Introduction.

1 Introdution

1.1 Dériver n'importe quoi ?

La théorie des distributions a été initialement créee pour répondre à des problèmes sur les équations aux dérivées partielles (EDP). Sans rentrer dans les détails, considérons l'exemple de l'équation des ondes sur $\mathbb{R}$:

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \tag{1}$$

on peut montrer que toutes les solutions sont de la forme $f(x-ct)+g(x+ct)$ avec f et g régulières (c'est-à-dire ici de classe C^2). Pourtant, une telle formule avec f et g seulement continues a toujours un sens et correspond effectivement à une solution physique. Il faudrait donc étendre l'équation des ondes dans un sens qui permettent d'obtenir aussi ces solutions.

De même, l'équation de transport (linéaire) est :

$$\frac{\partial \rho}{\partial t} - v.\vec{\nabla}\rho = 0.$$

Elle modélise l'évolution d'une densité de matière sous un champ de vitesses v. Une analyse rapide donne des solutions du type $f(x-tv)$ qui a un sens même pour f seulement continue. Mais ici, on a même envie de choisir des conditions initiales de type Dirac en 0 (c'est-à-dire que la masse initiale est intégralement en 0 à l'instant 0). On veut alors pouvoir définir la notion de dérivée pour une mesure.

Ce genre d'opérations est d'ailleurs utilisée de manière intuitive (et non rigoureuse) par les physiciens dès les années 1920. Ce sera Schwartz dans les années 50 qui apportera une réponse théorique satisfaisante avec la théorie des distributions (qui lui vaudra la médaille Fields). Bien sûr, ses idées sont dans la lignée de celles d'autres chercheurs : Leray, Sobolev, Courant... La force de la théorie des distributions est sa grande généralité et sa relative simplicité. Elle permet, à peu de frais, de définir la plupart des opérations courantes de l'analyse (dérivation, produits, convolutions, support ...) sur une très grande classe d'objet (cette classe contient les fonctions localement intégrables, les mesures ...).

Ne faisons pas durer le suspense de manière trop artificielle et voyons dans le cas de l'équation des ondes l'idée centrale pour définir une *solution faible* à cette équation, (ou encore une solution *au sens des distributions*). L'idée est la suivante : soit $u(x,t)$ une solution alors pour toute fonction $\varphi : \mathbb{R}^2 \to \mathbb{R}$, de classe C^2 et à support compact, on a en multipliant (1) et en intégrant que :

$$\int (\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2})\varphi(x,t)dxdt = 0.$$

Réciproquement, si u vérifie la condition ci-dessus pour toutes les fonctions de classe C^2 à support compact, alors u est solution (exercice). A l'aide d'intégrations par parties élémentaires, on obtient par ailleurs que l'équation ci-dessus s'interprète en :

$$\forall \varphi \in C_c^2, \int (\frac{\partial^2 \varphi}{\partial t^2} - c^2 \frac{\partial^2 \varphi}{\partial x^2})u(x,t)dxdt = 0. \tag{2}$$

Mais cette fois, plus aucune régularité n'est exigée sur u (u peut être dans L^1_{loc} ou une mesure de Radon et (2) a toujours un sens très clair). Ce sera donc notre définition d'une solution faible : une solution de (2). Si u est C^2, les deux notions coincident mais on obtient bien de nouvelles solutions non C^2 (qui ont un sens physique). De manière plus formelle, on voit les solutions comme des éléments du dual topologique d'un bon espace fonctionnel et on fait porter les notions de dérivation sur cet espace par transposée (une intégration par partie en l'occurence). Reste à trouver cet espace (spoiler : c'est $\mathcal{D}$).

Pour finir, notons que la grande généralité des distributions ne doit pas faire oublier que la plupart des problèmes se pose dans un cadre beaucoup plus restreint (mesure de Radon, fonction L^p, espaces de Sobolev, distributions tempérées...). Les outils des distributions, trop généraux dans ces cadres, donnent souvent des résultats moins optimaux, moins constructifs que des preuves directes. Par contre, les idées sous-jacentes à la théorie des distributions sont, elles, cruciales et permettent de forger l'intuition en mathématiques.

