cours 17, le jeudi 25 mars 2010

V.2.2. *Fubini général*

Théorème de Fubini. *On suppose que μ, ν sont deux mesures σ-finies. Pour toute fonction f à valeurs réelles ou complexes qui est $\mu \otimes \nu$-intégrable sur $\mathrm{X} \times \mathrm{Y}$, la fonction $y \to f(x, y)$ est ν-intégrable pour presque tout x ; la fonction presque partout définie $x \to \int_{\mathrm{Y}} f(x, y)\, \mathrm{d}\nu(y)$ est μ-intégrable et*

$$\int_{\mathrm{X}\times\mathrm{Y}} f\, \mathrm{d}(\mu \otimes \nu) = \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} f(x, y)\, \mathrm{d}\nu(y) \Bigr)\, \mathrm{d}\mu(x).$$

On a aussi, avec les mêmes précautions de langage,

$$\int_{\mathrm{X}\times\mathrm{Y}} f\, \mathrm{d}(\mu \otimes \nu) = \int_{\mathrm{Y}} \Bigl(\int_{\mathrm{X}} f(x, y)\, \mathrm{d}\mu(x) \Bigr)\, \mathrm{d}\nu(y).$$

Preuve. — La fonction $|f(x, y)|$ est mesurable positive, d'intégrale finie par l'hypothèse du théorème ; par Fubini positif on a

$$\int_{\mathrm{X}\times\mathrm{Y}} |f|\, \mathrm{d}(\mu \otimes \nu) = \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} |f(x, y)|\, \mathrm{d}\nu(y) \Bigr)\, \mathrm{d}\mu(x) < +\infty.$$

Puisque l'intégrale en $\mathrm{d}\mu(x)$ précédente est finie, l'ensemble

$$\mathrm{N} = \{x \in \mathrm{X} : \int_{\mathrm{Y}} |f(x, y)|\, \mathrm{d}\nu(y) = +\infty\} \in \mathcal{F}$$

est de mesure nulle pour μ. Pour tout $x \notin \mathrm{N}$, la fonction $y \to f(x, y)$ est ν-intégrable. Montrons le cas réel avec f^+ et f^- (la discussion sera analogue dans le cas complexe avec $\mathrm{Re}\, f$ et $\mathrm{Im}\, f$). Comme N est μ-négligeable

$$\int_{\mathrm{X}\times\mathrm{Y}} f^+\, \mathrm{d}\xi = \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} f^+(x, y)\, \mathrm{d}\nu(y) \Bigr)\, \mathrm{d}\mu(x) = \int_{\mathrm{N}} \Bigl(\int_{\mathrm{Y}} f^+(x, y)\, \mathrm{d}\nu(y) \Bigr)\, \mathrm{d}\mu(x)$$

et pareil pour f^-. Si x n'est pas dans N,

$$\int_{\mathrm{Y}} f^+(x, y)\, \mathrm{d}\nu(y) < +\infty, \quad \int_{\mathrm{Y}} f^-(x, y)\, \mathrm{d}\nu(y) < +\infty\,;$$

en faisant la différence on obtient une fonction $x \to \int_{\mathrm{Y}} f(x, y)\, \mathrm{d}\nu(y)$ définie μ-presque partout sur X, bien définie sur $\mathrm{X} \setminus \mathrm{N}$; d'après les calculs précédents, et par les conventions sur les intégrales des fonctions définies presque partout, on a

$$\begin{aligned}\int_{\mathrm{X}\times\mathrm{Y}} f\, \mathrm{d}\xi &= \int_{\mathrm{X}\times\mathrm{Y}} f^+\, \mathrm{d}\xi - \int_{\mathrm{X}\times\mathrm{Y}} f^-\, \mathrm{d}\xi = \int_{\mathrm{N}} \Bigl(\int_{\mathrm{Y}} (f^+(x, y) - f^-(x, y))\, \mathrm{d}\nu(y) \Bigr)\, \mathrm{d}\mu(x) \\ &= \int_{\mathrm{N}} \Bigl(\int_{\mathrm{Y}} f(x, y)\, \mathrm{d}\nu(y) \Bigr)\, \mathrm{d}\mu(x) = \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} f(x, y)\, \mathrm{d}\nu(y) \Bigr)\, \mathrm{d}\mu(x).\end{aligned}$$

