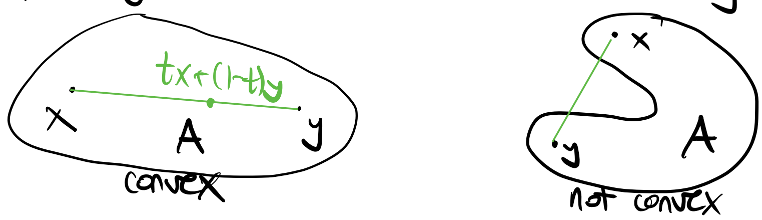
We now turn to some geometric aspects of infinite-dimensional spaces and their subsets, and relate this to their functional analytic properties.

Separation theorems

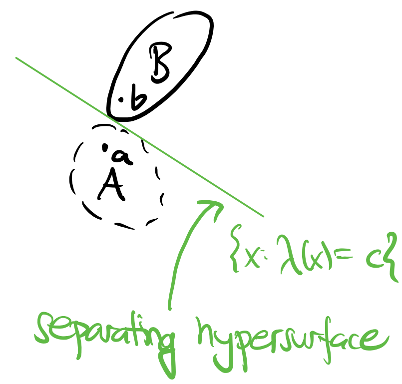
Definition A subset $A \subset X$ of a normed vector space X is *convex* if $x, y \in A$, $t \in [0,1] \Rightarrow tx + (1-t)y \in A$.



Theorem Let X be a normed $\mathbb{R}$-vector space. Let $A, B \subset X$ be non-empty convex sets with $A \cap B = \emptyset$.

(i) If A is open, then $\exists\, \lambda \in X^*$, $c \in \mathbb{R}$ s.t.

$$\lambda(a) < c \le \lambda(b) \quad \forall\, a \in A,\ b \in B.$$



(ii) If A is compact and B is closed, then $\exists\, \lambda \in X^*$, $c \in \mathbb{R}$ s.t.

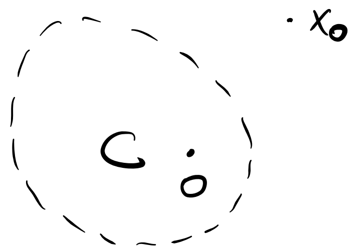
$$\sup_{a \in A} \lambda(a) < c < \inf_{b \in B} \lambda(b).$$

Proof (i) Fix $a_0 \in A$, $b_0 \in B$, $x_0 := b_0 - a_0 \neq 0$, and let

$$C = A - B + x_0 = \{a - b + x_0 : a \in A, b \in B\}.$$

Then C is non-empty, open (since A is open), convex (since A, B are convex), contains $0 = a_0 - b_0 + x_0$, and does not contain x_0 (since $a - b + x_0 = x_0$ for $a \in A$, $b \in B$ would imply $a = b$, but A and

B are disjoint).



We shall prove: $\exists\, \lambda \in X^*$ s.t. $\lambda(x) < 1 = \lambda(x_0)\ \forall x \in C$. $\circledast$

Granted this, we conclude as follows: for $a \in A$, $b \in B$, we have $a - b + x_0 \in C$ and therefore $\lambda(a) - \lambda(b) + \lambda(x_0) < 1 = \lambda(x_0)$, thus $\lambda(a) < \lambda(b)$. This implies

$$\sup_{a \in A} \lambda(a) \leq c := \inf_{b \in B} \lambda(b).$$

But since A is open, the supremum on the left is not attained at any $a_0 \in A$. (Otherwise $t \mapsto \lambda(a_0 + tv)$ would have a minimum at $t=0$ for all $v \in X$, so $\lambda(v) = 0\ \forall v \in X \Rightarrow \lambda = 0$ ↯.)

To prove $\circledast$, we proceed as follows.

