Chapitre premier
Rappels d'intégration

Nous allons dans ce chapitre introductif rappeler les notions les plus importantes de la théorie de la mesure et de l'intégration qui nous serviront en permanence par la suite. Le lecteur est néanmoins invité à se reporter à des textes spécifiques pour de plus amples détails (voir par exemple [10]).
Dans ce qui suit, E désigne un ensemble non vide.

1.1 Définition. *On appelle tribu sur E* un ensemble $\mathcal{E}$ de parties de E tel que

$$\begin{array}{ll} \text{(i)} & \emptyset \in \mathcal{E}. \\ \text{(ii)} & A \in \mathcal{E} \Rightarrow A^c \in \mathcal{E}. \\ \text{(iii)} & (A_n)_{n\in\mathbb{N}} \in \mathcal{E} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{E}. \end{array}$$

Le couple $(E,\mathcal{E})$ est dit **ensemble mesurable**.
Quand la condition (iii) est remplacée par la condition beaucoup plus faible

$$\text{(iii}'\text{)}\ A,B \in \mathcal{E} \Rightarrow A\cup B \in \mathcal{E}$$

on dit que $\mathcal{E}$ est une **algèbre de Boole**.
On vérifie immédiatement qu'une intersection de tribus est une tribu. De ce fait, quand $\mathcal{B}$ est un ensemble de parties de E, l'intersection de toutes les tribus sur E qui contiennent $\mathcal{B}$ est la plus petite tribu sur E qui contient $\mathcal{B}$: on l'appelle la **tribu engendrée par** $\mathcal{B}$ et on la note $\sigma(\mathcal{B})$.
Un exemple fondamental est donné quand $E=\mathbb{R}$ et $\mathcal{B}$ est l'ensemble des intervalles ouverts de $\mathbb{R}$. La tribu engendrée par $\mathcal{B}$ est dite **tribu borélienne** de $\mathbb{R}$ et notée $\mathcal{B}(\mathbb{R})$. Cette définition se prolonge au cas de $\mathbb{R}^k$ et plus généralement d'un espace métrique E quand $\mathcal{B}$ est la famille des ouverts de E.

1.2 Définition. *Soient $(E,\mathcal{E})$ et $(F,\mathcal{F})$ deux espaces mesurables et $f : E \to F$ une application. On dit que f est* **mesurable** si $f^{-1}(\mathcal{F}) \subset \mathcal{E}$ c'est-à-dire

$$\forall B \in \mathcal{F}, \{x \in E, f(x) \in B\} \in \mathcal{E}.$$

Un cas particulier simple de fonctions mesurables est celui des fonctions continues de $\mathbb{R}^k$ dans $\mathbb{R}^m$ quand ces deux espaces sont munis de leurs tribus boréliennes. Parmi les fonctions mesurables de $(E,\mathcal{E})$ dans $(\mathbb{R},\mathcal{B}(\mathbb{R}))$, on a avant tout les **fonctions étagées** de la forme

$$f = \sum_{k=0}^{n} \alpha_k \mathbb{I}_{A_k}$$

où les A_k sont des ensembles de $\mathcal{E}$. Rappelons aussi qu'une limite simple d'une suite de fonctions mesurables est une fonction mesurable. A ce sujet, on a l'utile résultat suivant :

1.3 Proposition. *Toute fonction numérique mesurable positive est limite simple d'une suite croissante de fonctions étagées positives.*

Quand $(f_i)_{i\in I}$ est une famille de fonctions de E dans l'espace mesurable $(F,\mathcal{F})$, la tribu engendrée par les ensembles $f_i^{-1}(A), A\in\mathcal{F}$ est la plus petite tribu sur $\mathcal{E}$ qui rende mesurables toutes les f_i : c'est la tribu engendrée par les f_i, notée $\sigma(f_i, i\in I)$.

