47.

48.

49.

50.

51. Mesures accélérées pour physiciens vigoureux

Ma (la ?) Bible, c'est le Briane-Pagès. On suit les chapitres 4 à 13 et en fait une exposition sommaire et accélérée pour le camarade physicien Raphaël Decan de Chatouville. On ne trouvera donc pas plus que des idées de preuve qui, au demeurant, sont très naturelles mais peuvent parfois s'avérer astucieuses. N'est présent que le minimum vital. Moralement, on va passer de ces horribles sommes de Riemann relatives à une subdivision $S(f,\sigma,\chi_\sigma) := \sum_{i=1}^n (a_i - a_{i-1}) f(\chi_\sigma)$ à une bien belle théorie qui s'assume parfaitement.

On rappelle simplement qu'un ensemble X est dit dénombrable s'il existe une injection de X dans $\mathbf{N}$, i.e. $\mathrm{card}(X) \leq \mathrm{card}(\mathbf{N})$. Exercice : montrer que $\mathbf{N}^2$ est dénombrable. Indication : penser à un escargot se déplaçant sur $\mathbf{N}\times\mathbf{N}$.

Il peut également être bon d'avoir en tête qu'un espace métrique est séparable si et seulement si il est à base dénombrable d'ouverts. C'est le cas de notre cher L^2. On y reviendra pas beaucoup mais bien avoir en tête que l'espace L^2 est vraiment très agréable (notamment pour faire de la mécanique quantique), c'est un joli Hilbert.

51.1. **Définition.** Soit X un ensemble. On appelle tribu (ou σ-algèbre) sur X toute famille $\mathscr{A}$ de parties de X vérifiant :

(i) $\emptyset \in \mathscr{A}$,

(ii) si $A \in \mathscr{A}$, alors $A^c \in \mathscr{A}$ [stabilité par complémentaire],

(iii) si $(A_n)_n$ est une suite d'éléments de $\mathscr{A}$, alors $\bigcup_{n\geq 1} A_n \in \mathscr{A}$ [stabilité par union *dénombrable*].

Le couple $(X,\mathscr{A})$ est appelé **espace mesurable**. On construira plus tard les espaces mesurés.

E.g. tribu grossière : $\mathscr{A} = \{\emptyset, X\}$; tribu triviale : $\mathscr{A} = \mathscr{P}(X)$; pour $A \subset X$ fixé, $\mathscr{A} = \{\emptyset, A, A^c, X\}$. Une intersection de tribus est une tribu. La plus petite tribu au sens de l'inclusion contenant $\mathscr{E}$ est appelée **tribu engendrée** par $\mathscr{E}$, on la note $\sigma(\mathscr{E})$. Par exemple, la tribu engendrée par les singletons de X est $\mathscr{A} = \{A \in \mathscr{P}(X) : A \text{ est dénombrable ou } A^c \text{ est dénombrable}\}$. Une tribu particulièrement importante (en particulier en probabilités) est la **tribu image**. Soit $f : X \to Y$ et $\mathscr{A}$ une tribu sur X. On appelle tribu image de $\mathscr{A}$ par f la tribu sur Y définie par $\mathscr{B} := \{B \in \mathscr{P}(Y) : f^{-1}(B) \in \mathscr{A}\}$. Attention, la terminologie "tribu *image*" est trompeuse. On remarquera que $\{f(A), A \in \mathscr{A}\}$ n'est pas une tribu sur Y en général.

51.2. **Définition.** La tribu engendrée la plus importante est sans doute la tribu borélienne. Soit $(X, \mathscr{O}(X))$ un espace topologique [14]. La tribu engendrée par les ouverts de X est appelée **tribu borélienne** de X. On la note $\mathscr{B}(X) := \sigma(\mathscr{O}(X))$.

Lorsque l'on parlera de tribu borélienne, on commettra l'abus fréquent qu'est celui de considérer $X = \mathbf{R}$. (Remarque : il n'est pas évidentissime d'exhiber un ensemble non borélien, i.e. qui n'est pas un élément de la tribu borélienne.) Pourquoi est-ce utile de se coltiner des tribus engendrées ? Si le monde est bien fait, on s'attend à ce que bien comprendre un "comportement" sur les éléments qui engendrent la tribu permette de comprendre totalement ce même "comportement" sur toute la tribu. Le monde est heureusement bien fait (détail mathématique, cf le **lemme de transport** : si $f : X \to Y$ et $\mathscr{E} \subset \mathscr{P}(Y)$, alors $\sigma(f^{-1}(\mathscr{E})) = f^{-1}(\sigma(\mathscr{E}))$). Dans le cas précis de $X = \mathbf{R}$, on comprend aisément qu'il suffit de saisir ce qu'il se passe sur les intervalles (car $\mathbf{R}$ est à base dénombrable d'ouverts). Remarquer que l'on a donc (en utilisant la densité de $\mathbf{Q}$ dans $\mathbf{R}$) :

$$\mathscr{B}(\mathbf{R}) = \sigma(\{[a,+\infty[, a \in \mathbf{Q}\}) = \sigma(\{]a,+\infty[, a \in \mathbf{Q}\}) = \sigma(\{]-\infty, a[\in \mathbf{Q}\}) = \sigma(\{]-\infty, a] \in \mathbf{Q}\}). \tag{222}$$