Define the Minkowski functional

$$p : X \longrightarrow \mathbb{R}, \qquad p(x) = \inf\{\lambda > 0 : \underset{\substack{\Updownarrow \\ \frac{x}{\lambda} \in C}}{x \in \lambda C}\}.$$

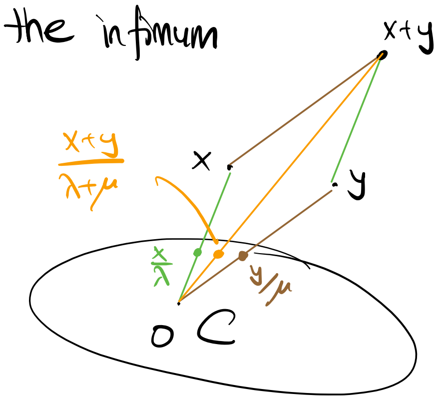
- Claim p is sublinear.

Indeed, $p(\alpha x) = \alpha p(x)$ for $\alpha \geq 0$, $x \in X$ follows directly from the definition. Let now $x, y \in X$, and let $\lambda, \mu > 0$ be s.t. $\frac{x}{\lambda}, \frac{y}{\mu} \in C$. Since C is convex,

$$\frac{\lambda}{\lambda+\mu}\frac{x}{\lambda} + \frac{\mu}{\lambda+\mu}\frac{y}{\mu} = \frac{x+y}{\lambda+\mu} \in C.$$

Thus $p(x+y) \le \lambda+\mu$, so taking the infimum over all possible λ, μ gives

$$p(x+y) \le p(x)+p(y).$$



- Since C is open, $p(x)<1$ for $x \in C$; since $x_0 \notin C$, $p(x_0) \ge 1$.

Define $f: \operatorname{span}\{x_0\} \to \mathbb{R}$, $f(tx_0)=t$. Then

$$f(tx_0) = t \le \begin{cases} t\,p(x_0) = p(tx_0), & t \ge 0, \\ 0 \le p(tx_0) & t<0. \end{cases}$$

By Hahn–Banach, we can extend f to a linear map $\lambda: X \to \mathbb{R}$ satisfying $\lambda(x) \le p(x)$ $\forall x \in X$, and thus $\circledast$.

- We still need to show that $\lambda \in X^*$. But

$$|\lambda(x)| = \max\{\lambda(x), \lambda(-x)\} \le \max\{p(x), p(-x)\}. \quad \boxplus$$

If $R>0$ is s.t. $B_R(0) \subset C$, then $\frac{Rx}{2\|x\|} \in C$ for all $x \in X$, $x \ne 0$, and thus $p(\pm x) \le \frac{2}{R}\|x\|$. $\boxplus$ then gives

$$|\lambda(x)| \le \frac{2}{R}\|x\|,$$ as desired.

(ii) We deduce this from (i) as follows. We claim that $\exists r>0$ s.t. $\mathcal{U} := \bigcup_{a \in A} B_r(a)$ (which is open and convex) satisfies $\mathcal{U} \cap B = \emptyset$.

Granted this, we get from part (i) $\lambda \in X^*$, $c \in \mathbb{R}$ s.t. $\lambda(u) < c \le \lambda(b)$ $\forall u \in \mathcal{U}$, $b \in B$. Since A is compact, $\exists a_0 \in A$ s.t. $\lambda(a_0) = \sup_{a \in A} \lambda(a)$; thus $\sup_{a \in A} \lambda(a) = \lambda(a_0) < c \le \inf_{b \in B} \lambda(b)$, as desired.

To prove the claim, suppose that for some sequence $\{r_k\} \subset \mathbb{R}$ with $r_k \downarrow 0$, $\exists\, a_k \in A$, $b_k \in B$ s.t. $b_k \in B_{r_k}(a_k)$. Since A is compact, WLOG $a_k \longrightarrow a \in A$. But then

$$\begin{aligned} \| b_k - a \| &\leq \| b_k - a_k \| + \| a_k - a \| \\ &< r_k \quad + \| a_k - a \| \\ &\xrightarrow{k \to \infty} 0, \end{aligned}$$

hence $b_k \longrightarrow a$. But B is closed $\Rightarrow a \in B$. This contradicts our assumption that $A \cap B = \emptyset$. $\square$

Application 1: convex sets and weak closures/weak convergence.