Exercice 1.1

Supposons $\mu(X)<\infty$. Soit (f_n) une suite de fonctions complexes mesurables qui converge simplement sur X. Alors quelque soit $\varepsilon>0$, il existe un sous ensemble mesurable E de X tel que $\mu(X\setminus E)<\varepsilon$ et que (f_n) converge uniformément sur E. (Théorème d'Egoroff)

Indication : poser

$$S(n,k)=\cap_{i,j>n}\left\{x : |f_i(x)-f_j(x)|<\frac{1}{k}\right\},$$

et montrer qu'il existe une suite croissante d'entiers (n_k) tel que $E=\cap_k S(n_k,k)$ ait la propriété désirée.

1.4 Définition. *Soit* $(E,\mathcal{E})$ *un espace mesurable. Une* **mesure** sur $(E,\mathcal{E})$ est une application $\mu:\mathcal{E}\to\mathbb{R}^+\cup\{+\infty\}$ telle que pour toute suite (A_n) d'éléments de E disjoints deux à deux on ait

$$\mu(\bigcup_{n\geq 0}A_n)=\sum_{n\geq 0}\mu(A_n).$$

L'exemple le plus important de mesure est donné par la mesure de Lebesgue sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$, notée traditionnellement λ, qui à un intervalle $]a,b[$ associe $b-a$. Rappelons au passage que son existence n'est pas absolument immédiate et résulte d'un théorème. Une mesure μ sur $(E,\mathcal{E})$ permet de définir l'**intégrale** d'une fonction mesurable $f:(E,\mathcal{E})\to(\mathbb{R},\mathcal{B}(\mathbb{R}))$. On considère d'abord des fonctions étagées positives sous la forme

$$f=\sum_{k=0}^{n}\alpha_k\mathbb{I}_{A_k}$$

pour lesquelles on a

$$\int_E f(x)d\mu(x)=\sum_{k=0}^{n}\alpha_k\mu(A_k).$$

On prolonge ensuite cette définition pour une fonction mesurable positive f : par la proposition 1.3, celle-ci est limite croissante de fonctions étagées f_k et on définit

$$\int_E f(x)d\mu(x)=\lim\uparrow\int_E f_k(x)d\mu(x).$$

Noter que ce nombre peut être infini. On montre qu'il ne dépend pas de la suite (f_k) choisie.

Enfin, quand f est une fonction mesurable, $f=f^+-f^-$ (où $a^+=\max(a,0)$ et $a^-=\max(-a,0)$) ; si $\int_E f^+(x)d\mu(x)<+\infty$ et $\int_E f^-(x)d\mu(x)<+\infty$, on dit que f est intégrable (sur E et par rapport à μ) et on pose

$$\int_E f(x)d\mu(x)=\int_E f^+(x)d\mu(x)-\int_E f^-(x)d\mu(x).$$

Noter que si f est intégrable, $|f|$ l'est aussi (puisque $|f| = f^+ + f^-$).
L'ensemble des fonctions intégrables sur $(E, \mathcal{E}, \mu)$ (ou plus exactement comme usuellement l'ensemble des classes d'équivalence de fonctions intégrables pour la relation d'égalité μ- presque partout) est noté $L^1(E, \mathcal{E}, \mu)$. Plus généralement, si p est un réel ≥ 1, on définit $L^p(E, \mathcal{E}, \mu)$ comme l'ensemble des fonctions mesurables de $(E, \mathcal{E})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telles que $|f|^p \in L^1(E, \mathcal{E}, \mu)$. On note alors

$$\|f\|_p = \Big(\int_E |f|^p (x) d\mu(x) \Big)^{1/p}.$$

On définit aussi $L^\infty(E, \mathcal{E}, \mu)$ comme l'ensemble des fonctions mesurables de $(E, \mathcal{E})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telles qu'il existe $M > 0$ tel que $\mu(\{x \in E, |f(x)| > M\}) = 0$ (on dit que $|f| \leq M$, presque partout (p.p.)) et $\|f\|_\infty$ est la borne inférieure de ces M.
L'inégalité suivante permet d'établir des liens entre les espaces L^p.

