Le théorème de dualité des espaces L^p

Version provisoire du 2 avril 2006.

Dans un article de 1910 qui a profondément influencé l'Analyse mathématique du vingtième siècle, << Untersuchungen über Systeme integrierbarer Funktionen >>, Math. Annalen 69, 449-497, F. Riesz prouve (parmi d'autres choses) un résultat que l'on peut reformuler ainsi :

Théorème : *Soit X un espace mesuré*(*)*, soit* $1 < p < \infty$*, d'exposant conjugué* q *(*$\frac{1}{p} + \frac{1}{q} = 1$*), alors le dual de* $L^p(X,\mu)$ *est canoniquement* $L^q(X,\mu)$ *via l'appariement* $(u,v) \mapsto \int_X u(x)v(x)d\mu(x)$.

Riesz ne considère (je crois) que le cas $X = I$, $I \subset \mathbb{R}$ un intervalle, muni de la mesure de Lebesgue $d\mu = dx$ (j'avoue, piteusement, ne jamais encore avoir consulté les Oeuvres de Riesz, à l'exception de son livre d'Analyse Fonctionnelle avec Sz-Nagy et c'est donc aussi le moment tragique de vous révéler que je ne suis nullement un expert des choses que je suis censé(**) vous enseigner ; pire je n'ai réellement réfléchi au contenu de cette feuille que ces derniers jours).

Par l'inégalité de Hölder (qui, pour les intégrales, fut en fait énoncée par Riesz dans ce même article de 1910, je crois) on sait déjà qu'effectivement chaque $v \in L^q(X,\mu)$ définit une forme linéaire continue L sur $L^p(X,\mu)$ via $L(u) = \int_X u(x)v(x)d\mu(x)$, et $\|L\| \leq \|v\|_q$. Supposons $v \not\equiv 0$ dans L^q et prenons $u = v^{-1}|v|^q$ (pour les x avec $v(x) = 0$ on pose $u(x) = 0$). Notons (aussi pour ces x) que $|u|^p = |v|^{-p+pq} = |v|^q$ donc $u \in L^p$ et $(\|u\|_p)^p = (\|v\|_q)^q$. De plus pour ce u on constate que $L(u) = \int_X |v(x)|^q\, d\mu(x)$. Il en résulte $\|L\| \geq \frac{|L(u)|}{\|u\|_p} = \|v\|_q^q\, \|v\|_q^{-q/p} = \|v\|_q$. Donc en fait $\|L\| = \|v\|_q$.(***)

On a ainsi une injection isométrique $L^q \hookrightarrow (L^p)^*$ et il s'agit de montrer qu'elle est surjective, c'est-à-dire, pour toute forme linéaire continue L sur L^p trouver v dans L^q telle que $\forall u \in L^p \quad L(u) = \int_X uv d\mu$. Ce n'est pas (du tout)

(*) donc, muni d'une tribu et d'une mesure. Dans la notation $L^p(X,\mu)$ je ne fais figurer que la mesure, la tribu reste implicite.

(**) le suis-je vraiment ? ça n'a pas l'air d'être considéré si important que cela en fait ces jours-ci.

(***) à propos, l'inégalité de Minkowski $\|v_1+v_2\|_q \leq \|v_1\|_q+\|v_2\|_q$ peut se concevoir comme une conséquence de cela, pourquoi ?.

évident. Nous allons chercher autant que possible à imiter la preuve faite dans le cas $p = q = 2$ (théorème de Fréchet-Riesz[(****)]). En particulier un rôle important est joué par la notion de << projection >> sur un sous-espace vectoriel fermé. Cependant nous verrons que la situation est, si l'on attaque le problème frontalement, plus délicate pour $1 < p < 2$ que pour $p > 2$.

Nous pouvons supposer L non identiquement nulle, et nous notons W son noyau qui est un sous-espace vectoriel fermé de L^p. Nous prenons u_0 avec $L(u_0) = 1$, et, nous admettons (la discussion relative à ce point vient ensuite) qu'il existe $w_0 \in W$ qui a la propriété $\|u_0 - w_0\|_p = \inf_{w\in W} \|u_0 - w\|_p$.