Au fond, ce que l'on veut c'est intégrer des choses. Autrement dit, les mesurer. L'intégrale de Riemann est légèrement embêtante car elle ne permet pas de mesurer des objets auxquels naturellement on saurait donner une valeur (e.g. l'indicatrice de $\mathbf{Q}$ est nulle *presque partout* donc...). On veut même un peu plus ! On veut élargir la classe des fonctions intégrables [15] tout en espérant obtenir d'agréables théorèmes de convergence !

51.3. **Définition.** Soient $(X,\mathscr{A})$ et $(Y,\mathscr{B})$ deux espaces mesurables. Une fonction $f : (X,\mathscr{A}) \to (Y,\mathscr{B})$ est **mesurable** si pour tout $B \in \mathscr{B}$, on a $f^{-1}(B) \in \mathscr{A}$. Si on considère des tribus boréliennes sur X et Y, on parlera plutôt de **fonctions boréliennes**.

14. L'ensemble $\mathscr{O}(X)$ correspond évidemment aux ouverts de X.

15. Faut pas croire, mais on perd du monde en chemin. Par exemple $x \mapsto \sin(x)/x$ est Riemann-intégrable mais pas Lebesgue-intégrable sur $[0,1]$.

Une fonction constante est évidemment mesurable, de même l'indicatrice d'un ensemble $A \in \mathscr{A}$ est mesurable. Un petit jeu sur les ouverts de $\mathbf{R}$ et leur tribu engendrée nous donne que toute fonction continue est borélienne. On a les propriétés classiques : la composition, somme, produit de fonctions mesurables est mesurable[16], le min, le max, la valeur absolue de fonctions mesurables est mesurable. Que demande donc le peuple ? Et bien il aimerait bien que tout se passe bien en dimension supérieure (par exemple quand il joue avec des lois de couples en probabilités). C'est effectivement le cas : soit $f := (f_1, f_2) : (X, \mathscr{A}) \to (\mathbf{R}^2, \mathscr{B}(\mathbf{R}^2))$. Alors f est mesurable si et seulement si f_1 et f_2 le sont.

Spoilons un petit peu. On dispose de fonctions mesurables mais pour l'instant on sait pas vraiment comment les manier (en revanche, on a vu qu'elles avaient tout un tas de propriétés habituelles). L'idée est de se ramener à des suites d'éléments d'une sous-classe de fonctions mesurables qui, modulo quelques hypothèses, tend (en croissant, voire même uniformément) vers une fonction mesurable. C'est surtout le mode de convergence qui nous intéresse (et qui fait la force de la théorie de l'intégration de Lebesgue). La construction est alors très naturelle. On part de **fonctions mesurables étagées** (i.e. qui ne prennent qu'un nombre fini de valeurs) et... c'est tout.

51.4. **Lemme.** (fondamental d'approximation) *Soit $f : (X, \mathscr{A}) \to \mathbf{R}$ mesurable. Il existe alors une suite $(f_n)_{n\in\mathbf{N}}$ de fonctions étagées telle que $\lim_n f_n(x) = f(x)$ pour tout $x \in X$. Si de plus $f \geq 0$, on peut choisir la suite croissante et positive (i.e. $\forall n \geq 1, 0 \leq f_n \leq f_{n+1}$). Si f est bornée, on peut choisir la suite de façon à avoir une convergence uniforme (i.e. $\lim_n \sup_{x\in X} |f_n(x) - f(x)| = 0$).*

Dans les faits, les fonctions étagées sont de la forme suivante pour I **fini** : $f = \sum_{i\in I} \alpha_i 1_{A_i}$. Remarquons que la traditionnelle image du découpage régulier selon l'axe des abscisses chez Riemann est ici remplacé par un découpage régulier selon l'axe des ordonnées chez Lebesgue. La puissance de ce lemme est qu'il permet de ramener un problème à des considérations sur des indicatrices. En effet, supposons que l'on veuille prouver une propriété sur une classe de fonctions mesurables alors : on le prouve d'abord pour une indicatrice, puis une somme finie d'indicatrice (i.e. une fonction étagée) et ensuite on conclut par le lemme fondamental d'approximation. C'est par exemple comme cela que l'on montre la propriété d'additivité de l'intégrale. Bref, continuons notre chemin.