1.5 Théorème. (inégalité de Hölder) *Soient p et q deux réels tels que $\frac{1}{p} + \frac{1}{q} = 1$. Soient $f \in L^p(E, \mathcal{E}, \mu)$ et $g \in L^q(E, \mathcal{E}, \mu)$. Alors $fg \in L^1(E, \mathcal{E}, \mu)$ et*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Dans le cas d'une mesure de masse totale 1, on a aussi l'inégalité suivante.

1.6 Théorème. (inégalité de Jensen) *Soit $(E, \mathcal{E}, \mu)$ un espace mesuré tel que $\mu(E) = 1$. Soit $\phi : \mathbb{R} \to \mathbb{R}$ une fonction convexe. Pour toute f fonction intégrable telle que $\phi \circ f$ soit intégrable, on a*

$$\phi\Big(\int_E f(x)\mu(dx)\Big) \leq \int_E \phi \circ f(x)\mu(dx).$$

Exercice 1.2

Soit μ une mesure positive sur $(X, \mathcal{B})$, $f : X \to \mathbb{C}$ mesurable, et $p > 0$. On pose

$$\phi(p) = \int_X |f|^p d\mu = \|f\|_p^p.$$

On pose ensuite $E = \{p : \phi(p) < \infty\}$, et on suppose que $\|f\|_\infty > 0$.
a) Pour $r < p < s$, avec $r, s \in E$, montrer que $p \in E$.
b) Montrer que $\log \phi$ est convexe à l'intérieur de E et que ϕ est continue sur E.
c) D'après a), E est connexe. Est-il nécessairement ouvert, fermé ? Peut-il être réduit à un point ? Peut-il être n'importe quel sous ensemble connexe de $]0, +\infty[$?
d) Pour $r < p < s$, montrer que $\|f\|_p \leq \max\{\|f\|_r, \|f\|_s\}$. En conclure $L^p(\mu) \subset L^s(\mu) \cap L^r(\mu)$.
e) Supposons qu'il existe $r > 0$ tel que $\|f\|_r < \infty$. Montrer qu'alors $\|f\|_p \to \|f\|_\infty$.

◊ Exercice 1.3

On suppose toujours $\mu(X) = 1$. Soient f et g deux fonctions intégrables et positives telles que $fg \geq 1$. Montrer qu'alors

$$\int_X f d\mu \int_X g d\mu \geq 1.$$

Exercice 1.4
Toujours si $\mu(X) = 1$, soit $h > 0$ mesurable et $A = \int_X h d\mu$. Montrer que

$$\sqrt{1+A^2} \le \mathbb{E}_\mu(\sqrt{1+h^2}) \le 1+A.$$

Si μ est la mesure de Lebesgue sur $[0,1]$ et si h est continue, avec $h = f'$, les inégalités ci-dessus ont une interprétation géométrique simple. Laquelle ?

Exercice 1.5
Soit $1 < p < \infty$, $f \in L^p(]0,+\infty[)$, relativement à la mesure de Lebesgue. Posons pour $x > 0$,

$$F(x) = \frac{1}{x} \int_0^x f(t)dt.$$

a) Démontrer l'inégalité de Hardy

$$\|F\|_p \le \frac{p}{p-1}\|f\|_p,$$

qui assure que F envoie L^p dans lui-même.
b) Démontrer que l'on a l'égalité si et seulement si $f = 0$ p.s..
c) Démontrer que la constante $\frac{p}{p-1}$ est la plus petite possible.
d) Si $f > 0$ et $f \in L^1$, montrer que $F \notin L^1$.
Indications : pour le a), considérer d'abord $f \ge 0$ continue à support compact, et intégrer par parties pour obtenir $\|F\|_p^p = -p\int_0^\infty F^{p-1}(x)xF'(x)dx$. Remarquer que $xF' = f - F$, et appliquer l'inégalité de Hölder à $\int F^{p-1} f dx$.
Pour c), faire $f(x) = x^{-1/p}$ sur $[1,A]$, 0 ailleurs, pour A assez grand.

