, comme $y \in B(0,r)$,

$$\|h(x)\| < 2r \tag{1}$$

De plus, pour tout $x, x' \in \overline{B(0,2r)}$, on a

$$\|h(x) - h(x')\| = \|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\|$$

par le lemme, donc h est $\frac{1}{2}$-contractante.

Par le théorème du point fixe, il existe un unique point fixe x_0 de h dans $\overline{B(0,2r)}$. De plus, par (1), $x_0 = h(x_0) \in B(0,2r)$. On a donc montré : pour tout $y \in B(0,r)$, il existe un unique antécédent de y par f dans $B(0,2r)$. Donc la restriction de f à $U_0 = B(0,2r) \cap f^{-1}(B(0,r))$ est une bijection de U_0 sur $B(0,r)$. De plus, comme intersection de deux ouverts contenant 0, U_0 est un ouvert contenant 0.

On note $f^{-1} : B(0,r) \to U_0$ l'inverse de $f|_{U_0}$ pour la fin de la preuve.

18.4 Étape 2 : régularité de l'inverse local

D'après la proposition du cours précédent, on sait ce que doit être la différentielle de f^{-1} en y : ce devrait être $(df(f^{-1}(y)))^{-1}$. On veut donc vérifier la différentiabilité par la définition. Pour ça, on veut déjà que le candidat pour la différentielle soit bien défini, donc que $df(f^{-1}(y))$ soit inversible.

D'après la Proposition 18.1, $\mathrm{GL}(E)$ est ouvert. De plus, f est C^1, donc df est continue. Finalement, comme $df(0) = \mathrm{Id}_E \in \mathrm{GL}(E)$, on en déduit que $df(x) \in \mathrm{GL}(E)$ pour x assez proche de 0. Quitte à prendre r plus petit à l'étape 1, on peut supposer que $df(x)$ est inversible pour $x \in U_0$.

Vérifions maintenant que f^{-1} est différentiable. Soit $y \in B(0,r)$ et k assez petit. On veut vérifier que

$$f^{-1}(y+k) - f^{-1}(y) - \left(df(f^{-1}(y))\right)^{-1}(k) = o(k)$$

Pour montrer ça, on va se ramener à la seule information de ce type qu'on ait, c'est-à-dire que f est différentiable.

Posons $x := f^{-1}(y) \in U_0$ et $h(k) := f^{-1}(y+k) - f^{-1}(y)$, de sorte qu'on peut écrire

$$\begin{aligned} f^{-1}(y+k) - f^{-1}(y) - \left(df(f^{-1}(y))\right)^{-1}(k) &= h(k) - (df(x))^{-1}(f(x+h(k)) - f(x)) \\ &= (df(x))^{-1}(df(x)(h(k)) - (f(x+h(k)) - f(x))) \end{aligned}$$

Comme f est différentiable en x, on a

$$f(x+h(k)) = f(x) + df(x)(h(k)) + \|h(k)\|\varepsilon(h(k))$$

avec $\lim_{h\to 0}\varepsilon(h) = 0$. Donc on peut écrire

$$\begin{aligned} f^{-1}(y+k) - f^{-1}(y) - \left(df(f^{-1}(y))\right)^{-1}(k) &= (df(x))^{-1}(-\|h(k)\|\varepsilon(h(k))) \\ &= \|k\|\,(df(x))^{-1}\left(-\frac{\|h(k)\|}{\|k\|}h(k)\right) \end{aligned}$$

On aura fini si on montre que

$$\lim_{k\to 0}(df(x))^{-1}\left(-\frac{\|h(k)\|}{\|k\|}h(k)\right) = 0$$

Pour cela, il suffit de montrer que $\frac{\|h(k)\|}{\|k\|}$ est borné, et que $\lim_{k\to 0} h(k) = 0$.

Pour vérifier ça, on observe :

$$\begin{aligned} \|h(k)\| &= \|x + h(k) - x\| \\ &= \|x + h(k) - f(x+h(k)) + f(x+h(k)) - (x - f(x) + f(x))\| \end{aligned}$$

en utilisant la fonction g de l'étape 1, ça s'écrit

$$\begin{aligned} &= \|g(x+h(k)) + f(x+h(k)) - g(x) - f(x)\| \\ &\leq \|g(x+h(k)) - g(x)\| + \|y + k - y\| \\ &\leq \frac{1}{2}\|h(k)\| + \|k\| \end{aligned}$$

où on a utilisé le fait que g est $\frac{1}{2}$-Lipschitzienne. Finalement, on en déduit

$$\|h(k)\| \leq 2\|k\|$$

On a donc bien montré que $\frac{\|h(k)\|}{\|k\|}$ est borné, et que $\lim_{k\to 0} h(k) = 0$, donc f^{-1} est différentiable avec la différentielle attendue.

Finalement, il faut montrer que $d(f^{-1})$ est continue. On peut écrire

$$d(f^{-1}) = \mathrm{Inv} \circ df \circ f^{-1}$$

donc par la Proposition 18.1 et par composition, c'est bien le cas.

19 Différentielles supérieures

19.1 Définitions

Pour les fonctions d'une variable réelle, la dérivée seconde est la dérivée de la fonction dérivée, la dérivée troisième est la dérivée de la dérivée seconde, etc. Dans le cas des différentielles, on peut aussi définir de telles différentielles supérieures, par récurrence. Il faut faire attention cependant aux différences. Par exemple, la différentielle d'une fonction n'a, en plusieurs variables, pas le même espace d'arrivée que cette fonction.

Définition 19.1. Soit $k \geq 2$. Une fonction $f : U \subset E \to F$ est *k-fois différentiable* en $a \in U$ si f est différentiable sur un voisinage ouvert V de a dans U, et sa fonction différentielle $df : V \to L(E, F)$ est $k-1$ fois différentiable en a.

Dans ce cas, on appelle *différentielle k-ième* de f en a, et on note $d^k f(a)$ l'application k-linéaire de $E \times \cdots \times E$ dans F définie par :

$$d^k f(a)(h_1, \ldots, h_k) := \left(\left(d^{k-1}(df)(a)\right)(h_1, \ldots, h_{k-1})\right)(h_k).$$

Une fonction f est *de classe C^k sur U* si elle est différentiable sur U et df est de classe C^{k-1} sur U.

Une fonction f est *de classe C^∞* si elle est de classe C^k pour tout $k \in \mathbb{N}$.

Attention à la nature des objets. Par exemple, la différentielle seconde $d^2 f(a)$ existe si et seulement si la différentielle de f est différentiable en a, mais $d^2 f(a)$ n'est *pas* la différentielle de df en a. On retrouve l'une avec l'autre par la relation suivante : si h_1 et $h_2 \in E$,

$$d^2 f(a)(h_1, h_2) = (d(df)(a)(h_1))\,(h_2)$$

On a des implications simples à vérifier : f $(k+1)$-fois différentiable en a implique que f est de classe C^k au voisinage de a, et en particulier, ça implique que f est k-fois différentiable au voisinage de a.

Proposition 19.2. *Les fonctions affines et les fonctions bilinéaires sont de classe C^∞. L'application différentielle seconde $a \mapsto d^2 B(a)$ d'une fonction bilinéaire B est constante.*

Démonstration. On procède par étapes. On commence par le cas particulier d'une fonction nulle : elle est différentiable et sa différentielle est la fonction nulle (définie sur un espace différent). Par une récurrence immédiate, on en déduit que toutes les fonctions nulles sont de classe c^∞.

Si une fonction f est constante, alors elle est différentiable et df est nulle. Comme df est de classe C^∞ par la première étape, on en déduit que f est de classe C^∞.