51.5. **Définition.** Soit $(X, \mathscr{A})$ un espace mesurable. On appelle **mesure** sur $(X, \mathscr{A})$ toute application $\mu : \mathscr{A} \to \mathbf{R}_+ \cup \{+\infty\}$ vérifiant :

(i) $\mu(\emptyset) = 0$,

(ii) si $(A_n)_n$ est une suite d'éléments de $\mathscr{A}$ deux à deux disjoints, alors : $\mu\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} \mu(A_n)$.Cette dernière propriété est appelée σ-additivité (le σ référant évidemment à l'aspect dénombrable).

Lorsque $\mu(X) < +\infty$ on parle de mesure finie. En particulier, si $\mu(X) = 1$, on parle de mesure de probabilité (d'ailleurs, comment ne pas reconnaître ci-dessous les axiomes définissant une probabilité ? !). La mesure nulle valant 0 pour tout élément A de $\mathscr{P}(X)$ est effectivement une mesure. De même pour la mesure de Dirac ou encore la mesure de comptage valant $\mathrm{card}(A)$ si A est fini et $+\infty$ autrement. Néanmoins une mesure surpasse toutes les autres, la voici.

51.6. **Théorème.** *Il existe une unique mesure sur $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$ dite **mesure de Lebesgue** sur $\mathbf{R}^d$, notée λ_d, vérifiant :*

(i) $\lambda_d([0,1]^d) = 1$,

(ii) $\forall a \in \mathbf{R}^d, \forall A \in \mathscr{B}(\mathbf{R}^d), \lambda_d(a + A) = \lambda_d(A)$.

Lorsque d vaut 1 on se permet de noter la mesure de Lebesgue λ. L'invariance par translation est beaucoup plus importante que ce qu'elle paraît. La construction et surtout l'unicité de cette mesure sont un petit peu pénibles et techniques à démontrer. (Il existe par ailleurs tout un procédé pour construire des mesures à coup de π-systèmes, classes monotones et théorème de Carathéodory, faudrait que je creuse cet aspect là un jour, Briane et Pagès prennent à peu près 20 pages bien denses pour le faire !) On esquive les propriétés de régularité intérieure et extérieure de la mesure de Lebesgue mais faut essentiellement avoir en tête les principales propriétés d'une mesure (quelconque) : **croissante pour l'inclusion**, **forte additivité** (i.e. $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ pour tous A et B dans $\mathscr{A}$), **sous-additivité** (i.e. pour $(A_n)_{n\in\mathbf{N}}$ une suite quelconque d'éléments de $\mathscr{A}$, on a $\mu(\bigcup_n A_n) \leq \sum_n \mu_{A_n}$), **continuité à gauche** (i.e. pour une suite *croissante* $(A_n)_{n\in\mathbf{N}}$ d'éléments de $\mathscr{A}$, on a $\mu(\bigcup_n A_n) = \lim_n \mu(A_n)$), **continuité (partielle) à droite** (i.e. pour une suite *décroissante* $(A_n)_{n\in\mathbf{N}}$ d'éléments de $\mathscr{A}$

16. L'ensemble des fonctions mesurables de $(X, \mathscr{A})$ dans $(\mathbf{K}, \mathscr{B}(\mathbf{K}))$ muni des opérations usuelles $+, \cdot, \times$ sur les fonctions forme une $\mathbf{K}$-algèbre.

telle qu'il existe $n_0 \in \mathbf{N}$ avec $\mu(A_{n_0}) < +\infty$, on a $\mu(\bigcap_n A_n) = \lim_n \mu(A_n)$).

La construction de la mesure de Lebesgue prolonge les résultats intuitifs déjà connus. Par exemple, pour $I \subset \mathbf{R}$ un intervalle, on a $\lambda(I)$ qui est égal à la longueur de I. De même $\lambda_d(P)$ est égal au volume de P, où P est un pavé, c'est-à-dire que P est un produit cartésien d'intervalle. On trouve immédiatement que $\lambda_d(P)$ est alors égal au produit des longueurs de chaque intervalle constituant le pavé.

Je ne sais pas si cela te sera utile mais la mesure de Lebesgue est en particulier une **mesure de Borel**, c'est-à-dire que pour tout compact K de $\mathbf{R}^d$, $\mu(K)$ est fini. En effet, tout compact de $\mathbf{R}^d$ étant borné, il existe $M \in \mathbf{R}$ tel que $K \subset [-M, M]^d$ et donc $\lambda(K) \leq (2M)^d < \infty$ (remarquons que l'on a utilisé la croissance pour l'inclusion de la mesure).