L'importance de la construction de l'intégrale vient de la très grande souplesse apportée aux passages à la limite (contrairement par exemple au cas de l'intégrale de Riemann pour laquelle on a en général besoin de convergences très fortes (type convergence uniforme) pour pouvoir passer à la limite).

1.7 Théorème. (convergence monotone, Fatou-Beppo-Levi) *Soit (f_n) une suite de fonctions mesurables telle que $0 \le f_n \le f_{n+1}$. On note pour tout $x \in E, \lim_n f_n(x) = f(x)$. Alors $\int_E f(x)d\mu(x) = \lim_n \int_E f_n(x)d\mu(x)$.*

Rappelons le Lemme de Fatou, qui peut permettre d'obtenir des inégalités intéressantes sous des hypothèses très faibles :

1.8 Théorème. (Lemme de Fatou) *Soit (f_n) une suite de fonctions mesurables telle que $\forall n, 0 \le f_n$. On note $f = \liminf_n f_n$. Alors $\int_E f(x)d\mu(x) \le \liminf_n \int_E f_n(x)d\mu(x)$.*

On en déduit le résultat le plus important, le théorème de convergence dominée de Lebesgue

1.9 Théorème. *Soit (f_n) une suite de fonctions mesurables telle que $\forall n, | f_n | \le g$ où g est une fonction intégrable sur $(E, \mathcal{E}, \mu)$. On suppose que p.p. $\lim_n f_n(x) = f(x)$. Alors $\int_E f(x)d\mu(x) = \lim_n \int_E f_n(x)d\mu(x)$.*

Rappelons aussi qu'on dit que la mesure μ admet la **densité** f **par rapport à la mesure** ν si f est une fonction numérique positive mesurable sur $\mathbb{R}$ telle que pour tout $A \in \mathcal{B}(\mathbb{R})$,

$$\mu(A) = \int_A f(x)\nu(dx).$$

Exercice 1.6

Calculer, en les justifiant, les limites suivantes :

$$\lim_{n\to\infty} \int_0^n \left(1 - \frac{x}{n}\right)^n \mathrm{e}^{x/2} dx, \qquad \lim_{n\to\infty} \int_0^n \left(1 + \frac{x}{n}\right)^n \mathrm{e}^{-2x} dx.$$

◊ Exercice 1.7

On reprend les hypothèses de l'exercice 2 mais en outre on suppose que $\mu(X) = 1$.
a) Démontrer que $\|f\|_r \leq \|f\|_s$ si $0 < r < s \leq \infty$.
b) A quelles conditions peut-on avoir $0 < r < s \leq \infty$ et $\|f\|_r = \|f\|_s < \infty$?
c) En supposant que $\|f\|_r < \infty$ pour un certain $r > 0$, montrer que

$$\lim_{p\to 0^+} \|f\|_p = \exp\left(\int_X \log|f| d\mu\right),$$

en posant $\exp(-\infty) = 0$.

Une autre notion fondamentale concerne les espaces-produits.

1.10 Définition. *Soient $(E,\mathcal{E})$ et $(F,\mathcal{F})$ deux espaces mesurables. La tribu sur $E \times F$ engendrée par les pavés $A \times B$ où $A \in \mathcal{E}$ et $B \in \mathcal{F}$ est dite* **tribu produit** . On la note $\mathcal{E} \otimes \mathcal{F}$.

Sur l'espace mesurable $(E \times F, \mathcal{E} \otimes \mathcal{F})$, on définit le produit d'une mesure sur $(E,\mathcal{E})$ par une mesure sur $(F,\mathcal{F})$.

1.11 Théorème et Définition. *Soit μ (resp. ν) une mesure sur l'espace mesurable $(E,\mathcal{E})$ (resp. $(F,\mathcal{F})$). Il existe une unique mesure sur $(E \times F, \mathcal{E} \otimes \mathcal{F})$, notée $\mu \otimes \nu$ et appelée* **produit de** μ **et** ν telle que

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B), A \in \mathcal{E}, B \in \mathcal{F}.$$

Les résultats fondamentaux suivants permettent d'intervertir l'ordre d'intégration dans le calcul d'une intégrale par rapport à une mesure produit. Le premier résultat sert notamment à démontrer l'intégrabilité d'une fonction de deux variables. Le second permet quant à lui de calculer une intégrale double par le biais du calcul de deux intégrales simples.