Soit maintenant $u_1 = u_0 - w_0$, alors à nouveau $L(u_1) = 1$. Remplacer u_0 par u_1[(*)] c'est comme remplacer w_0 par 0, et on voit qu'effectivement on a :

$$\forall w \in W \quad \|u_1\|_p \leq \|u_1 - w\|_p$$

Considérons, pour $w \in W$ fixé, et $(t, x) \in \mathbb{R} \times X$:

$$f(t, x) = |u_1(x) - tw(x)|^p \qquad f(t) = \int_X f(t, x)d\mu(x) = (\|u_1 - t\,w\|_p)^p$$

Pour $x \in X$ fixé, et en t_0 tel que $u_1(x) - t_0 w(x) \neq 0$ la fonction $t \mapsto f(t, x)$ est dérivable : pour ϵ réel, $|\epsilon| \ll 1$:

$$\begin{aligned}|u_1(x) - t_0 w(x) - \epsilon w(x)|^p &= |u_1(x) - t_0 w(x)|^p (1 - \epsilon\, 2\, \Re \frac{w(x)}{u_1(x) - t_0 w(x)} + \mathcal{O}(\epsilon^2))^{p/2} \\ &= |u_1(x) - t_0 w(x)|^p (1 - \epsilon\, p\, \Re \frac{w(x)}{u_1(x) - t_0 w(x)} + \mathcal{O}(\epsilon^2))\end{aligned}$$

$$\begin{aligned}f'(t_0, x) &= -p|u_1(x) - t_0 w(x)|^p\, \Re \frac{w(x)}{u_1(x) - t_0 w(x)} \\ &= -p\,|u_1(x) - t_0 w(x)|^{p-1}\, \Re \left(\frac{|u_1(x) - t_0 w(x)|}{u_1(x) - t_0 w(x)} w(x) \right) \\ |f'(t_0, x)| &\leq p|u_1(x) - t_0 w(x)|^{p-1}\, |w(x)|\end{aligned}$$

Si $u_1(x) - t_0 w(x) = 0$ alors $f(t_0 + \epsilon, x) = |\epsilon|^p\, |w(x)|^p$ est dérivable en $\epsilon = 0$ de dérivée nulle ($p > 1$; aussi on a imposé à u_1 et w d'être partout finies). On peut considérer que la formule ci-dessus est encore valable. Donc $f(t, x)$ est dérivable en tout t, pour chaque x. De plus on peut, pour $-T < t < T$, majorer $|f'(t, x)|$ par $p\,(|u_1(x)| + T|w(x)|)^{p-1}|w(x)|$ qui est une fonction intégrable (pourquoi ?) indépendante de t. Par le théorème de dérivation sous le signe somme, on peut donc affirmer que la fonction $f(t)$ est une fonction dérivable, et en prenant en particulier $t_0 = 0$, on a :

$$f'(0) = -p \int_X |u_1(x)|^{p-1} \Re \left(\frac{|u_1(x)|}{u_1(x)}\, w(x) \right) d\mu(x)$$

(****) le théorème de Fréchet-Riesz détermine le dual d'un Hilbert V quelconque. Spécialisé à $V = L^2(X, \mu)$ il dit que toute forme linéaire s'écrit $u \mapsto \int_X u\overline{v}d\mu$. On remplacera v par $\overline{v}$ qui est encore dans L^2 pour voir cela comme un cas particulier pour $p = 2$ du Théorème de Dualité des L^p.

(*) désolé, l'annexe précédente a je crois les notations inversées.

