Rappels Calcul Différentiel pour le M1

Francis Nier

Dans ce qui suit on travaille dans des espaces vectoriels normés sur $\mathbb{R}$, $(E, \| \ \|_E)$, $(F, \| \ \|_F)$ ... qui sauf mention du contraire seront supposés être des espaces de Banach. Sur une algèbre munie d'une unité, on considèrera tout particulièrement le cas des normes d'algèbre vérifiant $\|AB\| \le \|A\|\|B\|$ et $\|1\| = 1$. Une norme $\| \ \|_E$ étant fixée sur E , $\|A\| = \sup_{x\in E\,,x\neq 0} \frac{\|Ax\|}{\|x\|}$ définit une norme d'algèbre sur l'ensemble $\mathcal{L}(E)$ des applications linéaires continues de E dans E . On se rappellera en particulier les résultats correspondants à la dimension finie $E \sim \mathbb{R}^n$. Enfin il est très commode d'utiliser les notations $g(x) = \mathcal{O}_{x\to a}(f(x))$ (resp.$g(x) = o_{x\to a}(f(x))$) avec $a \in E$, $g : E \to F$ et $f : E \to \mathbb{R}_+$, correspondant à

$$
\begin{aligned}
& \exists C > 0\,, \exists \delta > 0\,, \forall x \in B(a, \delta)\,, \|g(x)\|_F \le C f(x) \\
\text{resp.} \quad & \forall \varepsilon > 0\,, \exists \delta_\varepsilon > 0\,, \forall x \in B(a, \delta_\varepsilon)\,, \|g(x)\|_F \le \varepsilon f(x)\,.
\end{aligned}
$$

Nous considèrerons typiquement des fonctions $f : U \to F$ avec U ouvert de E , ou voisinage ouvert du point a .

1 Différentiabilité, différentielle, dérivation partielle

1.1 Définitions et généralités

Definition 1.1. *Une application $H : U \subset E \to F$, U ouvert de E , est dite différentiable en $a \in U$ si il existe une application linéaire continue $L \in \mathcal{L}(E; F)$ telle que*

$$H(x) = H(a) + L(x - a) + o_{x\to a}(\|x - a\|_E)\,.$$

On note $L = DH_a$ ou $L = D_xH_a$ ou $L = dH(a)$ suivant les ouvrages (la première étant privilégiée ici, la deuxième étant pour préciser que l'on différentie par rapport à x). L'application $E \ni x \mapsto H(a) + DH_a(x - a)$ est l'application affine tangente.

Si H est différentiable en a , elle est continue en a et sa différentielle est unique. Si $E = \mathbb{R}$, et $U = I$ est un intervalle ouvert de $\mathbb{R}$ contenant a , la différentiabilité en a équivaut à la dérivabilité, mais attention il convient de ne pas confondre la dérivée (éventuellement vectorielle) $H'(a) \in F$ et la différentielle qui est l'application $DH_a \in \mathcal{L}(\mathbb{R}; F)$, $s \mapsto DH_a(z) = s.H'(a)$.
Si $E_1 \subset E$ est un sous-espace fermé de E , on dit que la fonction H est différentiable en a dans la direction E_1 si la fonction $H\big|_{U\cap(a+E_1)}$ restreinte au sous espace affine $a + E_1$ est différentiable en a . En particulier quand $E = E_1 \oplus E_2$ avec la décomposition $x = x_1 \oplus x_2$ des vecteurs, il est utile de spécifier la variable de différentiation $D_{x_1}H_a \in \mathcal{L}(E_1, F)$ ou $D_{x_2}H_a \in \mathcal{L}(E_2, F)$.
Bien évidemment une fonction différentiable en a est différentiable dans toutes les directions.
Dans le cas $E_1 = \mathbb{R}v$ avec $v \neq 0$, cela permet de définir la dérivée partielle le long du vecteur v : $\partial_v H(a) = DH_a(v) = \lim_{t\to 0,t\neq 0} \frac{1}{t}(H(a + tv) - H(a))$.
Cas particuliers usuels :1) Si $E \sim \mathbb{R}^m$ et $(e_1, \ldots, e_m)$ une base de E avec l'écriture $x = x_1e_1 + \cdots + x_me_m$, on note $\frac{\partial H}{\partial x_i}(a)$ ou simplement $\partial_{x_i}H(a)$ la dérivée partielle $\partial_{x_i}H(a) = DH_a(e_i) = \lim_{t\to 0 t\neq 0} \frac{H(a+te_i)-H(a)}{t}$ de H suivant le vecteur de base e_i (En pratique cela consiste à figer toutes les autres coordonnées $x_{i'}$, $i' \neq i$ et à dériver en une seule variable).

2) Si $E \sim \mathbb{R}^m$ et $F \sim \mathbb{R}^n$ avec les bases $\mathcal{E} = (e_1, \ldots, e_m)$ et $\mathcal{F} = (f_1, \ldots, f_n)$ et $H(x) = \begin{pmatrix} H_1(x_1, \ldots, x_m) \\ \vdots \\ H_n(x_1) \ldots, x_m) \end{pmatrix}$
ou encore $H(\sum_{1\le i\le m} x_ie_i) = \sum_{1\le j\le n} H_j(x)f_j$, la matrice jacobienne de H en a n'est autre que la matrice de DH_a dans les bases $\mathcal{E}$ et $\mathcal{F}$ est

$$\begin{pmatrix} \partial_{x_1}H_1(a) & \ldots & \partial_{x_m}H_1(a) \\ \vdots & \vdots & \vdots \\ \partial_{x_1}H_n(a) & \ldots & \partial_{x_m}H_n(a) \end{pmatrix}$$

L'implication " H différentiable implique H dérivable dans toutes les directions" est toujours vraie. Sa réciproque en revanche, et ce même en dimension finie, n'est vraie que dans la version $\mathcal{C}^1$: Si $E \sim \mathbb{R}^m$ alors H est $\mathcal{C}^1$ (cf définition ci-dessous) au voisinage de a si et seulement si toutes les dérivées partielles $\partial_{x_i}H$ définissent des fonctions continues au voisinage de a .
On note $\mathcal{L}^k(E; F)$ l'ensemble des applications $\underbrace{E \times \cdots \times E}_{k \text{ fois}} \to F$ qui sont k-linéaires (linéaires par rapport à

chacune des variables $x_i \in E$, $1 \leq i \leq k$) et continues ($\|M(x_1,\ldots,x_k)\|_F \leq C_M\|x_1\|_E \times \|x_k\|_E$)). Et on note $\mathcal{S}ym^k(E;F) \subset \mathcal{L}^k(E;F)$ le sous-espaces des applications k-linéaires continues symétriques, i.e. vérifiant de plus $M(x_{\sigma(1)},\ldots,x_{\sigma(k)}) = M(x_1,\ldots,x_k)$ pour toute $\sigma \in \mathfrak{S}_k$.