1.12 Théorème. (Fubini-Tonelli) *Soit μ (resp. ν) une mesure sur l'espace mesurable $(E,\mathcal{E})$ (resp. $(F,\mathcal{F})$). On suppose que μ et ν sont σ-finies (μ est σ-finie s'il existe une suite $(A_n) \in \mathcal{E}$ telle que $\bigcup_n \uparrow A_n = E$ et $\mu(A_n) < +\infty$). Soit $f : E \times F \to \mathbb{R}^+, \mathcal{E} \otimes \mathcal{F}$-mesurable, alors*

$$\int_{E\times F} f(x,y) d(\mu \otimes \nu)(x,y) = \int_E [\int_F f(x,y) d\nu(y)] d\mu(x) = \int_F [\int_E f(x,y) d\mu(x)] d\nu(y).$$

En outre, la fonction f est intégrable sur $E \times F$ si et seulement si les quantités ci-dessus sont finies.

1.13 Théorème. (Fubini) *Soit μ (resp. ν) une mesure sur l'espace mesurable $(E,\mathcal{E})$ (resp. $(F,\mathcal{F})$). On suppose que μ et ν sont σ-finies. Soit $f : E\times F \to \mathbb{R}$, $\mathcal{E}\otimes\mathcal{F}$-mesurable telle qu'il existe $G : F \to \mathbb{R}$, intégrable sur $(F,\mathcal{F},\nu)$ pour laquelle ν-p.p.,*

$$\int_E |\, f(x,y) \,|\, d\mu(x) \leq G(y).$$

Alors f est intégrable sur $E\times F$ et

$$\int_{E\times F} f(x,y)d(\mu\otimes\nu)(x,y) = \int_E [\int_F f(x,y)d\nu(y)]d\mu(x) = \int_F [\int_E f(x,y)d\mu(x)]d\nu(y).$$

Une autre notion très importante est la suivante :

1.14 Définition. *Soit μ une mesure finie (i.e. telle que $\mu(\mathbb{R}^d) < +\infty$) sur $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$. On définit la* **transformée de Fourier** de μ par

$$\hat{\mu}(t) = \int_E \mathrm{e}^{-2i\pi<t,x>}\mu(dx).$$

$\hat{\mu}$ ainsi définie est une application continue sur $\mathbb{R}^d$, bornée par $\mu(\mathbb{R}^d)$.
La propriété fondamentale de la transformation de Fourier est qu'il existe une **formule d'inversion** permettant de retrouver μ à partir de $\hat{\mu}$.

1.15 Théorème. *Soit μ une mesure finie sur $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$. Soit h continue et bornée de $\mathbb{R}^d$ dans $\mathbb{R}$. On a alors*

$$\int_{\mathbb{R}^d} h(x)\mu(dx) = \lim_{\sigma\downarrow 0}\int_{\mathbb{R}^d} h(x)\Big(\int_{\mathbb{R}^d}\hat{\mu}(u)\mathrm{e}^{2i\pi<u,x>-2\pi^2\sigma^2|u|^2}du\Big)dx.$$

Quand f est une fonction Lebesgue intégrable sur $\mathbb{R}^d$, les résultats précédents s'adaptent de façon agréable. La transformée de f, notée $\hat{f}$, est définie comme celle de la mesure $f(u)\lambda_{\mathbb{R}^d}(du)$, soit

$$\hat{f}(u) = \int_{\mathbb{R}^d} \mathrm{e}^{-2i\pi<u,x>} f(x)dx.$$

Quand $\hat{f}(u)$ est elle-même intégrable, la formule d'inversion s'écrit

$$f(u) = \int_{\mathbb{R}^d} \mathrm{e}^{2i\pi<u,x>}\hat{f}(u)du.$$

Chapitre 2
Espaces et mesures de probabilités

Depuis Kolmogorov (1933), le modèle habituel pour la présentation d'un problème de probabilités rentre dans le cadre de la théorie de la mesure.