Si une fonction $f = \ell + y$ est affine (avec ℓ sa partie linéaire), alors elle est différentiable et df est *constante*, égale à ℓ partout. Donc df est C^∞, donc f aussi.

Finalement, on a vu qu'une fonction bilinéaire $B : E_1 \times E_2 \to F$ est différentiable en tout point, et que sa différentielle en un point (a_1, a_2) est donnée par $dB(a_1, a_2) : E_1 \times E_2 \to F, (v_1, v_2) \mapsto B(a_1, v_2) + B(v_1, a_2)$. Donc la fonction $dB : E_1 \times E_2 \to L(E_1 \times E_2, F)$ est une application linéaire (en particulier affine). On en déduit qu'elle est C^∞, donc que B aussi. Remarquons, en passant, que la différentielle seconde d'une application bilinéaire est constante :

$$d^2B(a_1, a_2)\left((v_1, v_2), (w_1, w_2)\right) = B(v_1, w_2) + B(w_1, v_2)$$

□

19.2 Opérations et différentiabilité supérieure

Proposition 19.3. *Les résultats d'opération sur la différentiabilité s'étendent aux différentielles supérieures. Par exemple, quand ça a du sens,*

— *f, g k-fois diff. en a, $\lambda \in \mathbb{R} \Rightarrow \lambda f + g$ k-fois diff. en a*
— *f k-fois diff. en a, g k-fois diff en $f(a) \Rightarrow g \circ f$ k-fois diff en a*
— *f, g de classe $C^k \Rightarrow \lambda f + g$ de classe C^k*
— *f, g de classe $C^k \Rightarrow g \circ f$ de classe C^k*

Démonstration. La preuve pour la linéarité est très simple, et les deux autres procèdent de la même manière. Détaillons le deuxième point par récurrence sur k.

Par le théorème de différentielle d'une composée, le résultat est vrai pour $k = 1$. Supposons le résultat vrai (pour toute fonction) jusqu'au rang k.

On suppose $f : U \subset E \to F$ $(k+1)$-fois différentiable en a et $g : V \subset F \to G$ $(k+1)$-fois différentiable en $f(a)$, tels que $f(U) \subset V$. En particulier, quitte à restreindre U, $g \circ f$ est différentiable, et

$$d(g \circ f) : U \to L(E, G), dg(f(x)) \circ df(x)$$

On va écrire ceci comme une composition.

L'application $\mathcal{C} : L(E, F) \times L(F, G) \to L(E, G), (\alpha, \beta) \mapsto \beta \circ \alpha$ est bilinéaire, donc de classe C^∞. L'application $\phi : U \to L(E, F) \times L(F, G), x \mapsto (df(x), dg(f(x)))$ est k-fois différentiable, par hypothèse et hypothèse de récurrence. Comme $d(g \circ f) = \mathcal{C} \circ \phi$, par hypothèse de récurrence cette fonction est k-fois différentiable, et cond par définition $g \circ f$ est $(k+1)$-fois différentiable en a. □

Donnons un exemple d'application aux difféomorphismes.

Proposition 19.4. *La fonction* $\mathrm{Inv} : \mathrm{GL}(E) \subset L(E, E) \to L(E, E), u \mapsto u^{-1}$ *est un difféomorphisme de classe C^∞.*

Démonstration. Posons $\psi : L(E,E)\times L(E,E) \to L(L(E,E),L(E,E))$ l'application qui a (v,w) associe l'application linéaire de $L(E,E)$ dans $L(E,E)$ définie par $h \mapsto -v\circ h\circ w$. C'est une application bilinéaire, donc C^∞. Notons aussi $\mathcal{D} : L(E,E) \to L(E,E)\times L(E,E), u \mapsto (\mathrm{Inv}(u), \mathrm{Inv}(u))$. C'est une application qui a la même régularité que ses fonctions composantes, donc que Inv.

On a déjà montré dans la Proposition 18.1 que Inv est une fonction différentiable, et que sa différentielle est donnée par

$$\begin{aligned} d\mathrm{Inv}(u)(h) &= -u^{-1}\circ h\circ u^{-1} \\ &= (\psi\circ\mathcal{D})(u)(h) \end{aligned}$$

donc $d\mathrm{Inv} = \psi\circ\mathcal{D}$. Par composition, $d\mathrm{Inv}$ a la même régularité que Inv, ce qui n'est possible que si les deux sont de classe C^∞.

De plus, on a $\mathrm{Inv}\circ\mathrm{Inv} = \mathrm{Id}_{L(E,E)}$ donc Inv est bijective et son inverse est elle-même. Par conséquent, Inv est un difféomorphisme. □

Corollaire 19.5. *Si $f : U \to V$ est un difféomorphisme, et si f est aussi une fonction de classe C^k pour $k \geq 2$, alors f est un* difféomorphisme de classe C^k*, c'est-à-dire que f^{-1} est aussi de classe C^k.*

Attention, pour $k = 1$, ça fait partie de la définition de difféomorphisme. Comme on l'a vu plus tôt, si f est une bijection et une application de classe C^1, f^{-1} n'est pas forcément C^1.

Démonstration. On raisonne par récurrence sur $k \geq 1$. Pour $k = 1$ c'est dans la définition de difféomorphisme. Supposons que, pour tout difféomorphisme de classe C^{k-1}, f^{-1} est de classe C^{k-1}. Soit f un difféomorphisme tel que f est de classe C^k.

On se rappelle que $df^{-1}(y) = \left(df(f^{-1}(y))\right)^{-1}$, donc df^{-1} s'écrit comme une composition

$$df^{-1} = \mathrm{Inv}\circ df\circ f^{-1}$$

Par définition, df est de classe C^{k-1}, et par hypothèse de récurrence, f^{-1} est au moins de classe C^{k-1}. On vient de plus de montrer que Inv est de classe C^∞. Par composition, df^{-1} est donc au moins de classe C^{k-1}. Par définition, on en déduit que f^{-1} est de classe C^k. □

19.3 Dérivées directionnelles successives

Proposition 19.6. *Soit $f : U \subset E \to F$ une fonction 2-fois différentiable en a. Alors pour $v, w \in E$, on a*

$$d^2f(a)(v,w) = \partial_v\left(\partial_w f\right)(a)$$

Démonstration. On considère la fonction $x \mapsto \partial_w f(x)$. Comme $\partial_w f(x) = df(x)(w)$, on peut écrire cette fonction comme une fonction composée. Notons $e_w : L(E,F) \to F, \ell \mapsto \ell(w)$ l'application d'évaluation en w, qui est linéaire. Alors on a $\partial_w = e_w\circ df$.

Par linéarité, e_w est différentiable, et par hypothèse, df est différentiable, donc ∂_w est différentiable, et

$$\begin{aligned}\partial_v\partial_w f(x) &= (d(\partial_w f)(x))(v)\\ &= (d(e_w \circ df)(x))(v)\\ &= (e_w \circ d(df)(x))(v)\\ &= e_w(d(df)(x)(v))\\ &= (d(df)(x)(v))(w)\\ &= d^2 f(x)(v,w)\end{aligned}$$

□

On a une expression analogue pour les différentielles supérieures.
Si $E = \mathbb{R}^p$, par k-linéarité on a

$$d^k f(a)(v_1,\ldots,v_k) = \sum_{1\leq i_1,\ldots,i_k\leq p} v_{1,i_1}\cdots v_{k,i_k}\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(a)$$

où $v_j = (v_{j,1},\ldots,v_{j,p})$ en coordonnées, et $\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(a)$ est défini par récurrence par

$$\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(a) = \frac{\frac{\partial^{k-1} f}{\partial x_{i_2}\cdots\partial x_{i_k}}}{\partial x_{i_1}}(a)$$

Définition 19.7. Soit $f : U \subset \mathbb{R}^p \to \mathbb{R}$ une fonction deux fois différentiable en a. La *hessienne* de f en a, notée $H_f(a)$ est la matrice de taille $p\times p$ qui représente l'application bilinéaire $d^2f(a);\mathbb{R}^p\times\mathbb{R}^p \to \mathbb{R}$. C'est donc la matrice dont le coefficient à la i-ème ligne et j-ème colonne est $\frac{\partial^2 f}{\partial x_i\partial x_j}(a)$.