Definition 1.2. *Une application $H : U \to F$, U ouvert de E est $\mathcal{C}^1$ sur U si elle est différentiable en tout point et si l'application $DH : U \ni x \to DH_x \in \mathcal{L}(E;F)$ est continue. On note $\mathcal{C}^1(U;F)$ l'ensemble des applications $\mathcal{C}^1$ de U dans F .*
Par récurrence on dit qu'une application H est $\mathcal{C}^k$, $H \in \mathcal{C}^k(U;F)$, si $H \in \mathcal{C}^{k-1}(U;F)$ et $D^{k-1}H \in \mathcal{C}^1(U;\mathcal{L}^{k-1}(E;F))$, la différentielle d'ordre k est alors donnée par $D^kH = D(D^{k-1}H) \in \mathcal{C}^0(U;\mathcal{L}^k(E;F))$.
Enfin on dit qu'une application $H \in \mathcal{C}^\infty(U;F)$ si $H \in \mathcal{C}^k(U;F)$ pour tout $k \in \mathbb{N}$.

En vertu des isomorphismes $\mathcal{L}^k(E;F) \sim \mathcal{L}(\mathcal{L}^{k-1}(E;F);F) \sim \mathcal{L}^{k-1}(E;\mathcal{L}(E;F))$ la définition par récurrence ci-dessus est équivalente à $H \in \mathcal{C}^1(U;F)$ et $DH \in \mathcal{C}^{k-1}(U,F)$ et on a $D(D^{k-1}H) = D^kH = D^{k-1}(DH)$ (associativité de la différentiation) .

Lemme 1.3. *(Lemme de Schwarz) Si $H \in \mathcal{C}^2(U;F)$ avec U ouvert de $\mathbb{R}^2$ alors on a toujours $\partial_{x_1}\partial_{x_2}H(a) = \partial_{x_2}\partial_{x_1}H(a)$.*

Cela a pour conséquence que la différentielle d'ordre k en a d'une fonction $H \in \mathcal{C}^k(U;F)$ est une application k-linéaire continue symétrique, $D^kH_a \in \mathcal{S}ym^k(E;F)$.
Dans le cas particulier $E \sim \mathbb{R}^m$ avec la base $\mathcal{E} = (e_1,\ldots,e_m)$, on peut commuter les ordres des dérivées partielles. Il est alors commode d'introduire les multi-indices $\alpha = (\alpha_1,\ldots,\alpha_m) \in \mathbb{N}^m$ de longueur $|\alpha| = \alpha_1 + \cdots + \alpha_d$ avec la notation

$$\partial_x^\alpha f(a) = (\partial_{x_1}^{\alpha_1} \ldots \partial_{x_m}^{\alpha_m} f)(a)\,.$$

.

1.2 Règles de calculs, formules et contre-exemples

Ces règles de calculs permettent entre autre de différentier toutes les fonctions vectorielles de plusieurs variables exprimées à partir de fonctions usuelles.
Une combinaison linéaire d'applications différentiables est différentiable et la différentielle est la combinaison linéaire des différentielles.
La différentielle d'une application affine continue $H : E \ni x \mapsto Lx + b \in F$, $L \in \mathcal{L}(E;F)$, $b \in F$ est donnée par $DH_a = L$. Ainsi $DH : E \ni x \to L \in \mathcal{L}(E;F)$ est une application constante.
Si $M : E = E_1 \times \cdots \times E_k \to F$ est une application k-linéaire continue alors sa différentielle au point $a = (a_1,\ldots,a_k)$ est donnée par

$$DM_a(v_1,\ldots,v_k) = M(v_1,a_2,\ldots,a_k) + \cdots + M(a_1,\ldots,a_{k-1},v_k)\,.$$

En particulier si $\mathcal{A}$ est une algèbre de Banach (muni d'une norme d'algèbre) alors l'application produit $M : \mathcal{A} \times \mathcal{A} \ni (A_1,A_2) \to A_1A_2 \in \mathcal{A}$ est différentiable et $DM_{(A_1,A_2)}(B_1,B_2) = B_1A_2 + A_1B_2$.
L'ensemble des éléments inversibles $\mathcal{A}^*$ d'une algèbre de Banach est un ouvert (se rappeler en particulier la série de Neumann convergente $(1-A)^{-1} = \sum_{k=0}^\infty A^k$ convergente si $\|A\| < 1$) et l'application $\mathcal{A}^* \ni A \mapsto A^{-1} \in \mathcal{A}^*$ est différentiable avec $D(X \to X^{-1})_A(B) = -A^{-1}BA^{-1}$.
La composée de deux fonctions $H : E \subset U \to V \subset F$ et $K : F \supset V \to G$ différentiables respectivement en a et $b = f(a)$ est différentiable en a avec

$$D(K \circ H)_a = DK_{H(a)} \circ DH_a$$

A partir de là si $H : E \supset U \to V \subset F$ est un difféomorphisme (i.e. une bijection telle que H^{-1} est différentiable) alors $D(H^{-1})_{H(a)} = DH_a^{-1}$.
Quand $E = \mathbb{R}^m$, $F = \mathbb{R}^n$ et $G = \mathbb{R}^p$ on écrit le produit des matrices jacobiennes

$$\frac{\partial(K \circ H)_k}{\partial x_i}(a) = \sum_{j=1}^n \frac{\partial K_k}{\partial y_j}(H(a))\frac{\partial H_j}{\partial x_i}(a)\,.$$

Formule de Taylor avec reste intégral : Une récurrence avec une intégration par parties donne

$$H(b) = \sum_{k=0}^p \frac{1}{k!}DH_a^k(b-a,\ldots b-a) + \int_0^1 (p+1)(1-t)^p \frac{1}{(p+1)!}DH_{a+t(b-a)}^{p+1}(b-a,\ldots,b-a)\,dt\,.$$

Dans le cas $E \sim \mathbb{R}^m$ l'écriture avec les dérivées partielles s'obtient à partir de

$$D^k H_a(v,\dots,v) = k! \sum_{|\alpha|=k} \frac{1}{\alpha!}\partial_x^\alpha H(a) v^\alpha\,, \quad v^\alpha = v_1^{\alpha_1}\times\cdots\times v_m^{\alpha_m} \text{ avec } v=b-a \quad \alpha! = \alpha_1!\times\cdots\times\alpha_d!$$

Quelques contre-exemples :

a) Fonctions continues non différentiables : $\sqrt{|x|}$, $|x|$ sur $\mathbb{R}$.

b) Fonction admettant des dérivées partielles dans toutes les directions non différentiable : le prolongement par continuité de $\mathbb{R}^2 \ni (x,y) \to \frac{x^2y}{x^2+y^2} \in \mathbb{R}$ au point $(0,0)$.

c) Fonction ne vérifiant pas le Lemme de Schwarz : le prolongement par continuité H de $\mathbb{R}^2 \ni (x,y) \to x^2\sin(xy^{-1})$ pour laquelle $\partial_y\partial_x H(0,0) = 0$ et $\partial_x\partial_y H(0,0)=1$.

2 INégalité des accroissements finis et espaces fonctionnels

L'égalité des accroissements finis est déjà fausse pour une fonction de $\mathbb{R}\to\mathbb{R}^2$ (ex : $H(\theta)=e^{i\theta}\in\mathbb{C}\sim\mathbb{R}^2$, $\theta\in\mathbb{R}$). En revanche on a l'inégalité des accroissements finis qui contient les résultats suivants.