Produits de plus de deux facteurs
Considérons par exemple le produit de trois facteurs X_1, X_2, X_3, qui soient munis de mesures σ-finies μ_1, μ_2 et μ_3 ; si on pose $Y = X_2 \times X_3$, muni de la mesure $\mu_2 \otimes \mu_3$, on peut ensuite considérer que $X_1 \times X_2 \times X_3$ est naturellement isomorphe à $X_1 \times Y$, et introduire la mesure $\xi = \mu_1 \otimes (\mu_2 \otimes \mu_3)$, considérée comme mesure sur $X_1 \times X_2 \times X_3$; cette mesure ξ vérifie, pour tous $A_1 \in \mathcal{F}_1$, $A_2 \in \mathcal{F}_2$ et $A_3 \in \mathcal{F}_3$,

$$(*) \qquad \xi(A_1 \times A_2 \times A_3) = \mu_1(A_1)\mu_2(A_2)\mu_3(A_3)$$

Avec les conditions de σ-finitude, on voit que les produits $A_1 \times A_2 \times A_3$ de mesure finie forment une classe stable par intersection qui contient une suite croissante d'éléments dont la réunion est l'espace entier. Grâce au théorème d'unicité, on voit qu'il existe une seule mesure ξ qui vérifie $(*)$; elle est donc égale aussi à $(\mu_1 \otimes \mu_2) \otimes \mu_3$, où encore au dernier cas plus difficile à décrire où on groupe d'abord X_1 et X_3. On notera simplement cette mesure par $\xi = \mu_1 \otimes \mu_2 \otimes \mu_3$.

On s'intéressera particulièrement à la mesure de Lebesgue λ_n sur $\mathbb{R}^n$, qu'on peut considérer comme le produit tensoriel de n exemplaires de la mesure de Lebesgue sur $\mathbb{R}$.

V.3. Changement de variables

V.3.1. *Changement de variable linéaire (affine)*
On a vu le cas de dimension un,

$$\int_{\mathbb{R}} f(y)\,\mathrm{d}y = \int_{\mathbb{R}} f(ax+b)\,|a|\,\mathrm{d}x.$$

On passe à la dimension finie quelconque.

Théorème. *Soient* A *une application linéaire bijective de* $\mathbb{R}^n$ *sur* $\mathbb{R}^n$ *et* $b \in \mathbb{R}^n$ *un vecteur fixé ; on considère le changement de variable affine* $y = Ax + b$*, où* $x, y \in \mathbb{R}^n$*. Pour toute fonction* $f : \mathbb{R}^n \to [0, +\infty]$ *borélienne, on a*

$$\int_{\mathbb{R}^n} f(y)\,\mathrm{d}\lambda_n(y) = \int_{\mathbb{R}^n} f(Ax+b)\,|\det A|\,\mathrm{d}\lambda_n(x).$$

On notera que le changement de mesure $\mathrm{d}y = |a|\,\mathrm{d}x$ provenant du changement de variable $y = ax + b$ en dimension un est remplacé par $\mathrm{d}\lambda_n(y) = |\det A|\,\mathrm{d}\lambda_n(x)$ quand $y = Ax + b$ en dimension n.

Preuve. — Elle se fait par récurrence sur la dimension n. On connaît le cas $n = 1$. On va raconter l'histoire avec beaucoup de mots, sans doute beaucoup trop...

Montrons le passage de n à $n+1$ dans le cas particulier $n = 2$; on va découper le changement de variable affine $y = Ax + b$ en deux changements qui seront un peu plus simples ; le premier sera écrit $z = A_1 x + b'$, suivi d'un changement $y = Lz + b''$. Le changement global sera $y = Ax + b = L(A_1 x + b') + b'' = LA_1 x + Lb' + b''$, ce qui montre que la partie linéaire A du changement global sera factorisée sous la forme $A = LA_1$, donc $\det A = (\det L)(\det A_1)$. Si on démontre la formule de changement de variable pour chacun de deux changements séparément,

$$\int_{\mathbb{R}^3} f(y)\,\mathrm{d}\lambda_3(y) = \int_{\mathbb{R}^3} f(Lz + b'')\,|\det L|\,\mathrm{d}\lambda_3(z)$$

et pour toute fonction borélienne positive g sur $\mathbb{R}^3$,

$$\int_{\mathbb{R}^3} g(z)\,\mathrm{d}\lambda_3(z) = \int_{\mathbb{R}^3} g(\mathrm{A}_1 x + b')\,|\det \mathrm{A}_1|\,\mathrm{d}\lambda_3(x),$$

on aura en posant $g(z) = f(\mathrm{L}z + b'')$

$$\int_{\mathbb{R}^3} f(y)\,\mathrm{d}\lambda_3(y) = \int_{\mathbb{R}^3} f(\mathrm{L}z + b'')\,|\det \mathrm{L}|\,\mathrm{d}\lambda_3(z) = |\det \mathrm{L}| \int_{\mathbb{R}^3} g(z)\,\mathrm{d}\lambda_3(z)$$

$$= \int_{\mathbb{R}^3} f(\mathrm{L}(\mathrm{A}_1 x + b') + b'')\,|\det \mathrm{L}|\,|\det \mathrm{A}_1|\,\mathrm{d}\lambda_3(x) = \int_{\mathbb{R}^3} f(\mathrm{A}x + b)\,|\det \mathrm{A}|\,\mathrm{d}\lambda_3(x),$$

la formule attendue.