20 Formule de Taylor-Young et applications

20.1 Théorème de Schwarz

Le but de cette section est de présenter la formule de Taylor-Young et de rappeler son application à l'étude des extrema. Nous ne ferons pas les preuves en toute généralité, mais il est important de noter que l'IAF est fondamentale pour prouver les résultats de cette section (sauf les conditions nécessaires d'extrema locaux, qui sont plus élémentaires). Avant de donner la formule de Taylor-Young, nous énonçons le Théorème de Schwarz, qui se prouve à l'aide de l'IAF.

Théorème 20.1 (Théorème de Schwarz)**.** *Soit f une fonction k-fois différentiable en a, et σ une permutation de l'ensemble* $\{1,2,\ldots,k\}$*. Alors*

$$d^k f(a)(h_1,h_2,\ldots,h_k) = d^k f(a)(h_{\sigma(1)},h_{\sigma(2)},\ldots,h_{\sigma(k)}).$$

On dit que $d^k f$ est une forme k-linéaire symétrique*.*

Nous ne faisons pas la preuve ici, vous devez avoir vu une preuve dans le cas C^2 en HAX404X. Pour interpréter le résultat, il est utile de regarder sa traduction pour un petit k, en coordonnées :

Corollaire 20.2. *Soit $f : U \subset \mathbb{R}^2 \to F$ une fonction deux fois différentiable en a, alors :*

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a)$$

Démonstration. On applique directement le Théorème de Schwarz :

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) = d^2 f(a)((1,0),(0,1)) = d^2 f(a)((0,1),(1,0)) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a)$$

□

En général, il faut retenir que, pour une fonction suffisamment différentiable, on peut faire les dérivées partielles dans l'ordre qu'on veut, et de même avec les dérivées directionnelles.

20.2 Formule de Taylor-Young

Théorème 20.3 (Formule de Taylor-Young)**.** *Soit $f : U \subset E \to F$ une fonction k fois différentiable en a. Alors*

$$f(a+h) = f(a) + \sum_{m=1}^{k} \frac{1}{m!} d^m f(a)(h, \dots, h) + o(\|h\|^k)$$

Autrement dit, une fonction f-fois différentiable admet un développement limité à l'ordre k, et les termes du développement sont encodés par les différentielles supérieures successives.

Démonstration. Nous ne faisons la preuve que pour $k = 2$. C'est suffisant pour illustrer l'étape d'hérédité qu'il faut faire en général dans la preuve par récurrence sur k.

On suppose donc que f est deux fois différentiable en a. Considérons la fonction

$$F : h \mapsto f(a+h) - f(a) - df(a)(h) - \frac{1}{2} d^2 f(a)(h,h)$$

qui est bien définie sur un voisinage ouvert de 0 dans E. *Attention, ici, c'est h la variable, alors que a est fixé. En particulier, $d^2 f(a)$ est juste une application bilinéaire fixée.*

Pour montrer le résultat, on veut montrer que F est un $o(\|h\|^2)$ près de 0. Pour l'instant on ne connait que des DL à l'ordre 1 (la différentiabilité), on va s'y ramener. La fonction F est différentiable car elle est obtenue par opérations à partir de f, de fonctions affines et de fonctions bilinéaires. On a

$$\begin{aligned} dF(h)(v) &= df(a+h)(v) - 0 - df(a)(v) - \frac{1}{2}\left(d^2 f(a)(v,h) + d^2 f(a)(h,v)\right) \\ &= df(a+h)(v) - df(a)(v) - d^2 f(a)(h,v) \end{aligned}$$

par le Théorème de Schwarz

$$\begin{aligned} &= df(a+h)(v) - df(a)(v) - (d(df)(a)(h))(v) \\ &= (df(a+h) - df(a) - d(df)(a)(h))\,(v) \end{aligned}$$

Comme f est deux fois différentiable en a, df est différentiable en a, et on a

$$df(a+h) - df(a) - d(df)(a)(h) = o(h)$$

Pour la norme triple sur $L(E, F)$ relative aux normes fixées sur E et F, cela s'interprète par : $\forall \varepsilon > 0$, $\exists \eta > 0$ tel que

$$\|h\| < \eta \Rightarrow |||df(a+h) - df(a) - d(df)(a)(h)||| \leq \|h\|\varepsilon$$

Par conséquent pour un tel couple (ε, η), et h fixé pour le moment avec $\|h\| < \eta$, on a, pour $\|y\| \leq \|h\|$, $|||dF(y)||| \leq \|y\|\varepsilon \leq \|h\|\varepsilon$.

On applique l'IAF à la fonction F sur la boule $B(0, \|h\|)$, pour obtenir

$$\|F(y)\| = \|F(y) - F(0)\| \leq \|h\|\varepsilon\|y - 0\|$$

En particulier, en passant à la limite quand $y \to h$, on a

$$\|F(h)\| \leq \varepsilon\|h\|^2$$

On a montré, $\forall \varepsilon > 0$, $\exists \eta > 0$, tel que

$$\|h\| < \eta \Rightarrow \varepsilon\|h\|^2$$

c'est-à-dire que $F(h) = o(\|h\|^2)$. □

20.3 Extrema locaux

Dans cette sous-partie, on ne considère que des fonctions à valeurs réelles.

Définition 20.4. La fonction $f : A \subset E \to \mathbb{R}$ *admet un minimum local en a* s'il existe un voisinage V de a dans E tel que, pour $x \in A \cap V$, $f(x) \geq f(a)$. Ce minimum local est *strict* si pour $x \in A \cap V \setminus \{a\}$, $f(x) > f(a)$. Enfin, le minimum local est *global* (ou minimum tout court) si on peut prendre V avec $A \cap V = A$.

La définition de maximum local, local strict, global, suit, en changeant le sens des inégalités. Attention, une fonction peut atteindre à la fois un maximum local et un minimum au local au même point, si elle est constante au voisinage de ce point.

Théorème 20.5 (Conditions nécessaires)**.** *Soit $f : U \subset E \to \mathbb{R}$ et $a \in U$ (U ouvert comme d'habitude). Si f atteint un minimum local en a et si f admet une dérivée directionnelle en a selon un vecteur $v \in E$, alors $\partial_v f(a) = 0$.*

En particulier, si f est différentiable en a, alors $df(a) = 0_{L(E,\mathbb{R})}$ (on dit que a est un point critique *de f).*

Si f atteint un minimum local en a et $\partial_v(\partial_v f)(a)$ existe pour un vecteur $v \in E$, alors $\partial_v(\partial_v f)(a) \geq 0$.

En particulier, si f est deux fois différentiable en a, alors $d^2 f(a)(v, v) \geq 0$ pour tout $v \in E$.