2.1 Espaces de probabilités

2.1.1 Définition. *On appelle* **espace de probabilités** tout espace mesuré $(\Omega, \mathcal{F}, \mathbf{P})$, où Ω est un ensemble non vide, $\mathcal{F}$ une tribu sur Ω et $\mathbf{P}$ une mesure sur l'espace mesurable $(\Omega, \mathcal{F})$ telle que $\mathbf{P}(\Omega) = 1$.

Rappelons que les éléments de $\mathcal{F}$ s'appelent des **événements** , Ω est l'événement certain, et $\emptyset$ l'événement impossible. Par ailleurs, un événement de probabilité 1 est dit presque sûr. Plus généralement, on dira qu'une propriété est vraie **P**-presque sûrement (en abrégé **P**-p.s.) pour signifier qu'elle est vraie avec probabilité 1.
Rappelons une des multiples versions d'un résultat crucial sur les tribus qui sera souvent utilisé. Un ensemble de parties de Ω, $\mathcal{M}$, est dit une **classe monotone** si elle est stable par réunion dénombrable croissante et par intersection dénombrable décroissante.
Alors on a

2.1.2 Théorème. (de classe monotone). *Soient $\mathcal{F}_0$ une algèbre de Boole et $\mathcal{M}$ une classe monotone. Si $\mathcal{F}_0 \subset \mathcal{M}$, alors $\sigma(\mathcal{F}_0) \subset \mathcal{M}$.*

PREUVE : Tout d'abord, de manière manifeste, on constate qu'une intersection de classes monotones est une classe monotone. On peut donc définir une notion de classe monotone engendrée par un ensemble de parties $\mathcal{B}$, c'est-à-dire une plus petite classe monotone contenant $\mathcal{B}$. Par ailleurs, il est manifeste qu'une algèbre de Boole qui est une classe monotone est une tribu. De ce fait, pour montrer le résultat, il suffit de montrer que la classe monotone engendrée par $\mathcal{F}_0$, notée $\mathcal{C}$, est une algèbre de Boole.
Soit $\mathcal{G} = \{A, A^c \in \mathcal{C}\}$. Puisque $\mathcal{C}$ est une classe monotone, $\mathcal{G}$ l'est aussi. Par ailleurs, $\mathcal{F}_0 \subset \mathcal{G}$ et donc $\mathcal{C} \subset \mathcal{G}$. Donc $\mathcal{C}$ est stable par complémentaire.
Soit maintenant $A \in \mathcal{F}_0$. Considérons $\mathcal{G}_1 = \{B, B \cup A \in \mathcal{C}\}$. Clairement, $\mathcal{F}_0 \subset \mathcal{G}_1$ et de plus $\mathcal{G}_1$ est une classe monotone. De ce fait, $\mathcal{C} \subset \mathcal{G}_1$ et l'union d'un élément quelconque de $\mathcal{C}$ avec un élément de $\mathcal{F}_0$ est donc dans $\mathcal{C}$.
Posons $\mathcal{G}_2 = \{B, B \cup C \in \mathcal{C}, \forall C \in \mathcal{C}\}$. D'après l'étape précédente, $\mathcal{F}_0 \subset \mathcal{G}_2$. Par ailleurs, $\mathcal{G}_2$ est une classe monotone contenant $\mathcal{F}_0$ donc $\mathcal{C}$ qui est stable par union finie. $\square$

L'application essentielle du résultat précédent est l'identification d'une mesure μ. Plus précisément, si deux mesures μ et ν sur le même espace mesurable $(\Omega, \mathcal{F})$ coïncident sur une algèbre de Boole $\mathcal{F}_0$, elles coïncident sur $\sigma(\mathcal{F}_0)$. En particulier, si $\sigma(\mathcal{F}_0) = \mathcal{F}$, les deux mesures sont égales. Notons au passage, à toute fin utile, qu'il n'est absolument pas

suffisant que les deux mesures coïncident sur un système générateur de la tribu pour avoir le résultat (penser aux mesures δ_0 et $2\delta_0$ sur l'ensemble des intervalles ouverts $]a,b[$ de $\mathbb{R}$ qui ne contiennent pas 0, où δ_0 désigne la mesure de Dirac en 0).
Ci-dessous, un autre résultat du même genre est proposé.

