A muggle's approach to the uniform convexity of L^p, and related questions

May 24, 2023

Captatio benevolentiæ Textbook proofs of Riesz' theorem "$(L^p)' = L^{p'}$, $1 < p < \infty$", rely: either (a) on the uniform convexity of the L^p spaces with $1 < p < \infty$, combined with properties of reflexive spaces (see, e.g., Brezis [2, Chapter IV]), or (b) on uniform convexity, combined with James' theorem (see, e.g., Lieb and Loss [6, Chapter 2] or Willem [7, Chapitre V]), or (c) on the Radon-Nikodym theorem, and in this case one has the unnecessary extra assumption that the underlying measure is σ-finite (see, e.g., Bogachev [1, Chapter 4]). In turn, uniform convexity is usually established via the well-known inequalities of Clarkson [4] or Hanner [5] (for the latter ones, Hanner gives credit to Beurling). The more difficult case is $1 < p < 2$, for which other inequalities are available (see, e.g., Morawetz' approach, [2, Exercise 4.12]).

Although the above inequalities have now elegant and relatively concise proofs, they are definitely non-intuitive when $1 < p < 2$, and still have the appearance of a magic trick. In addition, they require separate analysis for $1 < p < 2$ and $p > 2$. My main motivation is to present a cheap, muggle's, p independent proof of the following result, equivalent to uniform convexity.

Proposition 1. Let $1 < p < \infty$. For each $\varepsilon > 0$, there exists some $\delta = \delta(\varepsilon) > 0$ such that

$$[f, g \in L^p, \|f\|_p = 1, \|f+g\|_p^p + \|f-g\|_p^p \le 2+\delta] \implies \|g\|_p \le \varepsilon. \tag{1}$$

More delicate (and irrelevant for obtaining uniform convexity) is the question of the value of δ_{opt}, the optimal δ in (1), and the characterization of the couples (f, g) satisfying the equality case

$$\|f\|_p = 1, \|g\|_p = \varepsilon, \|f+g\|_p^p + \|f-g\|_p^p = 2+\delta_{\text{opt}}. \tag{2}$$

This echoes [4] and [5], where similar questions were raised for related inequalities. Note that, in principle, δ_{opt} depends not only on p and ε, but also on the underlying measured space $(X, \mathscr{T}, \mu)$. Let us also note that, when $p = 2$ and $L^2 \neq \{0\}$, the parallelogram identity yields $\delta_{\text{opt}} = 2\varepsilon^2$, and for this δ, equality on the the left- and the right-hand side of (1) are equivalent. When $p > 2$, this optimality issue was settled by Hanner, who proved the following result [5, Theorem 1].

Proposition 2. Let $2 < p < \infty$. Then

$$[f, g \in L^p, \|f\|_p = 1, \|g\|_p = \varepsilon] \implies \|f+g\|_p^p + \|f-g\|_p^p \ge 2+2\varepsilon^p, \tag{3}$$

with equality if and only if $fg = 0$ a.e.

In particular, if $(X, \mathscr{T}, \mu)$ contains two disjoint measurable sets A, B such that $0 < \mu(A), \mu(B) < \infty$, then $\delta_{\text{opt}} = 2\varepsilon^p$.

When $1<p<2$, we prove the following counterpart of [5, Theorem 2].

Proposition 3. Let $1<p<2$. Then

$$[f,g\in L^p, \|f\|_p=1, \|g\|_p=\varepsilon] \implies \|f+g\|_p^p+\|f-g\|_p^p \geq (1+\varepsilon)^p+|1-\varepsilon|^p, \tag{4}$$

with equality if and only if $|g|=\varepsilon|f|$ a.e.

In particular, if $(X,\mathscr{T},\mu)$ contains a measurable set A such that $0<\mu(A)<\infty$, then $\delta_{\text{opt}}=(1+\varepsilon)^p+|1-\varepsilon|^p-2$.

Note that, in Propositions 2 and 3, there is no smallness assumption on ε.