Suivant les idées précédemment introduites, on comprend aisément que pour définir **l'intégrale de Lebesgue**, il suffit de savoir le faire sur les fonctions étagées (en réalité, il suffit de savoir le faire sur les fonctions étagées *positives*). L'intégrale d'une fonction étagée (positive) f par rapport à la mesure μ est définie par :

(223) $$\int_X f \mathrm{d}\mu = \sum_{\alpha \in f(X)} \alpha \mu(\{f = \alpha\}) \in \mathbf{R}_+ \cup \{+\infty\}.$$

On trouvera également les notations suivantes $\int_X f(x)\mathrm{d}\mu(x)$ ou $\int_X f(x)\mu(\mathrm{d}x)$. Par exemple, si μ est la mesure de Dirac en a, on trouve que l'intégrale d'une fonction f par rapport à μ vaut $f(a)$. C'est physiquement logique. Pour la mesure de comptage m, lorsque $X = \mathbf{N}$, on constate que l'intégration de Lebesgue généralise toute la théorie des séries bien connue. En effet : $\int_{\mathbf{N}} f(n)\mathrm{d}m(n) = \sum_{n\in\mathbf{N}} f(n)$. L'intégrale a le bon goût d'être **additive** (i.e. $\int_X (f+g)\mathrm{d}\mu = \int_X f\mathrm{d}\mu + \int_X g\mathrm{d}\mu$), **croissante** (i.e. $f \leq g \longrightarrow \int_X f\mathrm{d}\mu \leq \int_X g\mathrm{d}\mu$) et **homogène** (i.e. $\forall \in \mathbf{R}$, $\int_X \lambda f\mathrm{d}\mu = \lambda \int_X f\mathrm{d}\mu$). On peut également intégrer sur une partie de X, il suffit alors de multiplier f par une indicatrice de cette partie. Il faut donc désormais définir l'intégrale d'une fonction mesurable non nécessairement étagée.

51.7. **Définition.** Si f est mesurable, on définit $\displaystyle\int_X f\mathrm{d}\mu = \sup\left\{\int_X \varphi\mathrm{d}\mu, \varphi \leq f, \varphi \text{ mesurable étagée (positive)}\right\}$.

Un théorème centralissime (qui sert d'ailleurs à démontrer beaucoup beaucoup des théorèmes suivants) est celui de Beppo Levi, également appelé théorème de convergence monotone.

51.8. **Théorème.** *Soit $(f_n)_{n\in\mathbf{N}}$ une suite croissante de fonctions mesurables*[17] *(i.e. $0 \leq f_n \leq f_{n+1}$), alors $f := \lim\limits_n f_n$ est mesurable et $\int_X f\mathrm{d}\mu := \int_X \lim_n f_n\mathrm{d}\mu = \lim_n \int_X f_n\mathrm{d}\mu$.*

Dit autrement, on va pouvoir permuter limite et intégrale ! Une première application de ce grandissime théorème est la suivante : pour f mesurable, on a l'équivalence suivante : $\int_X f\mathrm{d}\mu = 0 \iff \mu(f \neq 0) = 0$. On obtient donc une caractérisation des tels f plus fine que ce que l'on aurait obtenu pour f continue (par morceaux) avec la théorie de Riemann ! Avec Lebesgue, on sent bien que f va être nul mais pas partout... f va être nul (μ-)presque partout ! Plus généralement, une fonction va vérifier une propriété $\mathscr{P}$ μ-presque partout si la mesure des éléments ne vérifiant pas cette propriété est nulle ! Le μ-p.p. (abrégé de "presque partout") c'est vraiment cool, ça permet d'oublier des points *qui ne comptent pas vraiment.* Par exemple, pour f et g deux fonctions mesurables, si $f = g$ μ-p.p., alors $\int_X f\mathrm{d}\mu = \int_X g\mathrm{d}\mu$. Logique.

On dit qu'une fonction est **μ-intégrable** si $|f|$ est μ-intégrable, i.e. $\int_X |f|\mathrm{d}\mu < +\infty$. On note comme suit l'ensemble des fonctions μ-intégrables définies sur X à valeurs dans $\mathbf{K}$:

(224) $$\mathscr{L}^1_{\mathbf{K}}(X, \mathscr{A}, \mu) := \{f : (X, \mathscr{A}) \to (\mathbf{K}, \mathscr{B}(\mathbf{K})) \text{ mesurable et } \mu\text{-intégrable}\}.$$

Par commodité et comme on ne travaillera essentiellement qu'avec des boréliens, on prendra l'habitude de ne noter que $\mathscr{L}^1(\mu)$. Octroyons nous le droit une fois de déroger à cette commodité ! Si on munit l'espace mesurable $(\mathbf{N}, \mathscr{P}(\mathbf{N}))$ de la mesure de comptage m, alors on retombe évidemment sur nos pattes : $\mathscr{L}^1(m) = \ell^1(m)$ l'espace des suites absolument convergentes.