1.2 Quelques espaces fonctionnels classiques

1.2.1 Espace L^p

Soit $\Omega \subset \mathbb{R}$, on rappelle la définition

$$L^p(\Omega) := \{f : \Omega \to \mathbb{C},\ f \text{ mesurable},\ \int_\Omega |f|^p d\lambda < +\infty\},$$

où $1 \leq p < \infty$. Le point de vue étant celui de l'intégration, on identifie deux fonctions qui coincident en dehors d'un ensemble de mesure nulle. Alors, $L^p(\Omega)$ muni de la norme $\|f\|_p := \int_\Omega (|f|^p d\lambda)^{\frac{1}{p}}$ est un espace de Banach.

Pour $p = +\infty$, on a :

$$L^\infty(\Omega) := \{f : \Omega \to \mathbb{C},\ f \text{ mesurable},\ \exists c \geq 0,\ xp.p,\ |f(x)| \leq c\},$$

et avec la même identification qu'au dessus, L^∞ est un espace de Banach lorsque l'on munit de la norme infinie (qui est l'inf des c de la définition).

On rappelle l'inégalité de Hölder : "si $f \in L^p$ et $g \in L^q$ avec $1/p + 1/q = 1$ alors $fg \in L^1$ et $\|fg\|_1 \leq \|f\|_p \|g\|_q$." Soit donc $p \in [1, +\infty]$ et q son exposant conjugué ($1/p + 1/q = 1$). On fixe $g \in L^q$. On a alors clairement que :

$$\begin{aligned} &L^p \to \mathbb{R} \\ \Lambda_g :\ &f \mapsto \int fg \end{aligned}$$

définie une forme linéaire continue sur L^p. On en déduit que $(L^p)' \subset L^q$ (vérifier que $g \mapsto \Lambda_g$ est injective). En fait, il y a l'égalité sauf dans le cas où $p = +\infty$ (on verra le cas $p = +\infty$ plus tard). Noter le rôle central de L^2 dans la théorie qui est son propre dual. L'exercice suivant est crucial (même s'il est écrit ici dans le cadre non optimal des espaces de Banach).

Exercice 1. Soient $(E, \|.\|_E)$ et $(F, \|.\|_F)$ deux espaces de Banach tels que $E \subset F$ et l'injection $E \to F$ est continue. Que signifie cette phrase en termes de norme ?

On suppose que E est dense dans F. Montrer que $F' \subset E'$.

Exercice 2. a) On suppose que Ω est borné. Pourquoi a-t-on $L^p \subset L^q$ pour $p \geq q$? Que dire par rapport à l'exercice précédent ? b) Donner un contre-exemple où l'on n'a pas $L^p \subset L^q$ pour $p \geq q$.

Une classe d'ensemble importants sont les espaces L^p_{loc}. Par définition, $\varphi \in L^p_{loc}(\Omega)$ si $\varphi \in L^p_{loc}(K)$ pour tout compact K de Ω. L'exercice précédent montre que L^1_{loc} est le plus grand de tous ces espaces (c'est donc l'hypothèse la plus faible d'intégrabilité que l'on puisse mettre). Notons enfin que si Ω n'est pas compact, les espaces L^p_{loc} ne sont pas des espaces de Banach. Ce sont en revanche des espaces de Fréchet (autrement dit, ils sont métrisables et complets).

Bien sûr, on n'est pas obligé de prendre comme mesure la mesure de Lebesgue et d'autres mesures donnent d'autres espaces.

Profitons que nous sommes sur les espaces L^p pour définir une notion de convergence faible. Soit $p < \infty$, on considère alors une suite (f_n) de L^p, on dira que (f_n) *converge faiblement dans* L^p vers $f \in L^p$ si :

$$\forall g \in L^q,\ 1/p + 1/q = 1, \int f_n g \to \int fg.$$

Ceci est noté par $f_n \rightharpoonup f$. Autrement dit, en notant $\langle f_n, g\rangle := \int f_n g$ le crochet de dualité, dire que $f_n \rightharpoonup f$ correspond à dire que la suite de forme linéaire $L^q \to \mathbb{R}$ qui à g associe $\langle f_n, g\rangle$ converge ponctuellement.