Theorem Let X be a normed $\mathbb{R}$-vector space, and let $A \subset X$ be convex. Then $\overline{A} = \overline{A}_w$.

Proof We already know $\overline{A} \subset \overline{A}_w$. Suppose $\exists\, x_0 \in \overline{A}_w \setminus \overline{A}$. By the separation theorem applied to $\{x_0\}$ and $\overline{A}$, $\exists\, \lambda \in X^*$ s.t. $\lambda(x_0) < \inf\limits_{x \in \overline{A}} \lambda(x) = \inf\limits_{x \in A} \lambda(x) =: c$. Therefore, $\{ x \in X : \lambda(x) < c \}$ is a weakly open neighborhood of x_0 disjoint from A. This contradicts $x_0 \in \overline{A}_w$. $\square$

Corollary (Mazur's Lemma) X = normed $\mathbb{R}$-vector space, $\{x_k\}_{k\in\mathbb{N}} \subset X$, $x \in X$ with $x_k \xrightarrow{w} x$. Then $\exists$ sequence $\{y_\ell\}_{\ell\in\mathbb{N}}$ of convex combinations of the x_k — more precisely,

$$y_\ell = \sum_{k=1}^{\ell} a_{\ell k} x_k \quad \text{with } 0 \le a_{k\ell} \le 1, \ \sum_{k=1}^{\ell} a_{\ell k} = 1$$

— so that $y_\ell \xrightarrow{\ell\to\infty} x$.

Proof Let $K = \overline{\operatorname{conv}\{x_1, x_2, \dots\}}$, where $\operatorname{conv}\Omega$ = (set of all convex combinations of elements of Ω) is the convex hull of $\Omega \subset X$. (This is the smallest convex set containing Ω.)

Then K is closed and convex, so $K = \overline{K} = \overline{K}_w$. Since $x_k \xrightarrow{w} x$, we have $x \in \overline{K}_w$. Therefore, $x \in K$ is indeed the limit of a sequence of convex combinations of the $\{x_k\}$. $\square$

This gives an important class of w.s.l.s.c. functionals for which the direct method of the calculus of variations applies:

Proposition Let X be a normed $\mathbb{R}$-vector space, let $M \subset X$ be closed and convex, and let $F: M \to \mathbb{R}$ be continuous and convex (i.e. $F(tx + (1-t)y) \le tF(x) + (1-t)F(y)$ $\forall x, y \in X$, $t \in [0,1]$).

Then F is w.s.l.s.c.

Proof Let $\{x_k\}_{k\in\mathbb{N}} \subset M$ with $x_k \xrightarrow{w} x_0 \in M$; upon passing to a subsequence, we may assume $F(x_k) \to \liminf_{\ell\to\infty} F(x_\ell) =: \alpha \ge -\infty$.

By Mazur's Lemma, applied to $\{x_{k_0}, x_{k_0+1}, \dots\}$ for some fixed $k_0 \in \mathbb{N}$, $\exists$ convex combinations $y_\ell = \sum_{k=k_0}^{\ell} a_{\ell k} x_k$ $(\ell \geq k_0)$ s.t. $y_\ell \to x_0$. Since F is continuous and convex,

$$F(y_\ell) \leq \sum_{k=k_0}^{\ell} a_{\ell k} F(x_k) \leq \sup_{k \geq k_0} F(x_k).$$

In the limit $\ell \to \infty$, get $F(x_0) \leq \sup_{k \geq k_0} F(x_k)$.