Je rappelle qu'aux x avec $u_1(x) = 0$ on a convenu qu'il fallait comprendre que l'intégrand était nul. Maintenant par construction la fonction $t \mapsto f(t)$ a un minimum en $t = 0$, donc $f'(0) = 0$. Ainsi :

$$\forall w \in W \quad 0 = \Re\left(\int_X u_1(x)^{-1}|u_1(x)|^p\, w(x)\, d\mu(x)\right)$$

Nous pouvons toujours trouver un nombre complexe ρ de module 1 de sorte que $\rho\int_X u_1(x)^{-1}|u_1(x)|^p\ w(x)\,d\mu(x) \in \mathbb{R}$. Maintenant nous appliquons ce qui précède à ρw au lieu de w, et finalement nous concluons :

$$\forall w \in W \qquad 0 = \int_X u_1(x)^{-1}|u_1(x)|^p\, w(x)\, d\mu(x)$$

Soit $v_1 = u_1^{-1}|u_1|^p$. Cette fonction appartient à L^q et $\|v_1\|_q^q = \|u_1\|_p^p$. Soit u quelconque dans L^p, et soit $w = u - L(u)u_1$. De $L(w) = L(u) - L(u) = 0$ on a $w \in W$ et donc par ce qui précède :

$$\begin{aligned} 0 = \int_X v_1(x)w(x)\,d\mu(x) &= \int_X v_1(x)u(x)\,d\mu(x) - L(u)\int_X v_1(x)u_1(x)\,d\mu(x) \\ &= \int_X v_1(x)u(x)\,d\mu(x) - L(u)\|u_1\|_p^p \end{aligned}$$

Finalement on pose $v(x) = \|v_1\|_q^{-q}\,v_1$ et on a ce que l'on voulait :

$$\forall u \in L^p \qquad L(u) = \int_X v(x)u(x)\,d\mu(x)$$

Pour que la preuve du Théorème de Dualité de Riesz soit complète, il suffit donc de prouver le théorème suivant :(*)

Théorème de Projection : *soit* $W \subset L^p$ *un sous-espace vectoriel fermé, et soit* $u_0 \in L^p$. *Notons* $d(u_0, W) = \inf_{w\in W}\|u_0 - w\|_p$. *Alors il existe* $w_0 \in W$ *tel que* $\|u_0 - w_0\|_p = d(u_0, W)$.

Si W est de dimension finie alors il suffit de dire que la fonction continue $w \mapsto \|u_0 - w\|_p$, sur le compact $\|w\|_p \leq 2\|u_0\|_p$ atteint son minimum (si $\|w\|_p > 2\|u_0\|_p$ alors $2\|u_0\|_p < \|w\|_p \leq \|w - u_0\|_p + \|u_0\|_p$ donc $\|u_0 - w\|_p > \|u_0\|_p = \|u_0 - 0\|_p$ et ainsi l'infimum sur la boule $\{\|w\|_p \leq 2\|u_0\|_p\}$ est égal à l'infimum sur W tout entier). Mais, si W est de dimension infinie on ne peut pas procéder ainsi car F. Riesz, toujours lui, a prouvé : *dans un espace vectoriel normé* $(W, \|\cdot\|)$ *de dimension infinie la boule unité fermée n'est pas compacte.*

Comme c'est important, en voici une preuve. Si $Y \subset W$ est un sous-espace vectoriel de dimension finie, soit $x \notin Y$. Soit $d = d(x, Y) > 0$. Soit $y \in Y$ avec $\|x - y\| \leq 2d$. Soit $z = x - y$, alors $d(z, Y) = d(x, Y) = d \geq \frac{1}{2}\|z\|$ et z est non nul. Enfin soit $z' = \frac{z}{\|z\|}$. Alors $\|z'\| = 1$ et $d(z', Y) = \|z\|^{-1}d(z, Y) \geq \frac{1}{2}$. Maintenant je prouve que la boule unité n'est pas compacte. On prend z_1 de norme

(*) On peut remplacer << sous-espace vectoriel fermé >> par << ensemble convexe fermé >>, nous n'en avons pas besoin ici.