Proposition 2.1. *Si $H\in\mathcal{C}^{p+1}(U;F)$, $p\in\mathbb{N}$, et U est conVexe alors*

$$\|H(b)-\sum_{k=0}^p \frac{1}{k!}D^kH_a(b-a,\dots,b-a)\|_F \le \frac{M_{[a,b]}^{(p+1)}\|b-a\|_E^{p+1}}{(p+1)!} \le \frac{M_U^{(p+1)}\|b-a\|_E^{p+1}}{(p+1)!}\,,$$

en notant $M_X^{(p+1)} = \sup_{x\in X}\|D^{p+1}H_x\|_{\mathcal{L}^{p+1}(E;F)}$.
Si U est un ouvert conNexe (donc connexe par arcs $\mathcal{C}^1$) alors pour tout $a,b\in U$ et tout $\gamma\in\mathcal{C}^1([0,1];U)$ tel que $\gamma(0)=a$ et $\gamma(1)=b$,

$$\|H(b)-H(a)\|_F = \|\int_0^1 DH_{\gamma(t)}\gamma'(t)\ dt\|_F \le L_\gamma M^{(1)}_{\gamma([0,1])}\,, \quad \text{avec } L_\gamma = \int_0^1 \|\gamma'(t)\|_F\ dt\,.$$

Le premier résultat se déduit de la formule de Taylor avec reste intégral et du lemme de la moyenne (Si $\Phi\in\mathcal{C}^0([t_0,t_1],E)$ et $\varphi\in\mathcal{C}^0([t_0,t_1];\mathbb{R})$, φ dérivable sur $]t_0,t_1[$, vérifient $\|\Phi(t)\|_E\le\varphi'(t)$ alors $\|\int_{t_0}^{t_1}\Phi(t)\ dt\|_E \le \varphi(t_1)-\varphi(t_0)$).
De façon générale, ce que l'on fait parfois avec l'égalité des accroissements finis pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ s'obtient à plusieurs variables avec la formule de Taylor avec reste intégral.
Une première conséquence est que, si U est un ouvert connexe, $H\in\mathcal{C}^1(U;F)$ est constante si et seulement si $DH_x=0$ pour tout $x\in U$ (et par conséquent $H\in\mathcal{C}^2(U;F)$ est affine si et seulement si $D^2H_x=0$ pour tout $x\in U$).
Contre-exemple pour la première inégalité dans le cas conNexe non conVexe : $U=\mathbb{R}^2\setminus(-\infty,0]\times\mathbb{R}$, $H(x,y)=\arctan(\frac{y}{x^2+y^2+x}) = \arctan(\theta/2)$ si $(x,y)=(r\cos(\theta), r\sin(\theta))$, et $a=(-1,-\varepsilon)$, $b=(-1,\varepsilon)$, $\varepsilon>0$.

Definition 2.2. *Pour $p\in\mathbb{N}$, note $\mathcal{C}_b^p(U;F)$ l'ensemble des fonctions $\mathcal{C}^p$ sur U dont toutes les différentielles $D^kH : U\to\mathcal{L}^k(E;F)$, $k\le p$, sont uniformément bornées. On munit cet espace de la norme $\|H\|_{\mathcal{C}_b^p} = \sup_{k\le p, x\in U}\|D^kH_x\|_{\mathcal{L}^k(E;F)}$, équivalente à $\sum_{k=0}^p\|H\|_{\mathcal{C}_b^k}$. L'ensemble $\mathcal{C}_b^\infty(U;F)$ est l'intersection $\cap_{k=0}^\infty \mathcal{C}_b^k(U;F)$.*

Proposition 2.3. *Si F est un espace de Banach, $\mathcal{C}_b^p(U;F)$ muni de la norme $\|H\|_{\mathcal{C}_b^p}$ est un espace de Banach. Sous la même hypothèse $\mathcal{C}^\infty(U;F)$ muni de la distance $d(H;K)=\sum_{p=0}^\infty 2^{-p}\frac{\|H-K\|_{\mathcal{C}_b^p}}{1+\|H-K\|_{\mathcal{C}_b^p}}$ est un espace métrique complet.*

Attention $\mathcal{C}_b^\infty(U;F)$ n'a pas de norme naturelle. c'est l'exemple typique d'un espace de Fréchet.
On dit qu'une inégalité (une convergence) est vérifiée localement uniformément sur U si pour tout point a on peut trouver $\delta>0$ telle que l'inégalité (la convergence) est vérifiée uniformément sur $B(a,\delta)\subset U$. Ainsi une suite $(H_\ell)_{\ell\in\mathbb{N}}$ est localement bornée (resp. converge localement vers H) dans $\mathcal{C}_b^p$, si pour tout $a\in U$ il existe $\delta>0$ et $C>0$ tel que pour tout $k\in\mathbb{N}$ $d_{\mathcal{C}_b^p(B(a,\delta);F)}(0;H_\ell)\le C$ (resp. $\lim_{\ell\to\infty} d_{\mathcal{C}_b^p(B(a,\delta))}(H_\ell,H)=0$).
Si E est de dimension finie, "localement uniformément" équivaut à "uniformément sur tout compact inclus dans U".
A partir de l'inégalité des accroissements finis on a les résultats de compacité suivants (qui peuvent être vus comme un cas particulier du théorème d'Ascoli). Il se démontre par récurrence sur p, en partant du cas $p=1$ et utilise un argument d'extraction diagonale (un deuxième argument d'extraction diagonale est utilisé pour le cas $p=\infty$).

Proposition 2.4. *On se place dans un ouvert U connexe, avec E et F de dimension finie, et $p \in \mathbb{N}$, $p \geq 1$. Si $(H_\ell)_{\ell\in\mathbb{N}}$ est une suite de $\mathcal{C}^p(U;F)$ localement bornée dans $\mathcal{C}_b^p$, alors on peut extraire une sous-suite $(H_{\ell_k})_{k\in\mathbb{N}}$ qui converge localement dans $\mathcal{C}_b^{p-1}$. Le résultat est encore vrai pour $p = p-1 = \infty$.*

3 Inversion locale et applications

3.1 Difféomorphisme local, coordonnées locales

Definition 3.1. *On dit que $H : E \supset U \to V \subset F$ est un $\mathcal{C}^k$-difféomorphisme, $1 \leq k \leq \infty$, de U sur V si c'est une bijection telle que $H : U \to V$ et $H^{-1} : V \to U$ sont des applications $\mathcal{C}^k$.*
On dit que $H : U \to V$ est un $\mathcal{C}^k$-difféomorphisme local en $a \in U$, s'il existe deux voisinages ouverts U_a^E de a et V_b^F de $b = H(a)$ tels que $H : U_a^E \to V_b^F$ est un $\mathcal{C}^k$-difféomorphisme.
Si cette dernière propriété est vraie pour tout point $a \in U$, on dit que H est un $\mathcal{C}^k$-difféomorphisme local sur U.

Attention : Un $\mathcal{C}^k$-difféomorphisme local sur U n'est pas forcément un difféomorphisme. Ex : l'application $\mathbb{R} \setminus \{0\} \ni x \to x^2 \in (0, +\infty)$ est un difféomorphisme local qui n'est pas injectif. En fait un difféomorphisme local injectif sur ouvert U est un difféomorphisme de U sur $H(U)$.
L'étude locale des fonctions $\mathcal{C}^k$ en dimension finie est souvent grandement simplifiée en introduisant un bon système de coordonnées.