Taking $k_0 \to \infty$ gives $F(x_0) \leq \alpha$. (In particular, $\alpha > -\infty$.) $\square$

Application 2: reflexivity and weak compactness

Our goal is to prove:

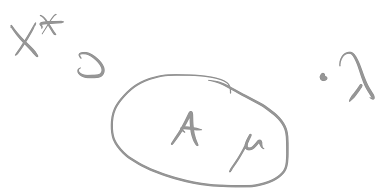
Theorem A Banach space X is reflexive if and only if the unit ball $B = \{x \in X : \|x\| \leq 1\}$ is weakly compact.

We already saw "$\Rightarrow$", as an immediate application of Banach–Alaoglu. The converse requires more work.

Lemma Let X be a real Banach space, and let $A \subset X^*$ be a convex weak-$*$-closed set. Let $\lambda \in X^* \setminus A$.

Then $\exists x_0 \in X$ s.t. $\lambda(x_0) > \sup_{\mu \in A} \mu(x_0)$

Remark The separation theorem only gives $\Phi \in X^{**}$ s.t. $\Phi(\lambda) > \sup_{\mu \in A} \Phi(\mu)$.



Proof Since A is weak-$*$-closed, $\exists$ weak-$*$-open neighborhood $\mathcal{U} = \bigcap_{n=1}^{N} \{\mu \in X^* : |\mu(x_n)| < \varepsilon\}$ (with $x_1, \dots, x_N \in X$, $\varepsilon > 0$)

s.t. $(\lambda + \mathcal{U}) \cap A = \emptyset$ still, or equivalently $\lambda \notin A + \mathcal{U}$.

Applying the separation theorem to $\{\lambda\}$ and $A+\mathcal{U}$ gives

$\Phi \in X^{**}$ s.t. $\sup_{\mu \in A} \Phi(\mu) \le \sup_{\mu \in A+\mathcal{U}} \Phi(\mu) \le \Phi(\lambda)$.

Since $\Phi \neq 0$, $\exists$ $\nu \in \mathcal{U}$ s.t. $\Phi(\nu) =: \delta > 0$. $\Rightarrow \sup_{\mu \in A} \Phi(\mu) + \delta \le \Phi(\lambda)$, so

$$\sup_{\mu \in A} \Phi(\lambda) < \Phi(\lambda).$$

Claim: This Φ is necessarily of the form $\Phi = \iota(x_0)$, $x_0 \in X$.

(In some sense, Φ is much more restricted than in the Remark since it separates $\{\lambda\}$ from the rather elongated set $A+\mathcal{U}$.)

Indeed, note that for $\mu \in \mathcal{U}$, we have

$$\Phi(\mu) < \Phi(\lambda) - \Phi(a) \quad \text{for all } a \in A.$$

$$\Rightarrow \sup_{\mu \in \mathcal{U}} \Phi(\mu) =: C < \infty.$$

Now, if $\mu \in X^*$, $\mu(x_1) = \cdots = \mu(x_n) = 0$, then not only $\mu \in \mathcal{U}$, but also $\alpha\mu \in \mathcal{U}$ $\forall \alpha \in \mathbb{R}$; so

$$|\Phi(\mu)| \le \frac{C}{\alpha} \quad \forall\, \alpha > 0 \Rightarrow \Phi(\mu) = 0.$$

We now have a linear algebra setup: $\Phi : X^* \to \mathbb{R}$ is linear and vanishes on the space $\bigcap_{n=1}^{N} \ker \underbrace{\iota(x_n)}_{\text{linear maps } X^* \to \mathbb{R}}$. Therefore,

Φ is a linear combination of $\{\iota(x_1), \ldots, \iota(x_N)\}$. (Exercise.)

$$\Rightarrow \Phi = \iota\Big(\underbrace{\sum_{n=1}^{N} c_n x_n}_{=: x}\Big) \quad \text{for some } c_1, \ldots, c_N \in \mathbb{R}.$$

This completes the proof. $\square$

Theorem (Goldstine). Let X be a Banach space, and let $\imath : X \to X^{**}$ be the canonical inclusion. Let $B = \{x \in X : \|x\| \leq 1\}$, $B^{**} = \{\Phi \in X^{**} : \|\Phi\|_{X^{**}} = 1\}$. Then $\imath(B) \subset B^{**}$ is weak-$*$-dense.