1 quelconque. Supposons connus $z_1, \ldots, z_N$. Soit Y_N l'espace vectoriel qu'ils engendrent. Il existe par ce qui précède z' de norme 1 à distance $\geq \frac{1}{2}$ de Y_N. On pose $z_{N+1} = z'$. La suite (z_n) est construite de sorte que pour tout n on a $\|z_n\| = 1$ et pour $n \neq m$ on a $\|z_n - z_m\| \geq \frac{1}{2}$. Aucune suite extraite ne peut être convergente. La boule unité n'est pas compacte.

Pour le théorème de projection il faut donc une idée autre. Dans le cas $p = 2$, la preuve classiquement proposée (il y en a de légèrement différentes), a été résumée avec indications au début de l'Annexe précédente. Elle utilise l'identité du parallélogramme (pour alléger la notation j'écris $\|f - g\|$ au lieu de $\|f - g\|_2$ etc...) :

$$\|f - g\|^2 + \|f + g\|^2 = 2\|f\|^2 + 2\|g\|^2$$

Soit $(w_j)_{j\geq 1}$ une suite dans $W \subset L^2$ avec $\|u_0 - w_j\| \dashrightarrow d(u_0, W)$. Avec $f = w_j - u_0$, $g = w_k - u_0$, on obtient :

$$\|w_j - w_k\|^2 = 4\left(\frac{\|w_j - u\|^2 + \|w_k - u\|^2}{2} - \|\frac{w_j + w_k}{2} - u_0\|^2\right)$$

Compte tenu de $\|\frac{w_j+w_k}{2} - u_0\| \geq d(u_0, W)$ il est alors facile de prouver que (w_j) est de Cauchy dans ce cas L^2. Comme L^2 est complet cette suite converge et sa limite w_0 appartient à W puisque l'on a supposé ce dernier fermé.

Pour $p \neq 2$ cela est (semble-t-il) plus compliqué pour $1 < p < 2$ que pour $p > 2$. J'explique donc d'abord comment on peut procéder pour $p > 2$. L'identité du parallélogramme n'est pas valable au niveau des normes mais elle est valable ponctuellement :

$$|f(x) - g(x)|^2 + |f(x) + g(x)|^2 = 2|f(x)|^2 + 2|g(x)|^2$$

Je vais élever cela à la puissance $\frac{p}{2}$ (comme $p > 2$, la fonction $x \mapsto x^{p/2}$ est convexe et c'est cela qui va nous aider). Une petite étude préliminaire de $(a + b)^r$ versus $a^r + b^r$ pour $r = \frac{p}{2} > 1$ fixé , et $a, b > 0$ n'est pas superflue. Soit $\phi(x) = \frac{(1+x)^r}{1+x^r}$ pour $0 \leq x \leq 1$. La dérivée logarithmique $(0 < x < 1)$ est $r\frac{1}{1+x} - r\frac{x^{r-1}}{1+x^r} = r\frac{1-x^{r-1}}{(1+x)(1+x^r)} > 0$ donc $1 \leq \phi(x) \leq 2^{r-1}$ pour $0 \leq x \leq 1$ et(*)

$$a^r + b^r \leq (a + b)^r \leq 2^{r-1}(a^r + b^r) \qquad (r > 1; 0 \leq a, b \leq +\infty)$$

On écrit alors :

$$\begin{aligned}|f(x) - g(x)|^p + |f(x) + g(x)|^p &= (|f(x) - g(x)|^2)^{p/2} + (|f(x) + g(x)|^2)^{p/2} \\ &\leq \left(|f(x) - g(x)|^2 + |f(x) + g(x)|^2\right)^{p/2} \\ &= 2^{p/2}\left(|f(x)|^2 + |g(x)|^2\right)^{p/2} \\ &\leq 2^{p-1}\left(|f(x)|^p + |g(x)|^p\right)\end{aligned}$$

(*) vous pouvez retrouver l'inégalité de gauche comme un cas particulier de Minkowski et celle de droite comme un cas particulier de Hölder.