Definition 3.2. *Si $\dim E = m$ et U est un ouvert de E. Un système de coordonnées locales $\mathcal{C}^k$ centrées au point $a \in U$ est la donnée d'un $\mathcal{C}^k$-difféomorphisme $\varphi : U_0^{\mathbb{R}^m} \to U_a^E$ d'un voisinage ouvert $U_0^{\mathbb{R}^m}$ de 0 dans $\mathbb{R}^m$ dans un voisinage ouvert $U_a^E \subset U$ de a dans E.*

Avec un système de coordonnées locales, on repère alors un point $x \in U_a^0$ par les nombres (coordonnées) $x_1, \ldots, x_m \in \mathbb{R}^m$ tels que $\varphi(x_1, \ldots, x_m) = x$.
Les résultats de composition des fonctions $\mathcal{C}^k$ conduisent tout de suite aux résultats suivants dans les cas $E \sim \mathbb{R}^m$ et $F \sim \mathbb{R}^n$:

- Une fonction $H : U \to V$ est $\mathcal{C}^k$ (resp. un $\mathcal{C}^k$-difféomorphisme local) au voisinage de $a \in U$ avec $b = H(a) \in V$ ssi il existe des systèmes coordonnées locales $\varphi : U_0^{\mathbb{R}^m} \to U_a^E$ et $\psi : U_0^{\mathbb{R}^n} \to U_b^F$ tels que $\psi^{-1} \circ H \circ \varphi : U_0^{\mathbb{R}^m} \to U_0^{\mathbb{R}^n}$ est $\mathcal{C}^k$ (resp. un $\mathcal{C}^k$-difféomorphisme local).
- Comme $D\varphi_0$ et $D\psi_0$ sont inversibles, le rang de DH_a est égal au rang de $D(\psi^{-1} \circ H \circ \varphi)_0$.

3.2 Inversion locale, fonctions implicites

Le théorème suivant donne un moyen simple de vérifier qu'une application $\mathcal{C}^k$ est un difféomorphisme local.

Théorème 3.3. ***Théorème d'inversion locale :*** *Une application $H \in \mathcal{C}^1(U;F)$ est un $\mathcal{C}^1$-difféomorphisme local au point $a \in U$ si et seulement si $DH_a \in \mathcal{L}(E;F)$ est inversible (En particulier E doit être isomorphe à F).*
De façon plus générale $H \in \mathcal{C}^k(U;F)$, $k \geq 1$, alors H est un $\mathcal{C}^k$ difféomorphisme local en a si et seulement si DH_a est inversible.

La preuve de ce théorème repose sur l'étude de la suite récurrente $x_{n+1} = x_n - DH_a^{-1}(y - H(x_n))$, (on remplace l'équation $H(x) = y$ par son approximation $H(x_n) + DH_a(x - x_n) = y$ pour déterminer x_{n+1}) et le théorème de Picard pour les applications contractantes.
Un exemple classique consiste à vérifier que les coordonnées polaires $\phi : (r, \theta) \mapsto (r\cos(\theta), r\sin(\theta))$ donnent un système de coordonnées locales (non centrées) $\mathcal{C}^\infty$ au voisinage de tout point $(x, y) \in \mathbb{R}^2 \setminus \{(0,0)\}$. En effet le déterminant de la matrice jacobienne est r.
Une variante du théorème d'inversion locale est le théorème des fonctions implicites qui permet d'étudier la solution d'une équation dépendant de paramètres.

Théorème 3.4. ***Théorème des fonctions implicites :*** *Nous nous plaçons dans le cas $E = E_1 \oplus E_2$, E_1 et E_2 étant des supplémentaires topologiques avec l'écriture $x = (x_1, x_2) = x_1 + x_2$ et D_{x_j} désignant la différentielle dans la direction E_j. On considère $H \in \mathcal{C}^1(U;F)$ telle que pour $a = (a_1, a_2)$ $H(a) = 0$ et $D_{x_1}H_a \in \mathcal{L}(E_1;F)$ est un isomorphisme. Alors il existe un voisinage $U_{a_1}^{E_1}$ de a_1 dans E_1, un voisinage $U_{a_2}^{E_2}$ de a_2 dans E_2, et une application $K \in \mathcal{C}^1(U_{a_1}^{E_1}; U_{a_2}^{E_2})$ tels que*

$$\left(\begin{array}{l} H(x_1, x_2) = 0 \\ (x_1, x_2) \in U_{a_1}^{E_1} \times U_{a_2}^{E_2} \end{array} \right) \Leftrightarrow (x_2 = K(x_1))$$

De plus la différentielle de K est donnée par

$$\forall t \in U_{a_1}^{E_1}\,, \quad DK_t = -D_{x_1}H_{(t,K(t))} \circ D_{x_2}H_{(t,K(t))}$$

De plus si $H \in \mathcal{C}^k(U;F)$ alors $K \in \mathcal{C}^k(U_{a_1}^{E_1};E_2)$.

3.3 Sous-variétés

Une application la plus importante de ces théorèmes tourne autour de la notion de sous-variété qui permet de comprendre la structure locale d'ensembles courbes en toute dimension. Mais on peut aussi à l'aide du théorème des fonctions implicites étudier toute sorte d'équations dépendant d"'un" paramètre, $x_1 \in E_1$, dont l'inconnue est $x_2 \in E_2$.

Definition 3.5. *On dit qu'une application $H \in \mathcal{C}^k(U;F)$, $k \geq 1$, est une immersion (resp. submersion) au point a si DH_a est injective (resp. surjective).*
On dit qu'une application $H \in \mathcal{C}^k(U;F)$, $k \geq 1$, est un plongement $\mathcal{C}^1$ si c'est une immersion (en tout point de U) injective.

Definition 3.6. *On appelle sous-variété paramétrée $\mathcal{C}^k$, $k \geq 1$ de $\mathbb{R}^n$, de dimension m , l'image d'une application $H \in \mathcal{C}^k(U;\mathbb{R}^n)$, U ouvert de $\mathbb{R}^m$, qui est une immersion sur un ouvert dense $\tilde{U}$ de U . On dit que la sous-variété paramétrée est immersive si $\tilde{U} = U$.*
On appelle espace des vitesses au paramètre $s_0 \in U$, le sous-espace vectoriel $\mathcal{V}_{s_0} = DH_{s_0}(\mathbb{R}^m)$ de $\mathbb{R}^n$.

De façon générale si $H \in \mathcal{C}^1(U;\mathbb{R}^n)$, U ouvert de $\mathbb{R}^m$, que l'ensemble des points $x \in U$ tels que D_xH est de rang maximal (i.e. $\min(m,n)$) est un ouvert.
Sur l'ouvert $\tilde{U}$, la dimension de l'espace des vitesses est toujours m (d'où la notion de dimension). On rappelle qu'il n'y a pas de bonne notion de dimension pour $k = 0$, un exemple étant donné par la "courbe" de Peano qui passe par tous les points d'un carré.
La formule de composition des différentielles donne immédiatement une autre caractérisation de l'espace des vitesses :

$$\mathcal{V}_{s_0} = \Big\{(H \circ \gamma)'(0)\,, \gamma \in \mathcal{C}^{(1)}((-1,1);U)\,, \gamma(0) = s_0\Big\}\,.$$

Definition 3.7. *On dit qu'une partie X de $\mathbb{R}^n$ est une sous-variété $\mathcal{C}^k$, $k \geq 1$, de dimension m si pour tout point $a \in X$ il existe un voisinage $U_a^{\mathbb{R}^n}$ de a dans l'espace ambiant $\mathbb{R}^n$, tel que $U_a^{\mathbb{R}^n} \cap X$ est l'image d'un plongement $\mathcal{C}^k$. Pour un plongement $\mathcal{C}^k$, $H : U_0^{\mathbb{R}^m} \to U_a^{\mathbb{R}^n}$ tel que $H(0) = a$, l'espace tangent à X au point a , noté T_aX , n'est rien d'autre que $T_aX = DH_0(\mathbb{R}^m)$, l'espace des vitesses au paramètre 0 . Le sous-espace affine tangent à X au point a est alors $a + T_aX$.*

Si X est une sous-variété $\mathcal{C}^k$, l'espace tangent au point a , T_aX , ne dépend pas du plongement choisi et sa dimension est m .
A l'aide du théorème des fonctions implicites on a une caractérisation dans l'espace ambiant de la notion de sous-variété.