Proofs We start by giving the heuristics of the proof of Proposition 1. When $|g|$ is not much smaller than $|f|$, we prove that $|f+g|^p+|f-g|^p-2|f|^p$ dominates $|g|^p$, and then we are done. On the other hand, when $|g|\ll|f|$ we have $\|g\|_p\ll\|f\|_p=1$, and then we are done again. We conclude by combining the two above arguments. (This dichotomy type argument is similar, e.g., to the strategy of the proof of the Brezis-Lieb lemma [3].)

Proof of Proposition 1. Consider the function

$$(0,\infty)\ni t\mapsto F(t):=\frac{(1+t)^p+|1-t|^p-2}{t^p}.$$

Since $F>0$ (by strict convexity of $x\mapsto |x|^p$) and $\lim_{t\to\infty}F(t)=2$, for each $\lambda>0$ we have

$$0<C_\lambda:=\inf\{F(t);\, \lambda\leq t<\infty\}\leq 2. \tag{5}$$

By (5) and homogeneity when $f(x)\neq 0$, and inspection of (6) when $f(x)=0$, we have

$$|(f+g)(x)|^p+|(f-g)(x)|^p-2|f(x)|^p\geq C_\lambda|g(x)|^p \quad \text{if } |g(x)|\geq\lambda|f(x)|. \tag{6}$$

On the other hand,

$$|g(x)|^p<\lambda^p|f(x)|^p \quad \text{if } |g(x)|<\lambda|f(x)|. \tag{7}$$

Combining (6) and (7) and using the fact that (again by convexity) the left-hand side of (6) is non-negative on the whole underlying space X, we find that

$$\|g\|_p^p\leq\frac{1}{C_\lambda}(\|f+g\|_p^p+\|f-g\|_p^p-2\|f\|_p^p)+\lambda^p\|f\|_p^p,\ \forall f,\forall g,\forall\lambda>0. \tag{8}$$

We obtain (1) by letting, e.g., $\lambda^p=\varepsilon^p/2$ and $\delta=C_\lambda\varepsilon^p/2$. □

Proof of Proposition 3. Step 1. Proof when $f(x)\neq 0$, $\forall x\in X$. By considering the measure $|f|^p\,\mu$ instead of μ and the function g/f instead of f, we may assume that μ is a probability measure, that $f=1$, and then we have to prove that

$$\|1+g\|_p^p+\|1-g\|_p^p\geq\left(1+\|g\|_p\right)^p+\left|1-\|g\|_p\right|^p, \tag{9}$$

with equality if and only if $|g|$ is constant a.e.

With no loss of generality, we may assume that $g\geq 0$. Let $h:=g^p\geq 0$ and set

$$\Psi(t):=\left(1+t^{1/p}\right)^p+\left|1-t^{1/p}\right|^p,\ \forall t\geq 0.$$

Then (9) amounts to

$$\int \Psi(h) \geq \Psi\left(\int h\right), \forall h \in L^1, h \geq 0, \tag{10}$$

with equality if and only if h is constant a.e. In turn, (10) holds provided Ψ is strictly convex.

Set

$$X = X(t) := 1 + t^{1/p},\ Y = Y(t) := |1 - t^{1/p}|.$$

When $0 < t < 1$, we have $Y(t) = 1 - t^{1/p}$ and

$$\begin{aligned}
\Psi'(t) &= pX^{p-1}X' + pY^{p-1}Y' = t^{1/p-1}(X^{p-1} - Y^{p-1}),\\
\Psi''(t) &= -\frac{p-1}{p}t^{1/p-2}(X^{p-1} - Y^{p-1}) + \frac{p-1}{p}t^{2/p-2}(X^{p-2} + Y^{p-2})\\
&= \frac{p-1}{p}t^{1/p-2}\left[(t^{1/p} - X)X^{p-2} + (t^{1/p} + Y)Y^{p-2}\right]\\
&= \frac{p-1}{p}t^{1/p-2}\left[Y^{p-2} - X^{p-2}\right] > 0,
\end{aligned}$$

since $0 < Y < X$ and $p < 2$.

Similarly, when $t > 1$, we have $Y(t) = t^{1/p} - 1$ and

$$\Psi''(t) = \frac{p-1}{p}t^{1/p-2}Y\left[X^{p-2} + Y^{p-2}\right] > 0.$$

This completes Step 1.