Proof Since B^{**} is weak-$*$-compact, we certainly have that $A := \overline{\imath(B)}^{w*} \subseteq B^{**}$ is closed. Suppose $A \neq B^{**}$, and let $\Phi_0 \in B^{**} \setminus A$. The Lemma, applied with X^*, gives us $\lambda \in X^*$ s.t.
$$\sup_{\Psi \in A} \Psi(\lambda) = \sup_{\Psi \in \imath(B)} \Psi(\lambda) = \sup_{x \in B} \lambda(x) = \|\lambda\|_{X^*}$$
$$< \Phi_0(\lambda) \leq \underbrace{\|\Phi_0\|_{X^{**}}}_{\leq 1} \|\lambda\|_{X^*} \leq \|\lambda\|_{X^*},$$
which is absurd. $\square$

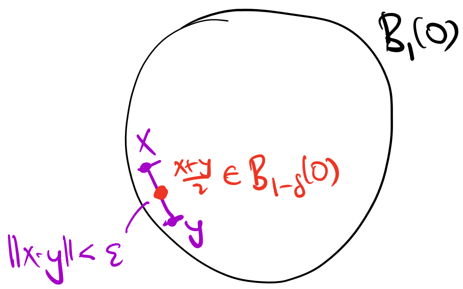
We can now finish the proof of the Theorem: suppose $B \subset X$ is weakly compact. Equivalently, $\imath(B) \subset X^{**}$ is weak-$*$-compact, in particular closed. But Goldstine's Theorem gives
$$\imath(B) = \overline{\imath(B)}^{w*} = B^{**}.$$
Therefore, also $\imath(X) = X^{**}$, and X is reflexive. $\square$

Application 3: uniform convexity and reflexivity

Definition A Banach space X is *uniformly convex* if

$\forall \varepsilon > 0 \ \exists \delta > 0$ s.t.

$$x, y \in X,\ \|x\| \le 1,\ \|y\| \le 1,\ \|x-y\| > \varepsilon \Rightarrow \left\| \frac{x+y}{2} \right\| < 1 - \delta.$$



Examples (1) Every Hilbert space is uniformly convex since

$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$ implies

$$\left\| \frac{x+y}{2} \right\|^2 = \frac{\|x\|^2 + \|y\|^2}{2} - \frac{\|x-y\|^2}{4} < \frac{1+1}{2} - \frac{\varepsilon^2}{4} = 1 - \frac{\varepsilon^2}{4}$$

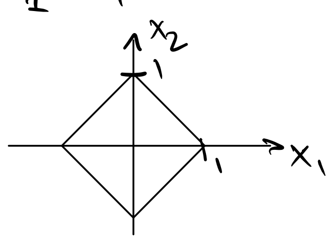
when $\|x\| \le 1, \|y\| \le 1, \|x-y\| > \varepsilon$.

So $\delta = 1 - (1 - \frac{\varepsilon^2}{4})^{\frac{1}{2}}$ works.

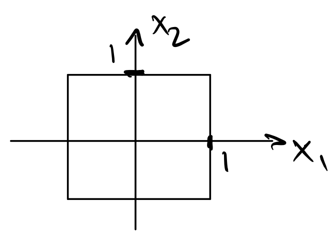
(2) $(\mathbb{R}^2, \|\cdot\|)$

(3) $(\mathbb{R}^2, \|\cdot\|_1)$ and $(\mathbb{R}^2, \|\cdot\|_\infty)$ are not uniformly convex: the unit spheres contain line segments.