Proposition 3.8. *Une partie X de $\mathbb{R}^n$ est une sous-variété $\mathcal{C}^k$, $k \geq 1$, de dimension m , si et seulement pour tout point $a \in X$ il existe un système de coordonnées locales dans $\mathbb{R}^n$ centrées au point a , $\varphi : U_0^{\mathbb{R}^n} \to U_a^{\mathbb{R}^n}$, que l'on note (x',x'') avec $x' = (x_1,\ldots,x_m) \in \mathbb{R}^m$ et $x'' = (x_{m+1},\ldots,x_n) \in \mathbb{R}^{m-n}$, et une application $\Psi \in \mathcal{C}^k(U_0^{\mathbb{R}^m};U_0^{\mathbb{R}^{m-n}})$, tels que dans ce système de coordonnées*

$$X \cap \varphi(U_0^{\mathbb{R}^m} \times U_0^{\mathbb{R}^{n-m}}) = \Big\{(x',\Psi(x'))\,, x' \in U_0^{\mathbb{R}^m}\Big\}\,.$$

L'espace tangent n'est autre que $T_aX = \{(t, D\Psi_0(t)), t \in \mathbb{R}^m\}$.

Autrement dit une sous-variété $\mathcal{C}^k$ de $\mathbb{R}^n$ est localement et dans un bon système de coordonnées locales, le graphe d'une application $\mathcal{C}^k$.
Quelques exemples à avoir en tête pour bien distinguer sous-variété paramétrée, qui privilégie l'espace de départ, et sous-variété qui privilégie l'espace ambiant :

- $\mathbb{R} \ni z \mapsto (s,|s|) \in \mathbb{R}^2$ n'est pas une sous-variété paramétrée au voisinage de $s = 0$. L'ensemble $\{(s,|s|)\,, s \in \mathbb{R}\}$ n'est pas non plus une sous-variété car il n'y a pas d'espace tangent en $(0,0)$.
- $\mathbb{R} \ni s \mapsto (s^2,s^3) \in \mathbb{R}^2$ est une sous-variété paramétrée de $\mathbb{R}^2$. Le paramétrage est injectif sur $\mathbb{R}$ mais non immersif au paramètre $s = 0$ et ce n'est pas une sous-variété au voisinage de $(0,0)$ (pas d'espace tangent) .

— $\mathbb{R} \ni s \mapsto (s^3, s^6) \in \mathbb{R}^2$ est une sous-variété paramétrée de $\mathbb{R}^2$. Le paramétrage est injectif sur $\mathbb{R}$ mais non immersif au paramètre $s = 0$. Toutefois l'ensemble $\{(s^3, s^6), s \in \mathbb{R}\} = \{(t, t^2), t \in \mathbb{R}\}$ est une sous-variété de $\mathbb{R}^2$.

— $\mathbb{R} \ni \theta \mapsto (\sin(2\theta)\cos(\theta), \sin(2\theta)\sin(\theta))$, autrement dit $\varrho(\theta) = \sin(2\theta)$ en coordonnées polaires, est une courbe paramétrée immersive mais non injective. La courbe à la forme d'un 8 incliné avec un point double en $(0,0)$ atteint pour $\theta = 0$ avec $\mathcal{V}_0 = \mathbb{R}e_x$ et $\theta = \frac{\pi}{2}$ avec $\mathcal{V}_{\frac{\theta}{2}} = \mathbb{R}e_y$. Ce n'est pas une sous-variété.

— Le cercle donné par le paramétrage $\mathbb{R} \ni \theta \mapsto (\cos(\theta), \sin(\theta)) \in \mathbb{R}^2$ non injectif est une sous-variété paramétrée immersive de $\mathbb{R}^2$. C'est aussi une sous-variété de $\mathbb{R}^2$.

Une partie de l'espace ambiant peut alternativement être vue comme l'image d'un paramètrage ou comme un ensemble de points vérifiant des équations. Dans ce dernier cas on parle alors de sous-variété implicite et leur étude repose encore sur le théorème des fonctions implicites. Ici on notera $\mathbb{R}^m$ l'espace ambiant.

Definition 3.9. *Une sous-variété implicite de U ouvert de $\mathbb{R}^m$ d'équation $H \in \mathcal{C}^k(U; \mathbb{R}^n)$, $m \geq n$, $H = \begin{pmatrix} H_1 \\ \vdots \\ H_n \end{pmatrix}$, est l'ensemble $X_H = \{x \in U, H(x) = 0\} = \{x \in U, H_1(x) = \ldots = H_n(x) = 0\}$.*
On dit qu'un point $x \in X_H$ est un point régulier de la sous-variété implicite (plus précisément de la paire (X_H, H)) si DH_x a le rang maximal n. Sinon on dit que c'est un point singulier de (X_H, H).

On note que l'ensemble des points réguliers de (X_H, H) est un ouvert (éventuellement vide, mais la plupart du temps un ouvert dense) de X_H muni de la topologie induite par celle de $\mathbb{R}^m$.

Proposition 3.10. *Si $X_H = \{x \in U; H(x) = 0\}$ est une sous-variété implicite avec $H \in \mathcal{C}^k(U; \mathbb{R}^n)$, $k \geq 1$, et $a \in X_H$ est un point régulier de (X_H, H) alors X_H est au voisinage de a une sous-variété $\mathcal{C}^k$ de U, de dimension $m - n$. De plus l'espace tangent en a est alors donné par $T_a X_H = \{v \in \mathbb{R}^m, DH_a v = 0\}$ et l'espace affine tangent à X_H au point a est égal à $a + T_a X_H = \{x \in \mathbb{R}^m, DH_a(x - a) = 0\}$.*

Quelques exemples

— La sphère $\mathbb{S}^{m-1} = \{x \in \mathbb{R}^m, x_1^2 + \cdots + x_m^2 = 1\}$ (en particulier le cercle $\mathbb{S}^1$ pour $m = 2$), est une sous-variété $\mathcal{C}^\infty$ de $\mathbb{R}^m$ de dimension $m - 1$. Le sous-espace tangent au point $a \in \mathbb{S}^{m-1}$ n'est rien d'autre que $T_a \mathbb{S}^{m-1} = a^\perp$.

— Les sous-variétés implicites $\{(x, y) \in \mathbb{R}^2, xy = 0\}$ et $\{(x, y) \in \mathbb{R}^2, y^2 - x^3 = 0\}$ ont toutes les deux un point singulier en $(0,0)$. Dans le premier cas c'est un point double et dans le deuxième cas un point de rebroussement.