Remarquons qu'on sent qu'il commence à se passer de drôles de choses par rapport à ce que l'on connaît. En effet, sur $\mathscr{L}^1(\mu)$, $||f||_1 := \int_X |f|\mathrm{d}\mu$ n'est pas une norme mais une semi norme (on n'a pas $||f||_1 = 0 \implies f = 0$ mais seulement f nul p.p.). Mais néanmoins, pour des fonctions bien régulières (continues par exemple), rien de nouveau sous le soleil ! Dans le cas des fonctions Riemann intégrables définies sur un compact, théorie de Lebesgue et Riemann coïncident.

17. Remarquer que je ne précise pas sur quoi qui est mesurable tant cela apparaît évident, puis essentiellement, on se limite à la tribu borélienne.

Passons désormais sur un arsenal de théorèmes de convergence (et c'est typiquement ce qui fait la force de la théorie de Lebesgue, rappelons le !). Il va bien falloir garder à l'esprit que le **théorème de Beppo Levi** également appelé **théorème de convergence monotone** permet de prouver les résultats qui vont suivre.

51.9. **Lemme.** (Fatou) *Soit $(f_n)_{n\in\mathbf{N}}$ une suite de fonctions mesurables* ***positives****, alors :* $\int_X \liminf f_n \mathrm{d}\mu \leq \liminf \int_X f_n \mathrm{d}\mu$.

51.10. **Théorème.** (**convergence dominée**) *Soit $(f_n)_{n\in\mathbf{N}}$ une suite d'éléments de $\mathscr{L}^1(\mu)$ vérifiant :*

(i) lorsque n tend vers l'infini, $f_n(x)$ converge μ-p.p.,

(ii) il existe $g \in \mathscr{L}^1(\mu)$ telle que, pour tout $n \geq 1$, on a $|f_n(x)| \leq g(x)$ μ-p.p.

Alors, il existe $f \in mathscrL^1(\mu)$ tel que $f_n(x)$ converge vers $f(x)$ μ-p.p. et on a :

$$\lim_n \int_X f_n \mathrm{d}\mu = \int_X f \mathrm{d}\mu \quad \text{et même} \quad \lim_n \int_X |f - f_n| \,\mathrm{d}\mu = 0. \tag{225}$$

Encore une fois, on peut agréablement permuter limite et intégrale ! En bidouillant un petit peu ce théorème et celui de convergence monotone, on obtient que pour des fonctions φ_n mesurables, si $\sum_{n\geq 1}\int_X |\varphi_n| \mathrm{d}\mu < +\infty$, alors les fonctions φ_n, $\sum_{n\geq 1}|\varphi_n|$ et la fonction définie μ-p.p. $\sum_{n\geq 1}\varphi_n$ sont μ-intégrables. En bonus, on a :

$$\int_X \left(\sum_{n\geq 1}\varphi_n\right)\mathrm{d}\mu = \sum_{n\geq 1}\int_X \varphi_n \mathrm{d}\mu. \tag{226}$$

Grâce à cela, on peut montrer le lemme de Borel-Cantelli.

51.11. **Lemme.** *Soit $(A_n)_{n\in\mathrm{N}}$ une famille de parties de $\mathscr{A}$, alors :*

$$\sum_{n\geq 1}\mu(A_n) < +\infty \implies \mu(\limsup A_n) = 0. \tag{227}$$

Terminons cette série de lemme-théorème par deux résultats qui ne sont que des applications du théorème de convergence dominée.

51.12. **Théorème.** (continuité sous le signe intégrale) *Soit $f : \mathbf{R}\times X \to \mathbf{R}$ et $\tilde{u} \in \mathbf{R}$. Si :*

(i) pour tout $u \in \mathbf{R}$, l'application $x \mapsto f(u,x)$ est mesurable,

(ii) l'application $u \mapsto f(u,x)$ est continue en $\tilde{u}$ μ-p.p.,

(iii) il existe $g \in \mathscr{L}^1(\mu)$ telle que $\forall u \in \mathbf{R}, |f(u,x)| \leq g(x)$ μ-p.p.

Alors la fonction $F(u) := \int_X f(u,x)\mu(\mathrm{d}x)$ est définie en tout point $u \in \mathbf{R}$ et est continue en $\tilde{u}$.