Expliquons la factorisation. Comme la matrice A est inversible on obtient, en développant le déterminant suivant la dernière colonne, que

$$0 \neq \det \mathrm{A} = \sum_{i=1}^{3} (-1)^{i+3} a_{i,3} \det \mathrm{A}^{(i,3)}$$

où $\mathrm{A}^{(i,3)}$ est la matrice 2×2 obtenue en enlevant de A la ligne i et la colonne 3. Il existe donc i tel que $\det \mathrm{A}^{(i,3)} \neq 0$, et *quitte à permuter les lignes de* A, on supposera pour le confort de l'écriture que

$$\mathrm{A}' := \mathrm{A}^{(3,3)} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

est inversible. On définit le changement $z = \mathrm{A}_1 x + b'$ par

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix},$$

de sorte que $z_1 = y_1$, $z_2 = y_2$ et $z_3 = x_3$. On note que

$$\begin{pmatrix} z_1 - a_{1,3} z_3 - b_1 \\ z_2 - a_{2,3} z_3 - b_2 \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathrm{A}' \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

et comme A' est inversible on peut exprimer x_1, x_2 comme fonctions affines de z,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{pmatrix} \begin{pmatrix} z_1 - a_{1,3} z_3 - b_1 \\ z_2 - a_{2,3} z_3 - b_2 \end{pmatrix},$$

où la matrice C est l'inverse de A' ; on définit le deuxième changement $y = \mathrm{L}z + b''$ par les formules $y_1 = z_1$, $y_2 = z_2$ et

$$y_3 = a_{3,1} x_1 + a_{3,2} x_2 + a_{3,3} x_3 + b_3 = u z_1 + v z_2 + w z_3 + b''_3,$$

pour des coefficients u, v, w qu'il n'est pas utile d'expliciter. Les parties linéaires des deux changements successifs donnent la factorisation

$$\mathrm{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ u & v & w \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ 0 & 0 & 1 \end{pmatrix}.$$

Le calcul des déterminants donne $\det \mathrm{A} = w \det \mathrm{A}'$. Pour finir la preuve il suffit de vérifier la formule de changement de variable affine séparément pour les deux changements successifs. On choisit pour le premier changement l'ordre d'intégration suivant,

$$\int_{\mathbb{R}^3} g(z)\,\mathrm{d}\lambda_3(z) = \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}^2} g(z_1, z_2, z_3)\,\mathrm{d}z_1\,\mathrm{d}z_2 \Bigr)\,\mathrm{d}z_3.$$

Pour t réel fixé considérons le changement de variable affine dans $\mathbb{R}^2$ donné par

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_{1,3}t \\ a_{2,3}t \end{pmatrix} = \mathrm{A}' \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_{1,3}t \\ a_{2,3}t \end{pmatrix},$$

l'hypothèse de récurrence permet d'appliquer la formule de changement de variable affine en deux variables à la fonction borélienne positive $(z_1, z_2) \to g(z_1, z_2, t)$, ce qui donne

$$\int_{\mathbb{R}^2} g(z_1, z_2, t)\,\mathrm{d}z_1\,\mathrm{d}z_2 =$$

$$= \int_{\mathbb{R}^2} g(a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}t, a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}t, t)\,|\det \mathrm{A}'|\,\mathrm{d}x_1\,\mathrm{d}x_2.$$

On termine le calcul de l'intégrale en intégrant en t, mais comme le nom de la variable muette d'intégration n'a pas d'importance, on peut écrire le résultat sous la forme

$$\int_{\mathbb{R}^3} g(z)\,\mathrm{d}\lambda_3(z) = \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}^2} g(z_1, z_2, z_3)\,\mathrm{d}z_1\,\mathrm{d}z_2 \Bigr)\,\mathrm{d}z_3 =$$

$$= \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}^2} g(a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3, a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3, x_3)\,|\det \mathrm{A}'|\,\mathrm{d}x_1\,\mathrm{d}x_2 \Bigr)\,\mathrm{d}x_3,$$

ce qui règle le premier changement. Le deuxième est analogue, mais il faut cette fois faire porter l'intégrale intérieure sur z_3 ; sur $\mathbb{R}$, pour t_1 et t_2 fixés on effectue le changement affine $y_3 = wz_3 + (ut_1 + vt_2 + b''_3)$, obtenant ainsi

$$\int_{\mathbb{R}} f(t_1, t_2, ut_1 + vt_2 + wz_3 + b''_3)\,|w|\,\mathrm{d}z_3 = \int_{\mathbb{R}} f(t_1, t_2, y_3)\,\mathrm{d}y_3.$$