Démonstration. Il suffit de se ramener à la fonction d'une seule variable $g : t \mapsto f(a+tv)$ déjà vue plusieurs fois. Elle vérifie $g'(t) = \partial_v f(a+tv)$ là où ça a du sens, et elle admet un minimum local en 0, donc par la première section de ce cours, on a $g'(0) = 0$. D'autre part, on a

$$g''(0) = \lim_{t \to 0} \frac{g'(t) - g'(0)}{t} = \lim_{t \to 0} \frac{\partial_v f(a+tv) - \partial_v f(a)}{t} = \partial_v(\partial_v f)(a)$$

Par formule de Taylor pour g, on a

$$g(t) = g(0) + tg'(0) + \frac{t^2}{2} g''(0) + o(t^2)$$

et comme $g'(0) = 0$, on peut écrire

$$g''(0) = \lim_{t \to 0} \frac{2(g(t) - g(0))}{t^2}$$

On prends la limite de quelque chose de positif pour $t > 0$ assez petit, donc par passage à la limite, $g''(0) \geq 0$. $\square$

Théorème 20.6 (Condition suffisante de minimum local)**.** *Soit $f : U \subset \mathbb{R}^p \to \mathbb{R}$ une fonction deux fois différentiable en a. Si a est un point critique de f et si la hessienne de f en a est définie positive, alors f atteint un minimum local strict en a.*

Démonstration. On suppose f deux fois différentiable en $a \in U$, $df(a) = 0$ et $H_f(a)$ définie positive. Par formule de Taylor à l'ordre deux, on peut écrire

$$f(a+h) = f(a) + df(a)(h) + \frac{1}{2} d^2 f(a)(h,h) + \|h\|^2 \varepsilon(h)$$

avec $\lim_{h \to 0} \varepsilon(h) = 0$. Avec ls hypothèses, ça peut se réécrire en

$$f(a+h) - f(a) = \frac{1}{2} h^T H_f(a) h + \|h\|^2 \varepsilon(h)$$

où on identifie $h \in \mathbb{R}^p$ avec un vecteur colonne. On sait que $\frac{1}{2} h^T H_f(a) h \geq 0$ avec égalité si et seulement si $h = 0$, mais ce n'est pas suffisant a priori pour conclure.

Pour obtenir une meilleure borne sur cette quantité, on va utiliser le fait de travailler en dimension finie. L'application $Q : \mathbb{R}^p \to \mathbb{R}, h \mapsto h^T H_f(a) h$ est continue (car polynomiale). La sphère $\mathbb{S} := \{h \in \mathbb{R}^p \mid \|h\| = 1\}$ est un ensemble fermé (image réciproque d'un singleton par la norme, qui est une fonction continue), et borné (clairement inclus dans la boule fermée de centre 0 et de rayon 1), donc compact. La restriction de Q à $\mathbb{S}$ est une fonction continue sur un compact, donc elle est bornée et atteint ses bornes. En particulier, il existe $h_0 \in \mathbb{S}$ tel que $\min_{h \in \mathbb{S}} Q(h) = Q(h_0)$. Par hypothèse, $Q(h_0) > 0$.

Comme $\varepsilon(h) \to 0$, il existe $\eta > 0$ tel que, si $\|h\| < \eta$, $|\varepsilon(h)| < \frac{Q(h_0)}{4}$. Pour $h \neq 0$ avec $\|h\| < \eta$, on a donc

$$\begin{aligned} f(a+h) - f(a) &= \|h\|^2 \left(\frac{1}{2} \frac{h^T}{\|h\|} H_f(a) \frac{h}{\|h\|} + \varepsilon(h) \right) \\ &> \|h\|^2 (\frac{1}{2} Q(h_0) - \frac{Q(h_0)}{4}) \\ &> \|h\|^2 \frac{Q(h_0)}{4} > 0 \end{aligned}$$

Donc f atteint un minimum local en a. $\square$

21 Théorème des fonctions implicites

21.1 Énoncé

Théorème 21.1. *Soient E et F deux espaces vectoriels réels de dimension finie. Soit $k \in \mathbb{N}^* \cup \{\infty\}$. Soit U un ouvert de $E \times F$, et $f : U \to F$ une fonction de classe C^k sur U. Soient $a \in E$, $b \in F$ tels que $(a, b) \in U$, $f(a, b) = 0$, et l'application linéaire*

$$P : F \to F, \qquad h \mapsto df(a, b)(0, h)$$

est inversible. Alors il existe :
- *un voisinage ouvert V de (a, b) dans U,*
- *un voisinage ouvert W de a dans E,*
- *une fonction $\phi : W \to F$ de classe C^k,*

tels qu'on ait équivalence entre les deux assertions :
- *$(x, y) \in V$ et $f(x, y) = 0$*
- *$x \in W$ et $y = \phi(x)$.*

Ce résultat s'interprète de la manière suivante : le TFI donne une paramétrisation locale de l'ensemble de niveau $f^{-1}(0)$ par une fonction de x. Plus précisément, localement, $f^{-1}(0)$ coïncide avec le graphe d'une fonction ϕ.

Dans le cas particulier où $E = F = \mathbb{R}$, la condition sur l'application linéaire P s'interprète plus facilement. Si on note (x, y) les variables de f, alors P est l'application linéaire de $\mathbb{R}$ dans $\mathbb{R}$ donnée par $h \mapsto df(a, b)(0, h) = h\frac{\partial f}{\partial y}(a, b)$. Elle est inversible si et seulement si $\frac{\partial f}{\partial y}(a, b) \neq 0$. C'est encore équivalent à dire que le gradient $\nabla f(a, b)$ n'est *pas* horizontal. Comme le gradient est orthogonal aux lignes de niveaux, il faut penser que la condition traduit que $f^{-1}(0)$ n'est pas tangent, en (a, b), à une droite verticale en (a, b). Il est facile de trouver des situations (par exemple $x = y^2$ en $(0, 0)$), dans ce cas, où il est impossible d'identifier $f^{-1}(0)$ avec le graphe d'une fonction de x.

On ne peut pas, en général, donner ϕ de manière explicite (d'où le nom). Par contre, on peut facilement obtenir son développement limité à l'ordre k en a, en fonction de celui de f en (a, b).

21.2 Méthode pour obtenir le DL de ϕ

On se place dans le cadre du TFI. La première remarque à faire est que, comme $(a, b) \in V$ et $f(a, b) = 0$, $b = \phi(a)$ (par la conclusion du TFI). Ensuite, comme f et ϕ sont de classe C^k, la fonction composée $\psi : W \to F, x \mapsto f(x, \phi(x))$ est aussi de classe C^k. En fait, elle est nulle par la conclusion du TFI, mais on peut utiliser cette remarque pour calculer la différentielle de l'application nulle de deux manières.

On a, pour $h \in E$,

$$\begin{aligned}
0 &= d\psi(a)(h) \\
&= df(a, \phi(a))(h, d\phi(a)(h)) \\
&= df(a, b)(h, 0) + df(a, b)(0, d\phi(a)(h)) \\
&= df(a, b)(h, 0) + P(d\phi(a)(h))
\end{aligned}$$

Par hypothèse, P est inversible, donc

$$d\phi(a)(h) = -P^{-1}(df(a,b)(h,0)$$

Comme P est déterminé par $df(a,b)$, on obtient bien la différentielle de ϕ en a en fonction de la différentielle de f en (a,b). Par exemple, si on travaille dans $\mathbb{R}^p \times \mathbb{R}^q$ avec les bases canoniques, la matrice représentant P est une sous-matrice de la jacobienne de f en (a,b), dont les coefficients sont donc des dérivées partielles de composantes de f en (a,b).

Plus généralement, le raisonnement précédent s'applique sur un voisinage de a, et si $k > 1$, on peut réitérer le processus pour exprimer les différentielles supérieures de ϕ en a.