$$|f(x) - g(x)|^p \leq 2^p \left(\frac{|f(x)|^p + |g(x)|^p}{2} - \left| \frac{f(x) + g(x)}{2} \right|^p \right)$$

$$(\|f - g\|_p)^p \leq 2^p \left(\frac{\|f\|_p^p + \|g\|_p^p}{2} - \|\frac{f + g}{2}\|_p^p \right)$$

Et on y prend $f = w_j - u_0$, $g = w_k - u_0$, $d(u_0, W) = \inf_{w \in W} \|u_0 - w\|_p$, $\|u_0 - w_j\| \dashrightarrow d(u_0, W)$, on fait la remarque que $\|\frac{w_j + w_k}{2} - u_0\|_p \geq d(u_0, W)$, et alors :

$$\|w_j - w_k\|_p \leq 2 \left(\frac{\|u_0 - w_j\|_p^p + \|u_0 - w_k\|_p^p}{2} - d(u_0, W)^p \right)^{1/p}$$

et donc (w_j) est de Cauchy. Voilà qui complète la preuve du théorème de dualité pour $p > 2$.

Le cas 1 < p < 2 est plus difficile, par cette approche qui consiste à vouloir faire dans L^p comme dans L^2, avec le point-clé de la projection. Que cela soit encore possible pour 1 < p < 2 a été rattaché par Clarkson (<< Uniformly convex spaces >>, Trans. AMS 40 (1936), 396-414) a une notion de << convexité uniforme >> pour L^p (p > 1) ; cela veut dire qu'il y a des inégalités qui remplacent l'identité du parallélogramme et qui permettent de faire la preuve selon notre schéma général. Grâce à un travail de Hanner (<< On the uniform convexity of L^p and l^p >> ; Arkiv för Matematik, Band 3, nr 19, 1956, 239-244) on connaît l'inégalité optimale (en un certain sens). Cette inégalité sophistiquée, la voici pour $1 \leq p \leq 2$ (pour $p \geq 2$ l'inégalité est dans l'autre sens et n'a pas d'intérêt spécial dans le contexte de la recherche du $w_0 \in W$ minimisant $\|u_0 - w_0\|$) :

$$|\,\|f + g\|_p + \|f - g\|_p\,|^p + |\,\|f + g\|_p - \|f - g\|_p\,|^p \leq 2^p(\|f\|_p^p + \|g\|_p^p) \qquad (1 \leq p \leq 2)$$

Je vous laisse le challenge de l'utiliser pour compléter notre programme : si $(w_j)_{j \geq 1}$ est une suite dans W avec $\|u_0 - w_j\|_p \dashrightarrow d(u_0, W)$ alors elle est de Cauchy, donc convergente car L^p est complet. Sa limite w_0 = lim w_j est dans W car ce dernier est fermé, et w_0 fait l'affaire car de w_0 = lim w_j on déduit $\|u_0 - w_0\|_p = \lim \|u_0 - w_j\|_p = d(u_0, W)$ par continuité de $u \mapsto \|u\|_p$. Pour la justification de l'inégalité de Hanner et la notion d'uniforme convexité de Clarkson je vous renvoie à des livres tels que << Analyse fonctionnelle élémentaire >> de M. Willem (éd. Cassini) et << Analysis >> de E. Lieb et M. Loss (éd. AMS).

L'article de Riesz est pour $X = \mathbb{R}$ ou un intervalle ; sa méthode ne distingue pas $1 < p \leq 2$ de $p \geq 2$. Elle semble à première vue assez spécifique à $X = \mathbb{R}$ (voir le Chapitre Deux du livre de Riesz et Sz-Nagy pour une preuve probablement assez proche de celle de Riesz en 1910), mais l'ingéniosité d'homo sapiens est grande, et avec le Théorème de Radon-Nikodym, que je n'énoncerai pas ici, et la notion attenante de mesure absolument continue par rapport à une autre, on a le formalisme qui permet de suivre Riesz pour un espace mesuré X quelconque ; pour être précis disons pour les espaces mesurés σ-finis (c'est-à-dire tels que $X = \cup_{n \geq 1} X_n$ avec $\mu(X_n) < \infty$ pour chaque n). Cela donne une preuve que vous trouverez au Chapitre Six du livre de Rudin << Analyse réelle et complexe >>.