Il est aisé de vérifier que la convergence forte implique la convergence faible. La réciproque est fausse : en effet, pour tout $1 < p < +\infty$, de toute suite (f_n) bornée, on peut extraire une sous suite qui converge faiblement vers $f \in L^p$ et on a

$$\|f\|_p \leq \liminf \|f_n\|_p.$$

Il s'agit de résultat du cours d'analyse fonctionnel, les preuves et les détails seront donc dans ce cours-là (et pas ici).

1.2.2 Fonctions continues, dérivables et holomorphes

Soit Ω un ouvert de $\mathbb{R}^n$. Rappelons la notion de support d'une fonction. C'est une notion cruciale en théorie des distributions :

Definition 1.1. *Soit U le plus grand ouvert de Ω sur lequel f est identiquement nulle. On définit* $\mathrm{supp}(f) := U^c$.

On peut considérer les différents espaces de fonctions continues :
- $C(\Omega)$, l'ensemble des fonctions continues sur Ω que l'on pourrait munir de la norme infinie mais ce n'est pas un bon choix car une fonction continue n'est pas nécessairement bornée (sur un compact oui mais pas sur un ouvert). On choisit plutôt la topologie de la convergence uniforme sur les compacts (ce qui fait de nouveau de $C(\Omega)$ un espace de Fréchet).
- $C_b(\Omega)$, l'ensemble des fonctions continues et bornées sur Ω. Muni de la norme infinie, c'est un espace de Banach.
- $C_0(\Omega)$, l'ensemble des fonctions continues et qui tendent vers 0 sur le bord de Ω. Muni de la norme infinie, c'est un espace de Banach.
- $C_c(\Omega)$ (ou parfois $\mathcal{D}_0$), l'ensemble des fonctions continues à support compact. Muni de la norme infinie, ce n'est pas un espace de Banach. En effet son adhérence (dans $C_b(\Omega)$) est $C_0(\Omega)$. Une manière d'éviter ce problème consiste à durcir la condition de convergence : on dit qu'une suite (f_n) de $C_c(\Omega)$ converge si elle converge pour la norme infinie et s'il existe un compact fixé K qui contient le support de tous les f_n (en déduire que f est à support dans K). On peut montrer qu'il existe encore une topologie qui définit cette convergence mais cela ne nous intéressera pas dans ce cours.

On a les inclusions ensemblistes $C_c \subset C_0 \subset C_b \subset C$, on vérifie même qu'elles sont continues (c'est immédiat).

Soit $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ un multiindice (ou un n-uplet). On utilisera dans ce cours les conventions usuelles :
- $|\alpha| = \sum_{1 \leq i \leq n} \alpha_i$, la longueur de α,
- $\partial^\alpha = (\frac{\partial}{\partial x_1})^{\alpha_1} (\frac{\partial}{\partial x_2})^{\alpha_2} \ldots (\frac{\partial}{\partial x_n})^{\alpha_n}$.

Soit $k \in \mathbb{N}^*$ un ordre de dérivabilité, on note alors $C^k(\omega)$ l'espace des fonctions dont toutes les dérivées partielles de longueur $\leq k$ sont continue et pour K compact dans Ω on note :

$$\|f\|_{C^k(K)} = \max_{\alpha,\ |\alpha| \leq k} \|\partial^\alpha f\|_\infty.$$

On définit alors comme précédemment $C^k_b(\Omega)$, $C^k_0(\Omega)$ et $C^k_c(\Omega)$ (attention, les hypothèses de bornitude ou de limite portent sur f et ses dérivées jusqu'à l'ordre k, la condition de support est toujours exigée pour la convergence dans $C^k_c(\Omega)$).