Exemple 21.2. Illustrons la méthode sur un exemple élémentaire : il ne faut pas apprendre la formule par coeur mais suivre les étapes pour la retrouver. On considère la fonction $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto x^2 + y^2 - 1$. Cette fonction est de classe C^∞, car polynomiale. L'ensemble de niveau $f^{-1}(0)$ est le cercle unité, dont on pourrait trouver des paramétrisations explicites, mais on veut voir ce que donne le théorème des fonctions implicites.

On calcule d'abord

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

qui est non-nul si $y \neq 0$. Donc si (a,b) est sur le cercle unité, avec $b \neq 0$, on peut appliquer le théorème des fonctions implicites à f en (a,b). On obtient :

— un intervalle ouvert $I \subset \mathbb{R}$ contenant a,
— un voisinage ouvert V de (a,b) dans $\mathbb{R}^2$, et
— une fonction $\phi : I \to \mathbb{R}$ de classe C^∞

tels que, pour $(x,y) \in V$, $x^2 + y^2 = 1$ si et seulement si $x \in I$ et $y = \phi(x)$.

En particulier, on a

$$x^2 + (\phi(x))^2 = 0$$

et, en dérivant,

$$2x + 2\phi'(x)\phi(x) = 0$$

De plus $\phi(a) = b \neq 0$, donc

$$\phi'(a) = \frac{-a}{b}$$

Par continuité, $\phi(x)$ est non nul pour x assez proche de a, donc dans un voisinage de a on a

$$\phi'(x) = \frac{-x}{\phi(x)}$$

et donc

$$\phi''(x) = \frac{-\phi(x) + x\phi'(x)}{(\phi(x))^2}$$

en particulier,

$$\phi''(a) = \frac{-b + a\frac{-a}{b}}{b^2} = -\frac{1}{b} - \frac{a^2}{b^3}$$

Et ainsi de suite...

21.3 Preuve du TFI

On se place sous les hypothèses du théorème.

L'idée est d'appliquer le théorème d'inversion locale à la fonction

$$g : U \subset E \times F \to E \times F, (x, y) \mapsto (x, f(x, y))$$

au point (a, b). On veut donc vérifier les hypothèses.

D'abord, la fonction g est de classe C^k par composantes, car f est de classe C^k. Pour $(v, w) \in E \times F$, on a

$$\begin{aligned} dg(a, b)(v, w) &= (v, df(a, b)(v, w)) \\ &= (v, df(a, b)(v, 0) + df(a, b)(0, w)) \\ &= (v, df(a, b)(v, 0) + P(w)) \end{aligned}$$

On sait que P est inversible, d'inverse noté P^{-1}, et on veut montrer que $dg(a, b)$ est inversible. On constate que c'est bien le cas en trouvant l'inverse explicite, qui est donné par

$$(v', w') \mapsto (v', P^{-1}(w' - df(a, b)(v', 0)))$$

Par le théorème d'inversion locale, il existe un voisinage ouvert V de (a, b) dans U et un voisinage ouvert $\tilde{W}$ de $g(a, b) = (a, f(a, b)) = (a, 0)$ dans $E \times F$ tels que $g|_V$ soit un difféomorphisme de V sur $\tilde{W}$. Quitte à réduire V, on peut supposer que $\tilde{W} = g(V)$ est de la forme $W \times Z$ où W est un voisinage ouvert de a dans E, et Z est un voisinage ouvert de 0 dans F.

La réciproque $g^{-1} : W \times Z \to V$ de $g|_V$ peut s'écrire sous la forme $g^{-1}(x, z) = (x, \psi(x, z))$ avec $\psi : W \times Z \to F$ de classe C^k, car la première composante de g est l'identité. On a donc des équivalences entre les assertions suivantes :

— $(x, y) \in V$ et $f(x, y) = z$

— $(x, y) \in V$ et $g(x, y) = (x, z)$

— $(x, z) \in W \times Z$ et $(x, y) = g^{-1}(x, z)$

— $(x, z) \in W \times Z$ et $\psi(x, z) = y$

En prenant la première et la dernière avec $z = 0$, et en posant $\phi : W \to F, x \mapsto \psi(x, 0)$, on en déduit le théorème.

21.4 Exemple d'application

On s'intéresse à un problème d'optimisation sous contrainte. Soient $f : \mathbb{R}^2 \to \mathbb{R}$ et $g : \mathbb{R}^2 \to \mathbb{R}$ deux fonctions de classe C^1. On veut trouver les points $(x_0, y_0) \in \mathbb{R}^2$ tels que $g(x_0, y_0) = 0$ et, si $g(x, y) = 0$, alors $f(x_0, y_0) \geq f(x, y)$. Autrement dit, si C est l'ensemble de niveau $g^{-1}(0)$, on cherche les maximum de la restriction de f à C.

Si g est identiquement nulle, on cherche juste les maximums de f, et on sait qu'il suffit de les chercher parmi les points critiques. Dans l'exercice 65 du TD, on a résolu un tel problème en se ramenant à une fonction d'une seule variable, mais la courbe était donnée explicitement comme le graphe d'une fonction. Dans le cas général, on utilise le TFI pour s'y ramener et en déduire des informations sur (x_0, y_0).

Théorème 21.3. *Pour un tel point (x_0, y_0), les vecteurs $\nabla g(x_0, y_0)$ et $\nabla f(x_0, y_0)$ sont colinéaires.*

Dans le cas où $\nabla g(x_0, y_0) \neq 0$, il existe donc un $\lambda \in \mathbb{R}$ tel que $\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$, appelé multiplicateur de Lagrange. Utilisez ce mot clé si vous souhaitez en apprendre plus sur ce problème.

Démonstration. Si $\nabla g(x_0, y_0) = 0$, il n'y a rien à faire.

Supposons que $\frac{\partial g}{\partial y}(x_0, y_0) \neq 0$. On peut donc appliquer le TFI à g en (x_0, y_0). En particulier, on a un voisinage ouvert I de x_0 et une fonction C^1 $\phi : I \to \mathbb{R}$ telle que $y_0 = \phi(x_0)$, et pour tout $x \in I$, $g(x, \phi(x)) = 0$. En dérivant, on a

$$\frac{\partial g}{\partial x}(x_0, y_0) + \phi'(x_0)\frac{\partial g}{\partial y}(x_0, y_0) = 0$$

De plus, la fonction $\alpha : I \to \mathbb{R}, x \mapsto f(x, \phi(x))$ est dérivable par composition, et atteint un maximum en x_0. Sa dérivée est donc nulle en x_0, et donc

$$\alpha'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0) + \phi'(x_0)\frac{\partial f}{\partial y}(x_0, y_0) = 0$$

On a montré que $\nabla f(x_0, y_0)$ et $\nabla g(x_0, y_0)$ sont dans le noyau d'une même forme linéaire non-nulle sur $\mathbb{R}^2$, donc ce sont des vecteurs colinéaires.

Finalement, si $\nabla g(x_0, y_0) \neq 0$ mais $\frac{\partial g}{\partial y}(x_0, y_0) = 0$, alors $\frac{\partial g}{\partial x}(x_0, y_0) \neq 0$. On peut donc appliquer le TFI en échangeant le rôle des deux variables x et y. (En exercice : rédiger cette étape proprement.) $\square$

22 Flot d'un champ de vecteur

22.1 Définition et formule du flot

Définition 22.1. Une fonction $F : U \subset E \to E$ est (parfois) appelée un *champ de vecteur* (il faut alors penser que f associe à tout *point* de U un *vecteur* de E).

Une équation différentielle de la forme $y' = F(y)$ avec F un champ de vecteur est dite *autonome*.