51.13. **Théorème.** (dérivation sous le signe intégrale) *Soit $I \subset \mathbf{R}$ un intervalle ouvert non vide, $f : I \times X \to \mathbf{R}$ et $\tilde{u} \in I$. Si :*

(i) pour tout $u \in I$, $f(u,\cdot) \in \mathscr{L}^1(\mu)$,

(ii) $\frac{\partial f}{\partial u}(\tilde{u},x)$ existe pour μ presque tout x,

(iii) il existe $g \in \mathscr{L}^1(\mu)$ telle que $\forall u \in I, |f(u,x) - f(\tilde{u},x)| \geq g(x)|u-\tilde{u}|$ pour μ presque tout x.

Alors la fonction $F(u) := \int_x f(u,x)\mu(\mathrm{d}x)$ est définie en tout point u de I, et est dérivable en $\tilde{u}$ de dérivée $F'(\tilde{u}) = \int_X \partial_u f(\tilde{u},x)\mu(\mathrm{d}x)$.

On peut désormais considérer le "récital" de théorie de la mesure terminé. [18] On va considérer les espaces canoniques dans lesquels se déploie la théorie de l'intégration... sachant que l'on en a déjà vu un, le beau mais ténébreux $\mathscr{L}^1(\mu)$. En fait, j'ai menti, ce ne sont pas eux les *L ronds* les bons espaces mais les *L droits*. Venons y petit à petit.

51.14. **Définition.** Pour tout réel $p > 0$, on définit l'espace suivant :

$$\mathscr{L}^p_{\mathbf{K}}(X,\mathscr{A},\mu) := \left\{ f : (X,\mathscr{A}) \to (\mathbf{K},\mathscr{B}(\mathbf{K})) \text{ mesurable et } \int_X |f|^p \mathrm{d}\mu < +\infty \right\}. \tag{228}$$

18. Pour un point de vue de dynamicien sur le sujet, se référer au chapitre 0 du Mañé, *Ergodic theory and Differentiable dynamics.* Merci Caleb pour la référence !

Une nouvelle fois, le contexte rendant claire la situation, on prendra l'habitude de ne noter que $\mathscr{L}^p(\mu)$. Cet espace est évidemment un **espace vectoriel** (on verra ensuite qu'il a quelques propriétés supplémentaires qui le rendent enviable à merci à beaucoup de ses confrères ; par ailleurs, bienvenue dans l'analyse dans des espaces de dimension infinie !). On fera bien attention à ne pas trop croire en ses rêves : en général, on n'a pas $0 < p \leq q \implies \mathscr{L}^q(\mu) \subset \mathscr{L}^p(\mu)$ (pour un exemple [19]). Il est nécessaire (mais non suffisant ! penser à la mesure de comptage) que μ soit de masse finie (c'est par exemple le cas pour une mesure de probabilité).

Dans l'optique de faire de l'analyse, on se souvient du crédo de J. Dieudonné résumant en ces mots la démarche d'un analyste : *majorer, minorer, approximer.* Présentons donc les deux principales inégalités pour la (semi-)**norme** $\mathscr{L}^p$ usuelle $||f||_p := \left(\int_X |f|^p \mathrm{d}\mu\right)^{1/p}$. On verra comment passer d'une semi norme à une norme, ce n'est pas très important pour l'instant.

51.15. **Proposition.** (**Inégalité de Hölder**) *Soient f et g mesurables, p et q des exposants conjugués (i.e. $1/p + 1/q = 1$). Si $f \in \mathscr{L}^p(\mu)$ et $g \in \mathscr{L}^q(\mu)$, alors $fg \in \mathscr{L}^1(\mu)$ et $||fg||_1 \leq ||f||_p||g||_p$. On a égalité si et seulement si il existe $(\alpha, \beta) \in \mathbf{R}_+^2 \setminus \{(0,0)\}$ tels que $\alpha|f|^p = \beta|g|^q$ μ-p.p.*

Dans le cas où $p = q = 2$, on retrouve la célèbre inégalité de Cauchy-Schwarz. Bien noter que l'on ne demande pas aux exposants conjugués d'être forcément des entiers ! (Dans les faits, on considèrera les p et q **réels** tels que $1 \leq p \leq q \leq +\infty$, mais bien faire attention aux cas $p = 1$ ou $p = +\infty$.)

51.16. **Proposition.** (**Inégalité de Minkowski**) *Si $p \in [1, +\infty[$, alors $(\mathscr{L}^p(\mu), || \cdot ||_p)$ est un espace vectoriel (semi-)normé. En particulier $\forall f, g \in \mathscr{L}^p(\mu), ||f + g||_p \leq ||f||_p + ||g||_p$.*

Ces deux inégalités sont absolument fondamentales ! (Une troisième est fondamentale lorsque l'on dispose d'une mesure de masse finie : l'inégalité de Jensen. Je ne la présente pas car je doute que tu aies souvent à traiter des inégalités de convexité, au besoin, tu sais que ça existe.)