On complète en intégrant en (t_1, t_2) sur $\mathbb{R}^2$, ce qui donne à nouveau en changeant le nom des variables d'intégration

$$\int_{\mathbb{R}^2} \Bigl(\int_{\mathbb{R}} f(z_1, z_2, uz_1 + vz_2 + wz_3 + b''_3)\,|w|\,\mathrm{d}z_3 \Bigr)\,\mathrm{d}z_1\,\mathrm{d}z_2 =$$

$$= \int_{\mathbb{R}^2} \Bigl(\int_{\mathbb{R}} f(y_1, y_2, y_3)\,\mathrm{d}y_3 \Bigr)\,\mathrm{d}y_1\,\mathrm{d}y_2.$$

Remarque. La méthode de preuve est reliée au résultat général suivant : si A est une matrice $n \times n$ inversible, on peut écrire $\mathrm{A} = \mathrm{PLU}$, P matrice permutation, L matrice triangulaire inférieure et U matrice triangulaire supérieure. On aurait pu aller beaucoup plus vite en admettant ce résultat matriciel : il suffisait alors de déduire de Fubini que la formule de changement de variable est vraie dans le cas triangulaire.

V.3.1. *Changement de variable général*

Si Φ est une application de classe C^1 définie sur un ouvert U de $\mathbb{R}^d$, à valeurs dans $\mathbb{R}^d$, elle est donnée par ses d fonctions coordonnées,

$$\Phi(x_1, \ldots, x_d) = \begin{pmatrix} \varphi_1(x_1, \ldots, x_d) \\ \vdots \\ \varphi_d(x_1, \ldots, x_d) \end{pmatrix},$$

fonctions réelles φ_i de d variables définies sur U, où i varie de 1 à d ; la matrice jacobienne $(\mathrm{J}\Phi)(x)$ de Φ en un point $x \in \mathrm{U}$ est la matrice dont la ligne i contient les coordonnées de la différentielle de φ_i au point x, c'est-à-dire, les valeurs des dérivées partielles

$$\Bigl(\frac{\partial \varphi_i}{\partial x_j}(x) \Bigr).$$

Théorème. *On suppose que Φ est une bijection d'un ouvert* U *de $\mathbb{R}^n$ sur un ouvert* V *de $\mathbb{R}^n$, de classe C^1 ainsi que son inverse ; si on pose $y = \Phi(x)$, on a pour toute fonction borélienne $f : \mathrm{V} \to [0, +\infty]$*

$$\int_{\mathrm{V}} f(y)\, \mathrm{d}\lambda_n(y) = \int_{\mathrm{U}} f(\Phi(x))\, |\det \mathrm{J}\Phi(x)|\, \mathrm{d}\lambda_n(x).$$

Le théorème sera admis pour l'essentiel. Mentionnons seulement qu'il existe deux stratégies de preuve.

— Découper l'ouvert U en très petits morceaux, sur lesquels le comportement du changement de variable Φ est « presque le même » que celui de sa partie linéaire, donnée par la matrice jacobienne $(\mathrm{J}\Phi)(x_0)$ en un point x_0 du « petit morceau ». On utilise alors sur chaque petit morceau la formule linéaire qu'on a montrée, appliquée au changement affine approché

$$y = \Phi(x_0) + (\mathrm{J}\Phi)(x_0)(x - x_0)\,;$$

on doit contrôler la qualité de l'approximation, c'est la partie la plus délicate.

— Factoriser le changement de variable en deux changements plus simples. Imitons la preuve linéaire, qui faisait intervenir une matrice fixée A. On va poser $y = \Phi(x)$, qui s'écrit en composantes sous la forme $y_i = \varphi_i(x)$, $i = 1, 2, 3$. La matrice $\mathrm{A}(x)$ est maintenant variable, c'est la jacobienne de Φ au point $x \in \mathrm{U}$. Il existe une sous-matrice 2×2 inversible sélectionnée dans les deux premières colonnes, mais son choix dépend du point x. On va esquisser le cas simple où, comme on l'avait supposé dans la preuve linéaire, la sous-matrice $\mathrm{A}'(x) = \mathrm{A}^{(3,3)}(x)$ reste inversible pour tout $x \in \mathrm{U}$.