Remarque 22.2. Si $f : I \times U \subset \mathbb{R} \times E \to E$ et $y' = f(t, y)$ n'est pas une équation autonome, on peut considérer l'équation autonome associée $z' = F(z)$ où

$$F : \tilde{U} = I \times U \subset \tilde{E} = \mathbb{R} \times E \to \tilde{E}, (t, x) \mapsto F(t, x) = (1, f(t, x))$$

Alors si u est solution de $y' = f(t, y)$, la fonction $v : t \mapsto (t, u(t))$ est solution de $z' = F(z)$. Réciproquement, si $v :]a, b[\to \mathbb{R} \times E, t \mapsto (v_1(t), v_2(t))$ est solution de $z' = F(z)$, alors on a $v_1(t) = t + c$ pour un $c \in \mathbb{R}$, et $v_2'(t) = f(t+c, v_2(t))$. En posant $u(t) = v_2(t-c)$, on a $u'(t) = f(t, u(t))$ donc u est solution de l'équation différentielle $y' = f(t, y)$.

Si F est un champ de vecteur localement Lipschitzien, alors le théorème de Cauchy-Lipschitz (comme les théorèmes d'explosion en temps fini ou de sortie de tout compact) s'applique. En effet, une fonction localement Lipschitzienne est continue, donc F est continue, et on en déduit que la fonction $(t, x) \mapsto F(x)$ est continue et localement, Lipschitzienne en espace.

Étant donné $x \in U$, on note $u_x : I_x \to E$ la solution maximale au problème de Cauchy

$$\begin{cases} y' = F(y) \\ y(0) = x \end{cases}$$

Le flot, qu'on va définir tout de suite, est une manière de mettre toutes ces solutions ensemble.

Définition 22.3. Soit $F : U \subset E \to E$ un champ de vecteur localement Lipschitzien. Le *flot* de F est l'application

$$\Phi : \Omega \subset \mathbb{R} \times E \to E, (t, x) \mapsto \Phi(t, x)$$

définie par

$$\Omega = \{(t, x) \mid t \in I_x\}$$

et

$$\Phi(t, x) = u_x(t)$$

Exemple 22.4. Si $A \in M_n(\mathbb{R})$, et $F : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Ax$, le flot de F est donné par

$$\Phi : \Omega = \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, (t, x) \mapsto e^{tA}x$$

Définition 22.5. Un champ de vecteur localement Lipschitzien $F : U \subset E \to E$ est *complet* si son flot Φ est défini sur $\Omega = \mathbb{R} \times U$. Autrement dit, si toutes les solutions maximales de $y' = F(y)$ sont globales.

On a déjà vu au début du cours beaucoup d'exemples de champs de vecteurs qui ne sont pas complets, par exemple $F :]0, +\infty[\to \mathbb{R}, x \mapsto \frac{1}{x}$.

Proposition 22.6 (Formule du flot)**.** *Soit $F : U \subset E \to E$ un champ de vecteur localement Lipschitzien. Soient t_1, $t_2 \in \mathbb{R}$ et $x \in U$ tels que $(t_1+t_2, x) \in \Omega$ et $(t_2, \Phi(t_1, x)) \in \Omega$. Alors*

$$\Phi(t_1 + t_2, x) = \Phi(t_2, \Phi(t_1, x))$$

Démonstration. La fonction $t \mapsto \Phi(t_1 + t, x)$ est solution (maximale) du problème de Cauchy

$$\begin{cases} y' = F(y) \\ y(0) = \Phi(t_1, x) \end{cases}$$

C'est donc exactement la fonction $t \mapsto \Phi(t, \Phi(t_1, x))$. □

Plus généralement, pour $t_0 \in \mathbb{R}$ et $x \in U$, la solution maximale du problème de Cauchy

$$\begin{cases} y' = F(y) \\ y(t_0) = x \end{cases}$$

est la fonction $t \mapsto \Phi(t - t_0, x)$.

De plus comme corollaire de la preuve, on a

Corollaire 22.7. *La fonction $t \mapsto \Phi(t, x)$ est soit injective, soit périodique. Plus précisément, si $\Phi(t_1, x) = \Phi(t_2, x)$, alors cette fonction est $(t_2 - t_1)$-périodique.*

22.2 Continuité du flot

Théorème 22.8. *Soit F un champ de vecteur localement Lipschitzien, et $\Phi : \Omega \to E$ son flot. Alors Ω est ouvert, et Φ est continue.*

On va donner l'idée de la preuve, sans détailler toutes les étapes.

Remarque 22.9. On peut noter que, à x fixé, $t \mapsto \Phi(t,x)$ est solution de l'équation différentielle $y' = F(y)$, donc est continue, et $\{t \mid (t,x) \in \Omega\}$ est un intervalle ouvert. La difficulté consiste à faire varier x.

Démonstration. Fixons $(t_0, x_0) \in \Omega$, on veut montrer que Ω est un voisinage de (t_0, x_0), et que $\Phi(t,x) \to \Phi(t_0,x_0)$ si $(t,x) \to (t_0,x_0)$. On suppose pour simplifier que $t_0 > 0$ (mais le cas général devrait être clair à la fin de la preuve).

Soit $t_1 > t_0$ tel que $(t_1, x_0) \in \Omega$ (qui existe par la remarque qui précède la preuve). On considère un ensemble de la forme

$$K := \{x \in E \mid \exists t \in [0, t_1], \|x - \Phi(t, x_0)\| \leq r\}$$

pour un $r > 0$. Comme $[0, t_1]$ est un segment, donc compact, son image par l'application (continue par la remarque précédent la preuve) $t \mapsto \Phi(t, x_0)$ est un compact dans U. Il n'est pas difficile de se convaincre que, comme U est ouvert, pour r assez petit, on a $K \subset U$ (pour le faire proprement, utiliser la propriété de Borel-Lebesgue du compact $\{\Phi(t,x_0) \mid t \in [0,t_1]\}$). De plus, K est borné (c'est facile) et fermé (le prouver par critère séquentiel en utilisant la compacité de $\{\Phi(t,x_0) \mid t \in [0,t_1]\}$), donc K est compact.

Comme F est localement Lispchitzienne et K compact, il existe un $\lambda \geq 0$ tel que, pour x_1 et x_2 dans K,

$$\|F(x_1) - F(x_2)\| \leq \lambda \|x_1 - x_2\|$$

Soit $x \in U$ tel que $\|x - x_0\| < r$, alors pour t assez petit, $\Phi(t,x) \in K$, et tant que $\Phi(t,x) \in K$ et $t \geq 0$, on a, par formulation intégrale,

$$\begin{aligned}
\|\Phi(t,x) - \Phi(t,x_0)\| &= \|x - x_0 + \int_0^t (F(\Phi(s,x)) - F(\Phi(s,x_0)))\, ds\| \\
&\leq \|x - x_0\| + \int_0^t \|F(\Phi(s,x)) - F(\Phi(s,x_0))\|\, ds \\
&\leq \|x - x_0\| + \int_0^t \lambda \|\Phi(s,x) - \Phi(s,x_0)\|\, ds
\end{aligned}$$

Par le lemme de Grönwall, on en déduit

$$\|\Phi(t,x) - \Phi(t,x_0)\| \leq \|x - x_0\| e^{\int_0^t \lambda\, ds} = \|x - x_0\| e^{\lambda t} \leq \|x - x_0\| e^{\lambda t_1}$$

En particulier, si $\|x - x_0\| \leq r e^{-\lambda t_1}$, on a $\Phi(t,x) \in K$ si $t \in [0,t_1]$ et la solution maximale $t \mapsto \Phi(t,x)$ est définie en t. Par Théorème de sortie de tout compact, on en déduit que cette solution maximale est définie pour tout $t \in [0, t_1]$.