Je ne résiste pas à parler de l'espace suivant qui ne nous concerne pas vraiment dans ce cours... Soit $H(\Omega)$ l'espace des fonctions holomorphes sur Ω (cette fois bien sûr, il s'agit d'un ouvert de $\mathbb{C}^n$). On le munit de la topologie de la convergence uniforme sur les compacts. Cela en fait un espace de Montel : de toute suite bornée sur les compacts, on peut extraire une sous suite qui converge. Par contre, l'espace $H_0(\Omega)$ n'a pas d'intérêt (il ne contient que la fonction nulle, car...). Cela nous amène d'ailleurs à la remarque suivante :

Remark 1.2. *Il n'y a pas de fonction non nulle analytique et à support compact dans un ouvert.*

Exercice 3. Soit B la boule unité de $\mathbb{C}$. On considère les espaces $H(B)$, $H_2 := H(B) \cap L^2$, $H_b := H(B) \cap L^\infty$ et $C := \{\sum_n a_n z^n,\ \sum |a_n| < +\infty\}$. Les espaces $H(B)$ et H_b son muni de la topologie de la convergence uniforme. On met sur H_2 la norme euclidienne et enfin pour C, on prend $\| \sum_n a_n z^n \| = \sum |a_n|$.

1. Soit $K \subset B$ un compact, montrer qu'il existe une constante C_K telle que pour tout f dans H_2 :
$$\forall z \in K, |f(z)| \leq C_K \|f\|_2.$$
En déduire que H_2 est un espace de Hilbert et que l'inclusion $H_2 \subset H(B)$ est continue.
2. Montrer que l'on a les injections continues suivantes :
$$C \subset H_b \subset H_2 \subset H(B)$$
3. Quel est le dual topologique de C ? Passer au dual les inclusions précédentes.

La morale de l'exercice précédent est que les fonctions analytiques sont rigides et que donc leur dual est petit. Jusqu'ici, nous n'avons pas encore considérez les duals des espaces de fonctions continues. Il va s'agir d'espaces de mesures et c'est l'objet de la section suivante.

1.3 Mesures

Une mesure positive μ est une fonction d'une tribu $\mathcal{T}$ sur un espace X à valeurs dans $[0, +\infty]$ qui vérifie les propriétés suiivantes :
- $\mu(\emptyset) = 0$
- $\sum_{i \in I} \mu(A_i) = \mu(\cup_{i \in I} A_i)$ pour toute union dénombrable disjointe d'éléments de $\mathcal{T}$.

Pour Ω un ouvert de $\mathbb{R}^n$, on considèrera uniquement le cas où $\mathcal{T} = \mathcal{B}$ est la tribu des boréliens (engendré par les ouverts). De telles mesures sont dites *mesures de Borel.* On demande en plus que μ soit finie sur les compacts. De telles mesures sont automatiquement extérieurement et intérieurement régulières.

Rappelons alors la définition de mesure de Radon.

Definition 1.3. *Nous appellerons mesure signée μ, ou mesure de Radon, la différence de deux mesures de Borel, finies sur les compacts. Une telle mesure définit, par restriction aux parties d'un compact fixé de Ω une fonction σ-additive d'ensembles, à valeurs dans $\mathbb{R}$. On peut la décomposer de manière unique en la différence de deux mesures positives singulières l'une à l'autre, appellées sa partie positive μ^+ et sa partie négative μ^-. La variation totale d'une mesure signée μ sur un ensemble $A \subset \Omega$ est définie par $\|\mu\|_A := \mu^+(A) + \mu^-(A)$.*

On note alors les ensembles suivants :

- *$M(\Omega)$ l'ensemble des mesures de Radon finies (c'est un espace de Banach).*
- *$M_{loc}(\Omega)$ l'ensemble de toutes les mesures de Radon (c'est un espace de Fréchet).*
- *$M(A)$ l'ensemble de toutes les mesures de Radon porté par $A \subset \Omega$.*

Observons que L^1_{loc} s'injecte dans $M_{loc}(\Omega)$ via l'injection $f \mapsto (A \mapsto \int_A f)$. De même, L^1 s'injecte dans $M(\Omega)$. Néanmoins, $M_{loc}(\Omega)$ est plus grand que L^1_{loc}. On a par exemple δ_x, le Dirac en x qui définit une mesure de Radon positive finie. Une combinaison linéaire de Dirac ou même une mesure portée par un ensemble de mesure de Lebesgue nulle (un hyperplan par exemple) donnent d'autres exemples.