$\|x\|_1 = |x_1| + |x_2| = 1$



$\|x\|_\infty = \max\{|x_1|, |x_2|\} = 1$



In particular, (2) and (3) show that uniform convexity is a property of a norm, not of an equivalence class of norms (unlike completeness, weak(-*) topology, etc.)

(4) $\ell^1, \ell^\infty, L^1, L^\infty$ are not uniformly reflexive; ℓ^p, L^p $(1<p<\infty)$ are (see below).

Uniformly convex spaces have many nice properties.

Lemma X uniformly convex, $Y \subset X$ closed subspace, $x_0 \in X$
$\Rightarrow \exists!\ y_0 \in Y$ with $\|x_0 - y_0\| = d(x_0, Y)$.

Proof Exercise. $\square$

Theorem (Milman–Pettis.) X uniformly convex $\Rightarrow$ X reflexive.

Proof · Let $\Phi \in X^{**}$, $\|\Phi\|_{X^{**}} = 1$. Since $\iota(B) \subset X^{**}$ (with $B = \{x \in X : \|x\| \le 1\}$) is closed, it suffices to show:

$$\forall \varepsilon > 0\ \exists\ x \in X \text{ s.t. } \|\Phi - \iota(x)\|_{X^{**}} \le \varepsilon. \quad (*)$$

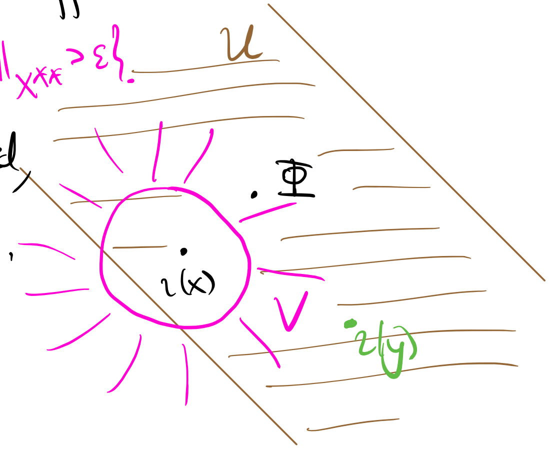
· Given $\varepsilon > 0$, fix $\delta > 0$ as in the definition of uniform convexity. Take $\lambda \in X^*$, $\|\lambda\|_{X^*} = 1$, with

$$\Phi(\lambda) > 1 - \frac{\delta}{2},$$

and set $\mathcal{U} = \{\Psi \in X^{**} : |\langle \Phi - \Psi, \lambda \rangle| < \frac{\delta}{2}\}$ (which is a weak-$*$-neighborhood of Φ in X^{**}).

· Goldstine's Theorem implies that $\exists\ x \in B$ s.t. $\iota(x) \in \mathcal{U}$.
We claim that x satisfies $(*)$. Suppose this is not the case. Set $V = \{\Psi \in X^{**} : \|\Psi - \iota(x)\|_{X^{**}} > \varepsilon\}$.
Since $\overline{B_\varepsilon}(\iota(x))$ is weak-$*$-closed, V is a weak-$*$-neighborhood of Φ.

$\overset{\text{Goldstine}}{\Longrightarrow} \exists\ y \in B : \ \iota(y) \in \mathcal{U} \cap V$.



Since $x, y \in \mathcal{U}$, $|\lambda(x) - \Phi(\lambda)| < \frac{\delta}{2}$,

$|\lambda(y) - \Phi(\lambda)| < \frac{\delta}{2}$.

$\Rightarrow 2(1-\frac{\delta}{2}) < 2\,\Phi(\lambda) < \lambda(x+y) + \delta \leq \|x+y\| + \delta$

$\Rightarrow \|\frac{x+y}{2}\| > 1-\delta \Rightarrow \|x-y\| < \varepsilon$, hence also

$\|\imath(x) - \imath(y)\| < \varepsilon$,

contradiction since $\imath(y) \in V$. $\square$

And finally, the long promised...