On commence par le premier changement de variable, qui comme dans le cas linéaire, consiste à changer z_1, z_2 mais à garder $z_3 = x_3$: on pose $z_1 = y_1 = \varphi_1(x_1, x_2, x_3)$, $z_2 = \varphi_2(x_1, x_2, x_3)$ et $z_3 = x_3$. La matrice jacobienne de z en fonction de x est égale à

$$\begin{pmatrix} \mathrm{A}'(x) & a_{1,3}(x) \\ 0 & 1 \end{pmatrix}.$$

Comme on a supposé $\mathrm{A}'(x)$ inversible, on va pouvoir, grâce au théorème des fonctions implicites (ou l'inversion locale), imiter l'inversion linéaire qui utilisait la matrice C, pour exprimer, pour chaque $z_3 = x_3$ fixé, le couple (x_1, x_2) en fonction de (z_1, z_2, z_3),

$$x_1 = c_1(z_1, z_2, z_3), \quad x_2 = c_2(z_1, z_2, z_3)\,;$$

on peut alors terminer la factorisation du changement de variable $y = \Phi(x)$ en posant

$$y_1 = z_1, \quad y_2 = z_2, \quad y_3 = \varphi_3(c_1(z_1, z_2, z_3), c_2(z_1, z_2, z_3), z_3).$$

On obtient le produit de deux changements partiellement triangulaires. L'hypothèse de récurrence (la dimension deux est connue) et Fubini permettent de traiter le changement de x en z ; la dimension un et Fubini traitent le changement de z en y.

Coordonnées polaires

À $r \geq 0$ et θ réel on associe le point $(r\cos\theta, r\sin\theta)$. Cette correspondance n'est pas bijective sur $[0, +\infty[\times\mathbb{R}$. On doit préciser un ouvert U sur lequel elle sera bijective. Il n'y a qu'un choix possible pour r, c'est de prendre $r > 0$, mais pour θ on doit choisir l'intervalle de longueur 2π où seront prises les valeurs des arguments.

On choisit pour U l'ensemble ouvert dans $\mathbb{R}^2$ formé des (r, θ), $0 < r < +\infty$ et (par exemple) $-\pi < \theta < \pi$, et on pose

$$\Phi(r, \theta) = (r\cos\theta, r\sin\theta).$$

L'application Φ est une bijection de U sur l'ensemble ouvert V donné par

$$\mathrm{V} = \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}.$$

On a

$$(\mathrm{J}\Phi)(r, \theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

dont le déterminant est r.

Le complémentaire de V est une demi-droite, dont la mesure est nulle pour λ_2 ; on peut donc écrire pour toute fonction borélienne positive f sur $\mathbb{R}^2$

$$\int_{\mathbb{R}^2} f(x, y)\,\mathrm{d}x\mathrm{d}y = \int_{\mathrm{V}} f(x, y)\,\mathrm{d}x\mathrm{d}y = \int_{-\pi}^{\pi} \Bigl(\int_0^{+\infty} f(r\cos\theta, r\sin\theta) r\,\mathrm{d}r\Bigr)\,\mathrm{d}\theta.$$

Exemple : intégrale gaussienne. En polaires,

$$\int_{\mathbb{R}^2} \mathrm{e}^{-x^2/2-y^2/2}\,\mathrm{d}x\mathrm{d}y = \int_{-\pi}^{\pi} \Bigl(\int_0^{+\infty} \mathrm{e}^{-r^2/2}\, r\,\mathrm{d}r\Bigr)\,\mathrm{d}\theta = 2\pi \int_0^{+\infty} \mathrm{e}^{-r^2/2}\, r\,\mathrm{d}r$$

$$= 2\pi \int_0^{+\infty} \mathrm{e}^{-u}\,\mathrm{d}u = 2\pi,$$

et avec Fubini positif

$$\int_{\mathbb{R}^2} \mathrm{e}^{-x^2/2-y^2/2}\,\mathrm{d}x\mathrm{d}y = \Bigl(\int_{\mathbb{R}} \mathrm{e}^{-x^2/2}\,\mathrm{d}x\Bigr)^2,$$

ce qui donne le moyen le plus classique de calculer l'intégrale gaussienne,

$$\int_{\mathbb{R}} \mathrm{e}^{-x^2/2}\,\mathrm{d}x = \sqrt{2\pi}.$$