Soit μ une mesure de Radon finie, alors elle définit (de manière injective) un élémént du dual topologique de $C_b(\Omega)$ par : $f \mapsto \int f d\mu$. Le théorème de Riesz, capital (et admis), montre que les mesure de Radon sont en fait exatement le dual des fonctions continues bornées, plus précisément :

Theorem 1.4 (Théorème de Riesz)**.** *Soit $\Omega \subset \mathbb{R}^n$, on peut alors identifier*

1. *une forme linéaire positive sur $C_c(\Omega)$ avec une mesure de Radon et l'espace $M_{loc}(\Omega)$ avec l'espace des formes linéaires sur $C_c(\Omega)$ telles que pour tout compact K il existe une constante C_K telle que pour toute fonction $\varphi \in C_c(\Omega)$ telle que* $\mathrm{supp}(\varphi) \subset K$*, on ait*

$$|\Lambda\varphi| \leq C_K \|\varphi\|_\infty.$$

2. *$M(\Omega)$ avec le dual (topologique) $C_0(\Omega)$*
3. *$M(K)$ avec le dual topologique de $C(k)$.*

Intéressons nous au premier point. Il est clair qu'une mesure de Radon μ vérifie :

$$|\int \varphi d\mu| \leq \|\mu\|_K \|\varphi\|_\infty.$$

Ce que dit le théorème de Riesz c'est que réciproquement, une forme qui vérifie ceci est représentable par une mesure de Radon. Cette inégalité est en fait exactement la continuité sur $C_c(\Omega)$ au sens de la convergence définie plus haut. En effet, si Λ vérifie une telle inégalité et si $\varphi_n \to \varphi$ in C_c, alors il existe K compact tel que le support de toutes les φ_i et φ sont dans K et alors la convergence est juste la continuité de Λ sur $C(K)$. A l'inverse, si Λ ne vérifie pas la condition, on dispose d'un compact fixé tel que pour tout $n \in \mathbb{N}$ il existe $\varphi_n \in C_c(K)$ telle que $\|\varphi_n\|_\infty = 1$ et $\Lambda(\varphi_n) > n$. La suite φ_n/n tend alors vers 0 dans $C_c(\Omega)$ mais $\Lambda(\varphi_n)$ ne tend pas vers $\Lambda(0) = 0$.

Ici encore, on a une notion de convergence faible, on dit que (μ_n) une suite de $M_{loc}(\Omega)$ (resp. $M(\Omega)$) converge faiblement dans $M_{loc}(\Omega)$ (resp. $M(\Omega)$) que l'on note $\mu_n \rightharpoonup to\mu$ si pour toute φ dans $C_c(\Omega)$ (resp. $C_0(\Omega)$) on a $\langle \mu_n, \varphi \rangle \to \langle \mu, \varphi \rangle$. Dans ce cas, on a alors automatiquement que μ est dans $M_{loc}(\Omega)$ (resp. $M(\Omega)$) et que $\|\mu\|_A \leq \liminf \|\mu_n\|_A$. On a le théorème suivant :

Theorem 1.5. *De toute suite de mesure localement équibornée, on peut extraire une sous suite qui converge faiblement.*

On peut à l'aide de ce théorème obtenir des résultats très puissants comme par exemple :

Theorem 1.6. *Soit $f : K \to K$ une application continue, alors il existe une mesure de probabilité (et donc positive) invariante par f, c'est-à-dire :*

$$\forall A \in \mathcal{B}, \mu(A) = \mu(f^{-1}(A))$$

Quelques remarques pour conclure : la théorie de la mesure est bien plus agréable sur les compacts. En terme de dualité, les mesures de Radon sont les éléments du dual des fonctions continues. Néanmoins, le calcul différentiel a lieu sur des ouverts et la théorie des distributions se place sur les ouverts de $\mathbb{R}^n$. On a vu que plus l'espace est petit, plus son dual est gros. L'idée de Schwartz est donc de prendre l'espace des fonctions C_c^∞ muni d'une convergence proche de celle sur C_c.