Theorem (Clarkson.) Let $1 < p < \infty$, $\Omega \subset \mathbb{R}^n$ measurable. Then $L^p(\Omega)$ is uniformly convex (and thus reflexive).

Lemma Let $x, y \in \mathbb{R}$, $\varepsilon > 0$. Then $\exists\, \delta > 0$ s.t.

$$4|x-y|^p \geq \varepsilon^p(|x|^p + |y|^p) \Rightarrow \left|\frac{x+y}{2}\right|^p \leq (1-\delta)\frac{|x|^p + |y|^p}{2} \quad \#$$

Proof Trivial for $x=y=0$. Otherwise, scale $(x,y) \rightsquigarrow (tx, ty)$ so that $|x|+|y|=1$. (Note: everything in $\#$ gets multiplied by t^p when $x, y \rightsquigarrow tx, ty$.) Suppose $\delta > 0$ does not exist; then for $\delta = \frac{1}{n}$ we get (x_n, y_n), $|x_n|+|y_n|=1$, so that $4|x_n - y_n|^p \geq \varepsilon^p(|x_n|^p + |y_n|^p)$ but

$$\left|\frac{x_n+y_n}{2}\right|^p > \left(1-\frac{1}{n}\right)\frac{|x_n|^p+|y_n|^p}{2}. \quad \circledast$$

Passing to a convergent subsequence $(x_n, y_n) \to (x_0, y_0)$, $|x_0|+|y_0|=1$, we obtain from $\circledast$ $\left|\frac{x_0+y_0}{2}\right|^p \geq \frac{|x_0|^p + |y_0|^p}{2}$

Exercise: for $p > 1$, this inequality holds iff $x_0 = y_0$. (Hint: $a \mapsto a^p$ is convex.)

But then $\circledast$ gives $x_0 = y_0 = 0$, $\lightning$ to $|x_0| + |y_0| = 1$. $\square$

Proof of the Theorem Let $\|f\|_{L^p} = \|g\|_{L^p} = 1$, $\|f-g\|_{L^p} > \varepsilon$.

Let $S := \{x \in \Omega : 4|f(x) - g(x)|^p \geq \varepsilon^p (|f(x)|^p + |g(x)|^p)\}$ be the set where f and g differ by "a fair bit". Then

$$\int_{\Omega \setminus S} |f(x) - g(x)|^p \, dx \leq \frac{\varepsilon^p}{4} \left(\|f\|_{L^p}^p + \|g\|_{L^p}^p \right) = \frac{\varepsilon^p}{2}$$

$$\Rightarrow \int_S |f(x) - g(x)|^p \, dx > \frac{\varepsilon^p}{2}.$$

By the Lemma, $x \in S$ implies $\left| \frac{f(x) + g(x)}{2} \right|^p \leq (1-\delta) \frac{|f(x)|^p + |g(x)|^p}{2}$, so

$$1 - \int_\Omega \left| \frac{f(x) + g(x)}{2} \right|^p dx = \int_\Omega \frac{|f(x)|^p + |g(x)|^p}{2} - \left| \frac{f(x) + g(x)}{2} \right|^p dx$$

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

$$\geq \delta \int_S \frac{|f(x)|^p + |g(x)|^p}{2} \, dx$$

$$\boxed{p > 1,\ a, b \in \mathbb{R} \Rightarrow \frac{a^p + b^p}{2} \geq \left(\frac{a+b}{2} \right)^p} \xrightarrow{a = f(x),\ b = -g(x)} \geq \delta \int_S \left| \frac{f(x) - g(x)}{2} \right|^p dx$$

$$> \frac{\delta}{2^p} \frac{\varepsilon^p}{2}, \text{ as desired.} \quad \square$$