On a donc $[0,t_1] \times B(x_0, re^{-\lambda t_1}) \subset \Omega$, et en particulier $(t_0, x_0) \in]0, t_1[\times B(x_0, re^{-\lambda t_1}) \subset \Omega$, donc Ω est bien un voisinage de (t_0, x_0).

Pour la continuité, prenons $(t,x) \in [0,t_1]\times B(x_0, re^{-\lambda t_1})$, et posons $M; = \sup\{\|F(x)\| \mid x \in K\}$, qui est fini car K est compact et F est continue. Alors

$$\begin{aligned}\|\Phi(t_0,x_0) - \Phi(t,x)\| &\leq \|\Phi(t_0,x_0) - \Phi(t_0,x)\| + \|\Phi(t_0,x) - \Phi(t,x)\| \\ &\leq \|x_0 - x\|e^{\lambda t_1} + \|\int_t^{t_0} F(\Phi(s,x))\,ds\| \\ &\leq \|x_0 - x\|e^{\lambda t_1} + M|t_0 - t|\end{aligned}$$

Et donc

$$\lim_{(t,x)\to(t_0,x_0)} \Phi(t,x) = \Phi(t_0,x_0)$$

□

23 Portrait de phase

On suppose dans toute la section que $F : U \subset E \to E$ est un champ de vecteur localement Lipschitzien.

23.1 Terminologie

Définition 23.1. Pour une équation différentielle autonome $y' = F(y)$ avec $F : U \subset E \to E$, on appelle (parfois) U l'*espace des phases*.

Attention, si on s'intéresse, par exemple, à une équation différentielle d'ordre 2 telle que $y'' = g(y)$, l'espace des phases celui de l'équation autonome de degré 1 associée $(y,v)' = (v, g(y))$. Si y s'interpréte comme la position d'une particule, alors l'espace des phases est l'espace des couples (position, vitesse).

Définition 23.2. On appelle *orbite*, ou *trajectoire de* $x \in U$ l'ensemble

$$\mathrm{orb}(x) := \{\Phi(t,x) \mid (t,x) \in \Omega\}$$

Les orbites forment une partition de U : si x_1 et x_2 sont dans U, on a soit $\mathrm{orb}(x_1) = \mathrm{orb}(x_2)$, soit $\mathrm{orb}(x_1) \cap \mathrm{orb}(x_2) = \emptyset$. Cette partition, ou sa représentation schématique est (parfois) appelé *portrait de phase*. En pratique, on représente quelques trajectoires remarquables qui permettent de deviner l'allure des autres.

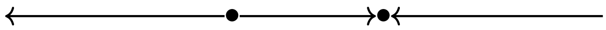
Définition 23.3. Un *point d'équilibre* pour F est un point $x \in U$ tel que $\mathrm{orb}(x) = \{x\}$.

Autrement dit, x est un point d'équilibre de F si la fonction constante égale à x est solution de l'équation différentielle $y' = F(y)$. Ces points sont donc exactement les solutions de l'équation $F(x) = 0$.

Parmi les trajectoires remarquables qu'on représente sur un portrait de phase, on inclus toujours les points d'équilibre. Voici quelques exemples très simples de portraits de phase, obtenu grâce aux cours ou TD précédents.

Exemple 23.4. Pour l'équation autonome $y' = -y$ dans $\mathbb{R}$, dont le flot est donné par $\Phi(t,x) = e^{-t}x$, on a trois orbites données par $\mathbb{R}_+^*$, $\mathbb{R}_-^*$ et $\{0\}$. De plus, on connaît le sens de parcours des trajectoires, qu'on peut aussi indiquer sur le portrait de phase :

Exemple 23.5. Pour l'équation logistique $y' = y(1-y)$ dans $\mathbb{R}$, on a cinq orbites données par $\mathbb{R}_-^*$, $\{0\}$, $]0,1[$, $\{1\}$ et $]1,+\infty[$. De plus, on connaît le sens de parcours des trajectoires, qu'on peut aussi indiquer sur le portrait de phase :



23.2 Exemple d'étude qualitative utilisant le portrait de phase

Exercice 93 du TD (modèle SIR)

23.3 Intégrales premières

Définition 23.6. On dit que $H : U \subset E \to \mathbb{R}$ est une *intégrale première* du champ de vecteur F si H est constante le long des orbites de F.

Proposition 23.7. *Si H est une fonction différentiable, c'est une intégrale première si et seulement si pour tout $x \in U$,*

$$dH(x)(F(x)) = 0$$

En particulier, si $E = \mathbb{R}^n$ muni du produit scalaire usuel, H est une intégrale première si et seulement si le gradient ∇H de H est orthogonal à F en tout point.

Démonstration. La définition de H intégrale première est que, pour tout $x \in U$, la fonction $t \mapsto H(\Phi(t,x))$ est constante. Si H est différentiable, cette fonction est a priori dérivable, et H est une intégrale première si et seulement si pour tout $x \in U$ et tout $t \in \mathbb{R}$ tel que $(t,x) \in \Omega$, $dH(\Phi(t,x))(\frac{\partial \Phi}{\partial t}(t,x)) = 0$. Comme $\frac{\partial \Phi}{\partial t}(t,x) = F(\Phi(t,x))$, on en déduit que H est une intégrale première si et seulement si pour tout $x \in U$ et tout $t \in \mathbb{R}$ tel que $(t,x) \in \Omega$, $dH(\Phi(t,x))(F(\Phi(t,x))) = 0$. Mais $\Phi(t,x) \in U$, donc il suffit de vérifier que, pour tout $x \in U$, $dH(x)(F(x)) = 0$. La dernière phrase de la proposition découle du fait que dans le cas où $E = \mathbb{R}^n$, $dH(x)(F(x))$ est égal au produit scalaire de $\nabla H(x)$ avec $F(x)$. $\square$

Une intégrale première donne des informations sur le portrait de phase : chaque orbite est contenue dans un ensemble de niveau de H. En particulier, si tous les ensembles de niveau de H sont des compacts de U, alors par le Théorème de sortie de tout compact, le champ de vecteur est complet.

Attention cependant : en général, une orbite peut être *strictement* incluse dans un ensemble de niveau de H, par exemple, on a toujours des intégrales premières triviales données par les fonctions constantes. De plus, en général il n'existe pas d'intégrales premières non-triviales (disons différentiables) pour un champ de vecteur donné (penser aux exemples de portraits de phase en dimension un donnés plus tôt).

Une intégrale première est particulièrement intéressante en dimension 2 : dans ce cas, les ensembles de niveaux sont en général des courbes, de même dimension que les orbites du champ de vecteurs.

24 Stabilité

On suppose pour toute la section que $F : U \subset E \to E$ est un champ de vecteur localement Lipschitzien, et on note $\Phi : \Omega \subset \mathbb{R} \times E \to E$ son flot.

24.1 Définitions

Définition 24.1. Un point d'équilibre x_0 de F est *stable* si :

$$\forall \varepsilon > 0, \exists \delta > 0, \mathrm{tq} \|x - x_0\| < \delta \Rightarrow \forall t \geq 0, \|\Phi(t, x) - x_0\| < \varepsilon$$

En une phrase moins précise : x_0 est stable si l'orbite future de tout point assez proche de x_0 reste proche de x_0.

Remarque 24.2. Par théorème de sortie de tout compact, si x_0 est un point d'équilibre stable, alors il existe $\delta > 0$ tel que $[0, +\infty[\times B(x_0, \delta) \subset \Omega$.

Exemple 24.3. Avec les exemples de portraits de phase avec sens de parcours des trajectoires donnés plus tôt, on voit que 0 est point d'équilibre stable pour $F : \mathbb{R} \to \mathbb{R}, x \mapsto \lambda x$ avec $\lambda \leq 0$, mais pas pour $\lambda > 0$.