— La sous-variété implicite $\{(x, y) \in \mathbb{R}^2, x^2 = 0\}$ ou plus précisément la paire (X_H, H) avec $H(x, y) = x^2$ n'a que des points singuliers. Mais on voit que si on écrit $X_H = X_K$ avec $K(x, y) = x$, alors il n'y a que des points réguliers. L'ensemble $X_H = X_K$ est bien une sous-variété (c'est la droite $x = 0$) de $\mathbb{R}^2$. Le bon choix d'un système d'équations n'est pas toujours évident et rentre, quand on travaille avec des fonctions "algébriques", dans le cadre de la géométrie algébrique.

4 Optimisation, calcul variationnel

L'objectif consiste à avoir des conditions nécessaires et des conditions suffisantes pour détecter les extrêma locaux ou globaux d'une fonction à valeur réelle.

Definition 4.1. *Soit $H : E \supset U \to \mathbb{R}$ une fonction définie sur une partie U de E et soit $a \in U$. On dit que*

— *a est un minimum local (resp. maximum local) si il existe un voisinage ouvert $U_a^E \subset U$ de a tel que*
$$\forall x \in U_a^E, \quad H(a) \leq H(x) \quad (resp. H(x) \geq H(a)).$$
On dit a est un extrêmum local si c'est un minimum ou un maximum. On parle de minimum (maximum, extrêmum) strict si les inégalités sont strictes pour $x \neq a$. $H(a)$ est la valeur minimale (resp. maximale).

— *a est un minimimum (resp. maximum) global si on peut prendre $U_a^E = U$ dans le cas précédent.*

— *a est un point critique de H, sous l'hypothèse supplémentaire que H est différentiable sur l'ouvert U, si $DH_a = 0 \in \mathcal{L}(E, \mathbb{R})$.*

Quand elle est définie, la différentielle seconde $D^2 H_a$ qui est une forme bilinéaire symétrique de $E \times E$ dans $\mathbb{R}$ est appelée Hessienne de H au point a. On dit qu'elle est positive si $D^2 H_a(v, v) \geq 0$ pour tout $v \in E$.

4.1 Optimisation sur un ouvert

Proposition 4.2. ***Conditions nécessaires :*** *Soit $H : E \supset U \to \mathbb{R}$ différentiable sur U et $a \in U$.*
Si a est un extrêmum local de H alors c'est un point critique.
Si de plus H est deux fois différentiable en a et si a est un minimum (resp. maximum) local alors se rajoute la condition $D^2H_a(v,v) \geq 0$ (resp. ≤ 0) pour tout $v \in E$.

Proposition 4.3. ***Condition suffisante :*** *Si on suppose $H \in \mathcal{C}^2(E;\mathbb{R})$, $a \in U$, $DH_a = 0$ et il existe $\alpha > 0$ tel que $D^2H_a(v,v) \geq \alpha\|v\|_E^2$ (resp. $-\alpha\|v\|_3$), alors a est un minimum (resp. maximum) local de H .*

Si E est de dimension finie, la deuxième condition consiste à dire que la Hessienne est définie positive (resp. négative).
L'optimisation globale est souvent étudiée sous des hypothèses de convexité (ou concavité). Rappelons les différentes caractérisations d'une fonction convexe.

Definition 4.4. *Une fonction $H : E \to \mathbb{R}$ est convexe si*

$$\forall x, y \in E\,, \forall t \in [0,1]\,, H((1-t)x + ty) \leq (1-t)H(x) + tH(y)\,.$$

Si H est différentiable sur E cela équivaut à

$$\forall x, y \in E\,, H(y) - H(x) \geq DH_x(y-x)$$

ou encore

$$\forall x, y \in E\,, \quad (DH_y - DH_x)(y-x) \geq 0\,.$$

Une fonction différentiable est dite α-convexe pour $\alpha > 0$ si

$$\forall x, y \in E\,, \quad (DH_y - DH_x)(y-x) \geq \alpha\|y-x\|_E^2\,.$$

En terme de différentielle seconde, une fonction $H \in \mathcal{C}^2(E;\mathbb{R})$ est convexe si $D^2H_x \geq 0$ pour tout $x \in E$ et elle est α-convexe si $D^2H_x(v,v) \geq \alpha\|v\|_E^2$ pour tout $x, v \in E$.

Proposition 4.5. *Pour une fonction convexe $H : E \to \mathbb{R}$ un minimum local est un minimum global.*
Si la fonction H est strictement convexe (inégalité stricte pour $t \in]0,1[$) un minimum (il peut ne pas y en avoir, ex : e^x) est forcément unique.
En dimension finie, une fonction α-convexe admet un unique minimum.

La dernière assertion, peut se généraliser en dimension infinie.

4.2 Optimisation sous-contraintes, optimisation sur une sous-variété

On se place ici dans le cas $E = \mathbb{R}^m$ (même si certains résultats peuvent se généraliser). On considère la sous-variété implicite $X_K = \{x \in U, K(x) = 0\} = \{x \in U, K_1(x) = \ldots = K_n(x) = 0\}$ avec $K \in \mathcal{C}^k(U;\mathbb{R}^n)$, $k \geq 1$. Le problème de minimisation (resp. maximisation) de la fonction $H : U \to \mathbb{R}$, $H \in \mathcal{C}^k(U;\mathbb{R})$, sous les contraintes $K_1 = 0, \ldots, K_d = 0$, consiste à étudier

$$\min_{x \in X_K} H(x) \quad \text{resp.} \max_{x \in X_K} H(x)\,.$$

Definition 4.6. *On dit qu'un point régulier $a \in X_K$ de (X_K, K) est un point critique de H relatif à X_K , si $DH_a \in Vect(DK_{1,a}, \ldots, DK_{n,a})$, autrement dit s'il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que*

$$DH_a = \lambda_1 DK_{1,a} + \cdots + \lambda_n DK_{n,a} \quad \text{dans } \mathcal{L}(E;\mathbb{R})\,.$$

Les coefficients $\lambda_1, \ldots, \lambda_n$ sont alors appelés multiplicateurs de Lagrange pour la fonction H au point a associés aux contraintes $K_1 = 0, \ldots, K_n = 0$.
Si de plus $H \in \mathcal{C}^2(U,\mathbb{R})$ et $K \in \mathcal{C}^2(U;\mathbb{R}^n)$, on appelle Hessienne H au point critique relatif a , relative aux contraintes $K_1 = \ldots = K_n = 0$, la différentielle seconde $D^2\tilde{H}_a$ avec $\tilde{H} = (H - \lambda_1 K_1 \cdots - \lambda_n K_n)$, les λ_i étant les multiplicateurs de Lagrange.

Proposition 4.7. ***Conditions nécessaire :*** *Si $a \in X_K$ est un point régulier de (X, K) et si c'est un minimum local de $H\big|_{X_K}$, i.e. un minimum de H relatif à X_K , alors c'est un point critique relatif.*
Si de plus $H \in \mathcal{C}^2(U;\mathbb{R})$ et $K \in \mathcal{C}^2(U;\mathbb{R}^n)$ alors la Hessienne relative en a doit vérifier $D^2\tilde{H}_a(v,v) \geq 0$ pour tout $v \in T_aX_K$.