La preuve de Radon-Nikodym chez Rudin suit un procédé de von Neumann qui introduit un espace de Hilbert-Lebesgue ad-hoc (c'est surprenant puisque rien ne faisait référence a priori dans Radon-Nikodym à des fonctions de carrés intégrables). Finalement on est toujours dans une histoire de projection mais elle est faite dans un Hilbert !

L'invocation de Radon-Nikodym fournit une fonction v dont on montre après qu'elle est dans L^q. Cette étape, nous allons la court-circuiter, car nous n'avons plus qu'à faire pour $p < 2$ et comme on sait que cela marche pour $p = 2$ on va avoir un v par ce biais. Je suis en cela Lax au Chapitre Huit de son livre << Functional Analysis >> (éd. Wiley). On suppose $\mu(X) = 1$. Je laisse en exercice l'extension du résultat au cas plus général d'un X qui est σ-fini. Comme $1 < p < 2$ et $\mu(X) = 1$ on a l'inclusion $L^2 \subset L^p$ et les inégalités $\|f\|_p \leq \|f\|_2$. Donc toute forme linéaire continue L sur L^p définit par restriction une forme linéaire continue sur L^2.(*) Par le théorème de Fréchet-Riesz il existe v dans L^2 tel que $L(u) = \int_X u(x)v(x)d\mu(x)$ pour tout u dans L^2. On va montrer que v est dans L^q. Pour cela définissons pour $N = 1, 2 \ldots$: $u_N(x) = \min(|v(x)|, N)^{q-1} \frac{\overline{v(x)}}{|v(x)|}$ si $v(x) \neq 0$ et $u_N(x) = 0$ si $v(x) = 0$. Alors $|u_N(x)| \leq N^{q-1}$ est borné donc dans L^2 et :

$$L(u_N) = \int_X \min(|v(x)|, N)^{q-1} |v(x)| d\mu(x) \geq \int_X \min(|v(x)|, N)^q d\mu(x)$$

Par ailleurs $|L(u_N)| \leq \|L\| \cdot \|u_N\|_p$. Or :

$$(\|u_N\|_p)^p = \int_X \min(|v(x)|, N)^{pq-p} d\mu(x) = \int_X \min(|v(x)|, N)^q d\mu(x)$$

Donc

$$\int_X \min(|v(x)|, N)^q d\mu(x) \leq |L(u_N)| \leq \|L\| \left(\int_X \min(|v(x)|, N)^q d\mu(x)\right)^{\frac{1}{p}}$$

Donc

$$\left(\int_X \min(|v(x)|, N)^q d\mu(x)\right)^{\frac{1}{q}} \leq \|L\|$$

On passe à la limite par le théorème de la convergence monotone et on obtient $\|v\|_q \leq \|L\|$. Donc v est bien un élément de L^q. Soit M la forme linéaire sur L^p définie par v. On a $L = M$ sur le sous-espace $L^2 \subset L^p$ et il s'agit d'un sous-espace dense car les fonctions simples (combinaisons linéaires finies de fonctions indicatrices de parties mesurables) sont contenues dans L^2 et sont denses dans L^p (si vous avez lu et tout compris dans toutes mes annexes vous êtes suffisamment calé(e) pour justifier cela sans peine).

Donc la forme linéaire continue L sur L^p est bien de la forme

$$u \mapsto L(u) = \int_X u(x)v(x) \, d\mu(x) \ ,$$

pour un certain $v \in L^q$. Cela complète la preuve du Théorème de Dualité.

(*) dans ce qui suit la notation $\|L\|$ est la norme de L comme forme linéaire sur L^p, pas L^2.