Pour $F : \mathbb{R} \to \mathbb{R}, x \mapsto x(1 - x)$ correspondant à l'équation logistique, le point d'équilibre 1 est stable, alors que le point d'équilibre 0 est instable.

Exemple 24.4. Voyons comment une intégrale première permet parfois de montrer qu'un point d'équilibre est stable. On considère le champ de vecteur $F : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (-y, x)$ (il est linéaire, donc on pourrait expliciter le flot, mais on illustre plutôt ici l'usage d'une intégrale première). Pour trouver une intégrale première, disons différentiable, on veut trouver $H : \mathbb{R}^2 \to \mathbb{R}$ telle que, pour tout $(x, y) \in \mathbb{R}^2$,

$$-y \frac{\partial H}{\partial x}(x, y) + x \frac{\partial H}{\partial y}(x, y) = 0$$

Comme les variables sont en quelque sorte séparées, on peut chercher H sous la forme $H(x, y) = f(x) + g(y)$ pour deux fonctions dérivables f et g. Alors l'équation ci-dessus s'écrit

$$-yf'(x) + xg'(y) = 0$$

et si $g'(y) = y$ et $f'(x) = x$, on a évidemment une solution. On choisit donc des primitives, et on obtient l'intégrale première

$$H(x, y) = \frac{x^2 + y^2}{2}$$

Ses ensembles de niveau sont les cercles centrés en 0. En particulier, ils sont compacts, donc le champ de vecteur est complet. De plus, si $\|x - 0\| < \varepsilon$, alors pour $t \geq 0$, $\Phi(t, x)$ est dans le cercle de centre 0 et de rayon $\|x\|$, donc $\|\Phi(t, x) - 0\| < \varepsilon$. On a montré que 0 est un point d'équilibre stable pour F.

Définition 24.5. Un point d'équilibre x_0 est *asymptotiquement stable* si il est stable et qu'il existe $\varepsilon > 0$ tel que

$$\|x - x_0\| < \varepsilon \Rightarrow \lim_{t \to +\infty} \Phi(t, x))x_0$$

Exemple 24.6. Les exemples de points d'équilibre stables donnés plus haut sont asymptotiquement stables, sauf 0 pour $F : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (-y, x)$.

24.2 Fonctions de Lyapunov

Définition 24.7. Soit $x_0 \in U$ et V un voisinage de x_0 dans U. On dit que $L : V \to \mathbb{R}$ est une *fonction de Lyapunov* pour F au voisinage de x_0 si L est continue, atteint un minimum strict en x_0, et est strictement décroissante le long des intersections d'orbites avec $V \setminus \{x_0\}$.

Pour préciser la dernière partie de la définition : si $x \in V \setminus \{x_0\}$, et $t_0 > 0$ est tel que $[0, t_0[\times \{x\} \subset \Omega$ et $\Phi(t, x) \in V \setminus \{x_0\}$ pour $t \in [0, t_0[$, alors la fonction $[0, t_0[\to \mathbb{R}, t \mapsto L(\Phi(t, x))$ est strictement décroissante.

Remarque 24.8. Si L est une fonction différentiable, qui atteint un minimum strict en x_0 et vérifie, pour tout $x \in V \setminus \{x_0\}$,

$$dL(x)(F(x)) < 0$$

alors c'est une fonction de Lyapunov.

Théorème 24.9. *Si L est une fonction de Lyapunov pour F au voisinage de x_0, alors x_0 est un point d'équilibre asymptotiquement stable.*

Remarque 24.10. Comme on va le voir dans la preuve, si on remplace *strictement décroissante le long des orbites* par *décroissante le long des orbites*, alors la preuve montre que x_0 est un point d'équilibre stable. En particulier, on peut appliquer cette version avec une intégrale première, comme dans l'exemple de la sous-section précédente.

Démonstration. Pour simplifier les notations on suppose que $L(x_0) = 0$.

Commençons par vérifier que x_0 est un point d'équilibre. On considère la fonction $[0, \varepsilon[\to \mathbb{R}, t \mapsto L(\Phi(t, x_0)$. Si $\Phi(t, x_0) = x_0$ pour $t \in [0, \varepsilon[$ avec un $\varepsilon > 0$ alors x_0 est un point d'équilibre, car

$$F(x_0) = \frac{\partial \Phi}{\partial t}(0, x_0) = 0$$

Donc il existe une suite (t_n) de réels strictement positifs, décroissante, telle que $\lim t_n = 0$ et pour tout n, $\Phi(t_n, x_0) \neq x_0$. La suite $(L(\Phi(t_n, x_0)))$ suite est strictement croissante : par hypothèse et par la formule du flot,

$$L(\Phi(t_{n+1}, x_0)) > L(\Phi(t_n - t_{n+1}, \Phi(t_{n+1}, x_0))) = L(\Phi(t_{n+1} + (t_n - t_{n+1}), x_0))$$

Par continuité de L, on a

$$L(x_0) = \lim L(\Phi(t_n, x_0))$$

comme limite strictement croissante, ce qui contredit l'hypothèse que L admet un minimum strict en x_0.

Montrons maintenant que le point d'équilibre x_0 est stable. Soit $r > 0$ tel que $\overline{B(x_0, r)} \subset V$, et soit $0 < \varepsilon < r$. La restriction de la fonction L à $\overline{B(x_0, r)} \setminus B(x_0, \varepsilon)$ est une fonction continue sur un compact, donc elle atteint son minimum b en un point. De plus, comme elle prends des valeurs strictement positive (0 est un minimum strict), $b > 0$.

Pour $c < b$, le fermé $K_c := L^{-1}([0, c])$ est un voisinage de x_0 contenu dans $B(x_0, \varepsilon)$. Soit $x \in K_c \setminus \{x_0\}$. Comme L est décroissante le long des orbites futures, on a, pour $t \geq 0$,

$$L(\Phi(t, x)) \leq L(x) \leq c < b$$

(On n'a pas encore besoin, ici, de la décroissance stricte.) En particulier, pour tout $t \geq 0$, $\Phi(t, x) \in K_c \subset B(x_0, \varepsilon)$. On a montré ainsi que x_0 est un point d'équilibre stable (car, comme K_c est un voisinage de x_0, il existe $\delta > 0$ tel que $B(x_0, \delta) \subset K_c$).

Montrons finalement que le point d'équilibre est asymptotiquement stable. Il suffit de montrer que, pour $x \in K_c$, $\lim_{t\to+\infty} \Phi(t, x) = x_0$. Soit $x \in K_c$. La fonction $t \mapsto L(\Phi(t, x))$ est strictement décroissante positive, donc admet une limite $a \in [0, c[$. Supposons, par l'absurde, que $\lim_{t\to+\infty} \Phi(t, x) \neq x_0$. Alors comme K_c est compact, il existe une suite (t_n) telle que $\lim_{n\to\infty} t_n = +\infty$ et $\lim_{n\to\infty} \Phi(t_n, x) = x_1 \in K_c \setminus \{x_0\}$. Ce point x_1 vérifie $0 < L(x_1) = a \leq c$ (où l'inégalité stricte est due au fait que L atteint un minimum strict en x_0). Pour $t > 0$, on a

$$L(\Phi(t, x_1)) = \lim_{n\to\infty} L(\Phi(t, \Phi(t_n, x)))$$

car L et Φ sont continues

$$= \lim_{n\to\infty} L(\Phi(t + t_n, x))$$

par formule du flot

$$= a$$
$$= L(x_1)$$

C'est une contradiction avec la décroissance stricte de L le long des orbites. $\square$

24.3 Exemples

Exercices 97 et 98 (pendule simple sans ou avec frottements).