Proposition 4.8. ***Condition suffisante :*** *Si $a \in X_K$ est un point régulier de (X, K) et un point critique de H relatif à X_K , avec $H \in \mathcal{C}^2(U;\mathbb{R})$, $K \in \mathcal{C}^2(U;\mathbb{R}^n)$ et si la Hessienne relative vérifie $D^2\tilde{H}_a(v,v) \geq \alpha\|v\|^2$ pour tout $v \in T_aX_K$ avec $\alpha > 0$, alors H est un minimum local de $H\big|_{X_K}$.*

5 Equations différentielles

5.1 Existence et unicité locale

Une équation différentielle porte sur une fonction $y \in \mathcal{C}^k(I;E)$ avec I intervalle de $\mathbb{R}$ et E espace de Banach et s'écrit de façon générale

$$\forall t \in I, G(t, y(t), y'(t), \ldots, y^{(k)}(t)) = 0\,.$$

L'ordre de l'équation différentielle est k le plus haut degré de dérivation. Un problème de Cauchy est une équations différentielle à laquelle on rajoute les conditions initiales au temps $t = t_0$:

$$y(t_0) = y_0\,, \ldots\,, y^{(k-1)}(t_0) = y_0^{(k-1)}\,,$$

$y_0, \ldots, y_0^{(k-1)}$ étant des vecteurs fixés de E. Si $G \in \mathcal{C}^1(I \times U^{E^{k+1}})$ avec $U^{E^{k+1}}$ ouvert de E^{k+1} et si on peut trouver $y_0^{(k)} \in E$ tel que $G(t_0, y_0, \ldots, y_0^{(k-1)}, y_0^{(k)}) = 0$ et tel que $D_{y^{(k)}}G(t_0, y_0\,, \ldots\, y_0^{(k)}) \in \mathcal{L}(E)$ est inversible le théorème des fonctions implicites ramène localement le problème à l'étude de

$$y^{(k)} = K(t, y(t), \ldots, y^{(k-1)}(t)) \quad y(t_0) = y_0 \ldots y^{(k-1)(t_0)} = y_0^{(k-1)}\,.$$

Enfin en posant $Y(t) = \begin{pmatrix} y(t) \\ y^{(1)}(t) \\ \vdots \\ y^{(k-1)}(t) \end{pmatrix}$ on se ramène à

$$\begin{cases} Y'(t) = H(t, Y(t)) \\ Y(t_0) = Y_0\,, \end{cases} \tag{5.1}$$

avec $H(t,Y) = \begin{pmatrix} y^{(1)}(t) \\ \vdots \\ y^{(k-1)}(t) \\ K(t, y(t), \ldots, y^{(k-1)}(t)) \end{pmatrix}$ et $Y_0 = \begin{pmatrix} y_0 \\ \vdots \\ y_0^{(k-1)} \end{pmatrix} \in E^k$. Dans la suite on travaille avec cette dernière forme en remplaçant E^k par E.

Le théorème fondamental qui repose encore sur le théorème de Picard pour les applications contractantes est le théorème de Cauchy-Lipschitz qui donne un résultat d'existence et d'unicité locale, ainsi qu'un résultat de régularité par rapport aux données.

Théorème 5.1. *Si l'application $H : I \times U \to E$, U ouvert de E contenant a et I intervalle ouvert contenant t_0, est continue et localement uniformément Lipschitzienne en Y, alors il existe un intervalle ouvert J, $t_0 \in J \subset I$ et un voisinage ouvert $U_a^E \subset U$ tel que pour tout $Y_0 \in U_a^E$, le problème de Cauchy* (5.1) *admet une unique solution $Y \in \mathcal{C}^k(J_0;E)$. De plus l'application qui à $(t_0, Y_0) \in I_0 \times U_a^E$ associe la solution $Y \in \mathcal{C}^k(I_0;E)$ est continue.*

Si E est de dimension finie, localement uniformément Lipschitzienne en Y, revient à dire uniformément Lipschitzienne en Y sur tout compact de $I \times E$.

Une condition suffisante est $H \in \mathcal{C}^1(I \times U)$, "localement" permettant de se ramener à des boules qui sont des convexes.

Enfin le point clé de ce théorème consiste à faire un argument de point fixe dans $\mathcal{C}^0(J_0;E)$ à l'aide du théorème de Picard pour résoudre la forme intégrale de l'équation

$$Y(t) = Y_0 + \int_{t_0}^{t} H(s, Y(s))\, ds\,.$$

5.2 Intervalle maximal

Sous les hypothèses du théorème de Cauchy-Lipschitz, si $Y_1 \in \mathcal{C}^1(J_1;E)$ et $Y_2 \in \mathcal{C}^1(J_2;E)$ sont deux solutions de $Y' = H(t,Y)$ sur les intervalles ouverts respectifs J_1 et J_2, tels que $Y_1(t_0) = Y_2(t_0) = Y_0$ pour un certain $t_0 \in J_1 \cap J_2$, alors il existe une unique solution de l'équation différentielle $Y \in \mathcal{C}^1(J_1 \cup J_2;E)$ de $Y' = H(t,Y)$ telle que $Y(t_0) = Y_0$ et de plus $Y\big|_{J_1} = Y_1$ et $Y\big|_{J_2} = Y_2$. Cela conduit à la définition suivante.

Definition 5.2. *Sous les hypothèses du théorème de Cauchy-Lipschitz, on note pour $t_0 \in I$ et $Y_0 \in U$, $J(t_0, Y_0)$ l'intervalle maximal sur lequel le problème de Cauchy* (5.1) *admet une unique solution $Y \in \mathcal{C}^1(J(t_0, Y_0);E)$.*

Proposition 5.3. *L'intervalle maximal est un intervalle ouvert* $J(t_0, Y_0) =]T_-, T_+[$ *, et pour tout compact* $K \subset U$ *il existe* $T_{K,,-} > T_-$ *et* $T_{K,+} < T_+$ *tel que la solution de* (5.1) *vérifie*

$$\forall t \in]T_-, T_+[\setminus [T_{K,-}, T_{K,+}] \, , Y(t) \notin K \, .$$

Dans le cas où $I = \mathbb{R}$ et $U = E$, avec E de dimension finie, cela a pour conséquence, $J(t_0, Y_0) =]T_-, T_+[$ avec : $\lim_{t \to T_\pm} \|Y(t)\| = +\infty$ ou $T_\pm = \pm\infty$.
Attention, il ne faut pas perdre de vue que ces résultats ne sont valables que sous les hypothèses du Théorème de Cauchy-Lipschitz. En particulier sans ces hypothèses (notamment le caractère localement uniformément Lipschitien en Y) il peut exister plusieurs voire une infinité de solutions sur l'intervalle $\mathbb{R}$ tout entier. L'exemple classique est $y' = 2\sqrt{|y|}$ pour laquelle les hypothèses sont valables tant que $y \neq 0$. Mais toutes les fonctions données par $y(t) = t^2$ pour $t \geq 0$, $y(t) = 0$ pour $t \in [-C, 0]$ et $y(t) = -(t+C)^2$ pour $t \leq -C$ sont des solutions pour la donnée initiale $y(1) = 1$.
Si la fonction $(t, y) \mapsto H(t, y)$ a une croissance sous-linéaire par rapport à y, le lemme de Gronwall ci-dessous conduit à des résultats d'existence et d'unicité globale (sous les hypothèses du théorème de Cauchy-Lipschitz).