Bon, pourquoi je m'embête depuis tout à l'heure à dire que l'on n'a pas une norme mais une semi-norme ? Simplement car $||f||_p = 0$ n'est pas équivalent à $f = 0$. Tout ce que l'on peut dire c'est que f va être nulle **presque partout**. Donc, sans surprise, pour obtenir une norme on va devoir raisonner sur un représentant d'une classe d'équivalence. La relation d'équivalence est toute trouvée : être égal presque partout. On quotiente $\mathscr{L}^p(\mu)$ par cette relation d'équivalence et l'on définit l'espace vectoriel quotient obtenu comme étant $\mathrm{L}^p(\mu)$. Prudence est mère de sûreté... Les éléments de cet espace sont des classes d'équivalence, donc dire que deux éléments sont égaux dans $\mathrm{L}^p(\mu)$ est un peu casse gueule. On commettra un nouvel abus qu'est celui d'oublier que l'on travaille avec des représentants de classes d'équivalence. On se dira naïvement que ce sont des éléments tout ce qu'il y a de plus normal (et dans l'absolue majorité des cas, ça ne pose pas de soucis).

L'espace vectoriel normé ainsi construit a le bon goût [20] d'être **complet** (théorème de Riesz-Fisher, 1907), c'est un Banach. Cela signifie que toute suite de Cauchy converge au sens de la norme $|| \cdot ||_p$. On vient de débloquer un nouveau mode de convergence ! Un corollaire immédiat, particulièrement utile en physique est le suivant.

51.17. **Corollaire.** *L'espace $L^2(\mu)$ muni du produit scalaire canonique est un espace de Hilbert.*

C'est en théorie à ce moment là que les formes linéaires continues font leur apparition. Tu as de très jolis résultats de représentations (Riesz-Fisher, Riesz) et de décomposition (Radon, Radon-Nikodym). Je passe (au moins momentanément), c'est trop technique pour moi pour l'instant. Il va falloir que je m'y penche sérieusement dessus.

Il ne manque qu'un seul petit ingrédient pour pouvoir intégrer sur des **espaces produits** : trouver une tribu naturelle à adjoindre à l'espace produit.

51.18. **Définition.** Soient $(X, \mathscr{A})$ et $(Y, \mathscr{B})$ deux espaces mesurables. On appelle **tribu produit** de $\mathscr{A}$ et $\mathscr{B}$ la tribu $\mathscr{A} \otimes \mathscr{B}$ égale à $\sigma(\{A \times B, A \in \mathscr{A}, B \in \mathscr{B}\})$.

D'un point de vue pédestre, la tribu produit n'est rien d'autre que la tribu engendrée par les pavés à côtés mesurables. Et là c'est banco ! La tribu est suffisamment bien faite pour qu'elle soit *géométriquement naturelle* [21], qu'elle jouisse de propriétés

---

19. Sur $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$ muni de la mesure de Lebesgue λ, on n'a aucune inclusion entre $\mathscr{L}^1$ et $\mathscr{L}^2$:

$$x \mapsto \frac{1_{]0,1]}}{\sqrt{x}} \in \mathscr{L}^1(\lambda) \setminus \mathscr{L}^2(\lambda) \quad \text{et} \quad x \mapsto \frac{1}{\sqrt{x^2+1}} \in \mathscr{L}^2(\lambda) \setminus \mathscr{L}^1(\lambda). \tag{229}$$

20. Il a également le bon goût de disposer de propriétés de densité bien joyeuses. Par exemple, dans $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$ muni de la mesure de Lebesgue λ, l'ensemble des fonctions en escalier à support compact est dense dans tous les espaces $\mathrm{L}^p(\lambda)$, $1 \leq p < +\infty$. C'est également le cas des fonctions continues à support compact. Remarque que les cas $+\infty$ ont été mis de côté. Il ne semble pas possible (?) de dire plus que les fonctions étagées sont denses dans $\mathrm{L}^\infty(\mu)$. De même, tu as des résultats de dualité qui foirent pour $p = +\infty$ (cf. les espaces réflexifs), de même pour des résultats de séparabilité.

21. *Géométriquement* en un sens volontairement vague, i.e. on cherche la plus petite tribu rendant les projections canoniques mesurables.

que l'on est bien en droit de lui demander. Par exemple une fonction est mesurable sur l'espace produit si et seulement si chacune de ses composante est mesurable. Néanmoins, on se gardera de ne pas manquer de quelques attentions. Pour X et Y deux espaces topologiques, on a toujours $\mathscr{B}(X) \otimes \mathscr{B}(Y) \subset \mathscr{B}(X \times Y)$. L'inclusion réciproque n'est pas toujours vraie, si X et Y sont à base dénombrable d'ouverts elle l'est par exemple. Par exemple, $\mathscr{B}(\mathbf{R}^2) = \mathscr{B}(\mathbf{R}) \otimes \mathscr{B}(\mathbf{R})$. Plus généralement, $\mathscr{B}(\mathbf{R}^d) = \mathscr{B}(\mathbf{R}) \otimes \cdots \otimes \mathscr{B}(\mathbf{R})$ d fois, noté $\mathscr{B}(\mathbf{R})^{\otimes d}$.