Exercice 2.1

Soit Ω un ensemble. On appelle π-système un ensemble de parties stable par intersection finie. On appelle λ-système un ensemble de parties qui contient l'ensemble vide, est stable par complémentation et par réunion dénombrable disjointe (cette dernière condition est donc moins exigente que pour une tribu).
a) Montrer que toute intersection de λ-systèmes est un λ-système.
b) Montrer qu'un π-système qui est aussi un λ-système est une tribu.
c) Soient $\mathcal{P}$ un π-système et $\mathcal{L}$ un λ-système tels que $\mathcal{P} \subset \mathcal{L}$. Montrer que $\sigma(\mathcal{P}) \subset \mathcal{L}$ (ce résultat est connu sous le nom de "$\pi\lambda$ **théorème** ").

Enfin, mentionnons sans démonstration la forme fonctionnelle des résultats précédents. Elle ne sera en fait utilisée qu'une seule fois dans ce cours, en 12.4.

2.1.3 Théorème. *Soit $\mathcal{H}$ un espace vectoriel de fonctions numériques bornées définies sur Ω, contenant les constantes, et tel que pour toute suite f_n croissante et bornée,* $\lim_n \uparrow f_n \in \mathcal{H}$.
Alors, si $\mathcal{H}$ contient le π-système $\mathcal{P}$ (au sens où si $A \in \mathcal{P}$, $\mathbb{I}_A \in \mathcal{H}$), il contient aussi les processus bornés $\sigma(\mathcal{P})$-mesurables.

2.2 Théorème d'extension

Le théorème d'extension nous permettra de construire effectivement des lois de probabilités. Une **probabilité sur une algèbre de Boole** $\mathcal{A}$ se définit à l'instar de ce qu'il en est d'une tribu, à ceci près qu'on va rajouter, pour la σ-additivité, que $\cup_n A_n \in \mathcal{A}$ (ce qui n'est pas automatique). Autrement dit, $\mathbf{P}$ probabilité sur l'algèbre de Boole $\mathcal{A}$ satisfait $\mathbf{P}(\Omega) = 1$, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$ pour A et B disjoints dans $\mathcal{A}$ et , pour $(A_n)_{n\geq 0}$ une suite d'éléments de $\mathcal{A}$ **telle que** $\bigcup_{n\geq 0} A_n \in \mathcal{A}$,

$$\mathbf{P}(A) \leq \sum_{n\geq 0} \mathbf{P}(A_n).$$

2.2.1 Théorème. *Soit* $\mathbf{P}$ *une probabilité sur une algèbre de Boole $\mathcal{A}$. Alors il existe une unique probabilité* $\tilde{\mathbf{P}}$ *sur $\sigma(\mathcal{A})$, qui coïncide avec* $\mathbf{P}$ *sur $\mathcal{A}$.*

PREUVE : La preuve repose sur l'utilisation de la **mesure extérieure** $\mathbf{P}^*$, et est pour l'essentiel dûe à Carathéodory. On pose

$$\mathbf{P}^*(E) = \inf_{\substack{A_n \in \mathcal{A}, \\ E \subset \cup_n A_n}} \sum_n \mathbf{P}(A_n), \ E \subset \Omega.$$

Considérons

$$\mathcal{M} = \{A \subset \Omega : \forall E \subset \Omega, \ \mathbf{P}^*(A \cap E) + \mathbf{P}^*(A^c \cap E) = \mathbf{P}^*(E)\}.$$