Lemme 5.4. ***Lemme de Gronwall :*** *Si* $y \in \mathcal{C}^1(J; E)$ *,* J *intervalle de* $\mathbb{R}$ *, vérifie*

$$\forall t \in J \, , \quad \|y'(t)\| \leq \alpha(t) + \beta(t)\|y(t)\|_E$$

avec $\alpha, \beta \in \mathcal{C}^0(J; \mathbb{R}_+)$ *, alors on a*

$$\forall t \in J \, , \quad \|y(t)\|_E \leq e^{\int_{[t_0,t]} \beta(s) \, ds} \|y(t_0)\| + \int_{[t_0,t]} e^{\int_{[s,t]} \beta(u) \, du} \alpha(s) \, ds$$

avec la convention $\int_{[t_0,t_1]} = \int_{\min\{t_0,t_1\}}^{\max\{t_0,t_1\}}$.

5.3 Equations différentielles linéaires

Un cas particulier important de la théorie générale est celui des équations différentielles et systèmes différentiels linéaires pour lesquels on peut à la fois utiliser le théorème de Cauchy-Lipschitz et le lemme de Gronwall.

— Pour une application $A \in \mathcal{C}^0(J; \mathcal{L}(E))$ et pour $t_0 \in J$, $Y_0 \in E$, on considère le problème de Cauchy $Y'(t) = A(t)Y$ avec $Y(t_0) = Y_0 \in E$. Alors pour $t_0 \in J$ et pour tout $Y_0 \in E$, l'intervalle maximal est $J(t_0, Y_0) = J$. Pour tout $t_0 \in J$, on note $\Phi(t, t_0) : E \to E$ qui à $Y_0 \in E$ associe $Y(t)$, la valeur en t de la solution du problème de Cauchy. Alors pour tout $t_0, t \in J$, l'application $\Phi(t, t_0)$ est linéaire continue inversible, d'inverse $\Phi(t_0, t)$. Pour tout $t_0 \in J$, l'application $E \ni Y_0 \mapsto Y \in \mathcal{C}^1(J; E)$ avec Y solution du problème de Cauchy, est un isomorphisme d'espace de Banach de E sur $\mathcal{S} = \left\{Y \in \mathcal{C}^1(J; E) \, , Y' = A(t)Y\right\}$.

— Pour un système différentiel affine $Y'(t) = A(t)Y + B(t)$ avec $A \in \mathcal{C}^0(J; E)$ et $B \in \mathcal{C}^0(J; E)$, on a aussi $J(t_0, Y_0) = J$. La solution du problème de Cauchy avec $Y(t_0) = Y_0$ est donnée par la formule de Duhamel qui généralise la méthode dite de "variation de la constante" : $Y((t)) = \Phi(t, t_0)Y_0 + \int_{t_0}^t \Phi(t, s)B(s) \, ds$, oú $\Phi(t, t_0)$ défini plus haut est associé au cas $B = 0$.

— Pour un système différentiel linéaire d'ordre k sous la forme

$$Y^{(k)}(t) + A_{k-1}(t)Y^{(k-1)}(t) + \cdots + A_0(t)Y(t) = 0$$

on se ramène avec la méthode usuelle au cas $k = 1$ en posant

$$\mathcal{Y} = \begin{pmatrix} Y \\ Y^{(1)} \\ \vdots \\ Y^{(k-1)} \end{pmatrix} \quad , \quad \mathcal{Y}' = \mathcal{A}(t)\mathcal{Y} \quad , \quad \mathcal{A}(t) = \begin{pmatrix} 0 & \mathrm{Id}_E & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \mathrm{Id}_E \\ -A_0(t) & -A_1(t) & \ldots & -A_{k-1}(t) \end{pmatrix}$$

— Un cas particulier est le cas d'une équation différentielle scalaire ($E = \mathbb{K}$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) d'ordre k : L'application qui aux données initiales à $t = t_0$, $(y_0, \ldots, y_0^{(k-1)}) \in \mathbb{K}^k$, associe $(y(t), y'(t), \ldots, y^{(k-1)}(t)) \in \mathbb{K}^k$ est un isomorphisme vectoriel de $\mathbb{K}^k$. De même l'application qui à $(y_0, \ldots, y_0^{(k-1)})$ associe la solution $y \in \mathcal{C}^k(J; \mathbb{K})$ est un isomorphisme d'espace vectoriel de $\mathbb{K}^k$ sur $\mathcal{S} = \left\{y \in \mathcal{C}^k(J; \mathbb{K}) \, , y^{(k)} + a_{k-1}(t)y^{k-1} + \cdots a_0(t)y = 0\right\}$. Ainsi l'ensemble des solutions est un espace vectoriel de dimension k. Si $y_1, \ldots, y_k \in \mathcal{S}$, ces solutions forment une base de $\mathcal{S}$ si et seulement si le Wronskien en t_0 (n'importe quel $t_0 \in J$) donné par

$$\mathcal{W}(y_1, \ldots, y_k)(t_0) = \det \begin{pmatrix} y_1(t_0) & \ldots & y_k(t_0) \\ y_1'(t_0) & \ldots & y_k'(t_0) \\ \vdots & \vdots & \vdots \\ y_1^{(k-1)}(t_0) & \ldots & y_k^{(k-1)}(t_0) \end{pmatrix} ,$$

ne s'annule pas.

— Un cas particulier à savoir traiter est le cas des systèmes différentiel linéaire à coefficients constants : La solution de $Y' = AY$ avec $A \in \mathcal{M}_n(\mathbb{K})$ est donnée par $Y(t) = e^{(t-t_0)A}Y_0$, où l'exponentielle d'une matrice se calcule par trigonalisation. On rappelle en particulier que si $A = \lambda \,\mathrm{Id} + N$ avec $N^r = 0$ on a

$$e^{tA} = e^{t\lambda} \sum_{k=0}^{r-1} \frac{(tN)^k}{k!}\,.$$

— Pour les équations différentielles linéaires scalaires à coefficients constants $y^{(k)} + a_{k-1}y^{(k-1)} + \ldots + a_0 y = 0$ on peut utiliser la méthode des systèmes avec la matrice $\begin{pmatrix} 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 \\ -a_0 & -a_1 & \ldots & -a_{k-1} \end{pmatrix}$ ou plus simplement résoudre l'équation caractéristique

$$X^k + a_{k-1}X^{k-1} + \cdots + a_1 X + a_0 = 0\,.$$

Si λ est une solution de l'équation caractéristique de multiplicité μ, on met dans la base de l'espace des solutions les fonctions linéairement indépendantes $e^{\lambda t}$, $e^{\lambda t}(t - t_0)$, ..., $e^{\lambda t}(t - t_0)^{\mu - 1}$. '

6 Remarques finales et références

Ce résumé n'est pas un cours.
Il y a quelques exemples et contre-exemples classiques, mais il est important de pratiquer ces notions en faisant des exercices.
Il manque aussi de figures (Ne pas négliger de faire des dessins quand on étudie un problème de topologie, calcul différentiel ou de géométrie différentielle).
Pour plus de détails, vous pouvez consulter les notes de cours manuscrites du cours de Francis Nier sur https ://www.math.univ-paris13.fr/ nier/ensP13.html
Vous pouvez aussi consulter les livres suivants adaptés au niveau Licence

— Cours de Topologie, Calcul Différentiel, Equations différentielles, pour la licence MAF, J.C. Yoccoz, Université de Paris-Sud Editions.

— Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation, de François Rouvière aux Editions Cassini.