Step 2. Proof in the general case. Let $A := \{x; f(x) \neq 0\}$. Set $B := X \setminus A$ and $s := \|g\|_{L^p(A)}$. By Step 1, we have

$$\int_A [|f+g|^p + |f-g|^p] \geq (1+s)^p + |1-s|^p, \tag{11}$$

with equality if and only if $|g| = s$ a.e. on A. On the other hand, we have

$$\int_B [|f+g|^p + |f-g|^p] = 2\int_B |g|^p = 2(\varepsilon^p - s^p). \tag{12}$$

In view of (11) and (12), in order to complete the proof it suffices to prove that

$$(1+s)^p + |1-s|^p + 2(\varepsilon^p - s^p) > (1+\varepsilon)^p + |1-\varepsilon|^p,\ \forall \varepsilon > 0, \forall 0 \leq s < \varepsilon. \tag{13}$$

In turn, (13) amounts to proving that the function

$$[0,\infty) \ni s \mapsto \Phi(s) := (1+s)^p + |1-s|^p - 2s^p$$

is (strictly) decreasing. Set $\alpha := p-1 \in (0,1)$. When $0 < s < 1$, we have

$$\Phi'(s) = p\left[(1+s)^\alpha - (1-s)^\alpha - 2s^\alpha)\right]. \tag{14}$$

Using the inequality

$$(x+y)^\alpha < x^\alpha + y^\alpha,\ \forall x, y > 0,$$

we find that

$$(1+s)^\alpha = (1-s+s+s)^\alpha < (1-s)^\alpha + 2s^\alpha, \tag{15}$$

and thus, by (14) and (15), $\Phi' < 0$ on $(0,1)$.

When $s > 1$, the inequality $\Phi'(s) < 0$ amounts to

$$(1+s)^\alpha + (s-1)^\alpha < 2s^\alpha,$$

which follows from the strict concavity of $x \mapsto x^\alpha$, $x > 0$.

The proof of Proposition 3 is complete. □

For the sake of completeness, we also present the

Proof of Proposition 2. It suffices to prove that, when $a, b \in \mathbb{R}^*$, we have

$$|a+b|^p + |a-b|^p > 2|a|^p + 2|b|^p. \tag{16}$$

By homogeneity, (16) amounts to

$$\Xi(t) := (1+t)^p + |1-t|^p - 2t^p > 2 = \Xi(0), \ \forall\, t > 0. \tag{17}$$

In order to obtain (16), we prove that Ξ is (strictly) increasing.

Set $\beta := p - 1 > 1$. When $0 < t < 1$, we have

$$\Xi'(t) = p\left[(1+t)^\beta - (1-t)^\beta - 2t^\beta\right], \tag{18}$$

and (as in the proof of (15)) the inequality $\Xi'(t) > 0$ is a consequence of (18) and

$$(x+y)^\beta > x^\beta + y^\beta, \ \forall\, x, y > 0.$$

When $t > 1$, we have

$$\Xi'(t) = p\left[(1+t)^\beta + (t-1)^\beta - 2t^\beta\right], \tag{19}$$

and the inequality $\Xi'(t) > 0$ follows from (19) and the strict convexity of $x \mapsto x^\beta$, $x > 0$. □

References

[1] V. I. Bogachev, Measure theory. Vol. I, Springer-Verlag, Berlin, 2007, xviii+500.

[2] H. Brezis, Functional analysis, Sobolev spaces and partial differential equations, Universitext, Springer, New York, 2011, xiv+599.

[3] H. Brezis and E. H. Lieb, *A relation between pointwise convergence of functions and convergence of functionals*, Proc. Amer. Math. Soc., **88** (1983), no 3, 486–490.

[4] J. A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc., **40** (1936), no 3, 396–414.

[5] O. Hanner, *On the uniform convexity of L^p and l^p*, Ark. Mat., **3** (1956), 239–244.

[6] E. H. Lieb and M. Loss, Analysis, Graduate Studies in Mathematics, vol. 14, 2nd edition, American Mathematical Society, Providence, RI, 2001, xxii+346.

[7] M. Willem, Principes d'analyse fonctionnelle, Nouvelle Bibliothèque Mathématique vol. 9, Cassini, Paris, 2007, iv+198.