Le dernier problème à régler est celui de la mesure produit. Peut-on trouver une mesure sur l'espace produit $(X \times Y, \mathscr{A} \otimes \mathscr{B})$ héritant des propriétés inhérentes à $(X, \mathscr{A})$ et $(Y, \mathscr{B})$? Parfaitement, oui.

51.19. **Proposition.** *Soient $(X, \mathscr{A}, \mu)$ et $(X, \mathscr{B}, \nu)$ deux espaces mesurés σ-finis*[22]*. Il existe alors une* ***unique*** *mesure sur $(X \times Y, \mathscr{A} \otimes \mathscr{B})$ notée $\mu \otimes \nu$ vérifiant : $\forall A \in \mathscr{A}, B \in \mathscr{B}, (\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$.*

Application directe à la mesure de Lebesgue ! On munit $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$ de la mesure produit $\lambda_d := \lambda^{\otimes d}$ égale à d produit de la mesure de Lebesgue sur $\mathbf{R}$. Essentiellement, tu n'as *que* deux théorèmes à bien savoir manipuler. On fera bien attention à remarquer les hypothèses de **positivité** et d'**intégrabilité.**

51.20. **Théorème.** (**Fubini-Tonelli**) *Soient $f : (X \times Y, \mathscr{A} \otimes \mathscr{B}) \to \mathbf{R}_+$ une fonction mesurable, μ et ν deux mesures σ-finis sur $(X, \mathscr{A})$ et $(Y, \mathscr{B})$. Alors les fonctions suivantes sont mesurables :*

$$x \mapsto \int_Y f(x,y)\nu(\mathrm{d}y), \quad y \mapsto \int_X f(x,y)\mu(\mathrm{d}x). \tag{230}$$

De plus, on a le classique :

$$\int_{X \times Y} f(x,y)\mathrm{d}(\mu \otimes \nu) = \int_X \left(\int_Y f(x,y)\nu(\mathrm{d}y) \right) \mu(\mathrm{d}x) = \int_Y \left(\int_X f(x,y)\mu(\mathrm{d}x) \right) \nu(\mathrm{d}y) \tag{231}$$

51.21. **Théorème.** (**Fubini-Lebesgue**) *Considérons l'espace mesuré $(X \times Y, \mathscr{A} \otimes \mathscr{B}, \mu \otimes \nu)$. Soit $f \in \mathrm{L}^1(\mu \otimes \nu)$. Alors :*

$$x \mapsto f(x,y) \in \mathrm{L}^1(\mu) \ \nu(\mathrm{d}y)\text{-p.p.}, \quad y \mapsto f(x,y) \in \mathrm{L}^1(\nu) \ \mu(\mathrm{d}x)\text{-p.p.} \tag{232}$$

$$x \mapsto \int_Y f(x,y)\nu(\mathrm{d}y) \in \mathrm{L}^1(\mu) \ \mu\text{-p.p.}, \quad y \mapsto \int_X f(x,y)\mu(\mathrm{d}x) \in \mathrm{L}^1(\nu) \ \nu\text{-p.p.} \tag{233}$$

On a alors :

$$\int_{X \times Y} f(x,y)\mathrm{d}(\mu \otimes \nu) = \int_X \left(\int_Y f(x,y)\nu(\mathrm{d}y) \right) \mu(\mathrm{d}x) = \int_Y \left(\int_X f(x,y)\mu(\mathrm{d}x) \right) \nu(\mathrm{d}y) \tag{234}$$

52. Probabilités avec Laure Coutin

53. Inégalité de Hoeffding

54.

22. Hypothèse technique mais absolument nécessaire. Lorsque la mesure n'est pas de masse finie, i.e. $\mu(X)$ non fini, on peut tout de même espérer que la mesure soit une agglutination de composantes finies. On dira qu'une mesure sur un espace mesurable $(X, \mathscr{A})$ est σ-finie, s'il existe une suite croissante $(E_n)_n$ d'éléments de $\mathscr{A}$ tels que X soit l'union croissante des E_n et que leur mesure soit finie pour tout n. Par exemple, la mesure de Lebesgue est σ-finie. En effet, considérer $E_n =]-n, n[$. C'est bien de mesure finie $\lambda(E_n) = 2n < +\infty$ pour tout $n \in \mathbf{N}$.