# $L^{p}$ Spaces and related Banach Space Theory for Math 502 

Eric A. Carlen ${ }^{1}$<br>Rutgers University

March 6, 2017

## Contents

1 Convex Functions ..... 1
1.1 Continuity and lower-semicontinuity of convex functions ..... 1
1.2 Convex functions on $\mathbb{R}$ ..... 2
1.3 The subgradient of a convex function ..... 4
1.4 The Legendre Transform ..... 6
$2 L^{p}$ norms and their close relatives ..... 9
2.1 The $L^{p}$ norm and Hölder's Inequality ..... 9
2.2 Orlicz norms ..... 11
2.3 Duality in Orlicz spaces ..... 14
3 Uniform convexity and uniform smoothness ..... 15
3.1 Uniform convexity ..... 15
3.2 First applications of uniform convexity ..... 16
3.3 Uniform smoothness ..... 19
4 Uniform convexity and smoothness in $L^{p}$ spaces ..... 24

## 1 Convex Functions

### 1.1 Continuity and lower-semicontinuity of convex functions

1.1 DEFINITION (Convex Function). A function $\phi$ on defined on a convex subset $C$ of a real vector space $X$ with values in $(-\infty, \infty]$ is convex in case for all $x, y \in C$ and all $\lambda \in(0,1)$,

$$
\begin{equation*}
\phi(\lambda x+(1-\lambda) y) \leq \lambda \phi(x)+(1-\lambda) \phi(y), \tag{1.1}
\end{equation*}
$$

and $\phi$ is strictly convex in case this inequality is strict for all $x \neq y$ and all $\lambda \in(0,1)$. A convex function $\phi$ on $X$ is proper in case $\phi(x)<\infty$ for at least one $x \in X$.

[^0]1.2 DEFINITION (Epigraph). Let $\phi$ be a function from some set $X$ to $(-\infty, \infty]$. The epigraph of $\phi, \operatorname{Epi}(\phi)$, is the subset of $X \times \mathbb{R}$ consisting of points $(x, t)$ such that $f(x) \geq t$. Note that $\operatorname{Epi}(\phi)=\emptyset$ if and only if $\phi(x)=\infty$ for all $x$.
1.3 LEMMA. Let $(X, \mathscr{O})$ be a topological vector space. Then a function $\phi$ on $X$ with values in $(-\infty, \infty]$ is convex if and only if $\operatorname{Epi}(\phi)$ is convex, and is lower-semicontinuous if and only if $\operatorname{Epi}(\phi)$ is closed in the product topology.

Proof. This is elementary and is left to the reader.
Let $(X, \mathscr{O})$ be a topological vector space, and let $\phi$ be a convex function on $X$. Then $\operatorname{Epi}(\phi)$ is convex, and since the closure of a convex set is convex, $\overline{\operatorname{Epi}(\phi)}$ is closed and convex. Define a convex function $\bar{\phi}$ by

$$
\bar{\phi}(x)=\left\{\begin{array}{cl}
\infty & \{x\} \times \mathbb{R} \cap \overline{\operatorname{Epi}(\phi)}=\emptyset \\
\inf \{t:(x, t) \in x \in \overline{\operatorname{Epi}(\phi)}\} & \text { otherwise } .
\end{array}\right.
$$

Then $\bar{\phi}$ is a lower-semicontinuous convex function such that $\bar{\phi}(x) \leq \phi(x)$ for all $x$. By construction and by Lemma 1.3, $\bar{\phi}$ is the largest (in the usual ordering) lower-semicontinuous convex function that is dominated by $\phi$ pointwise. The function $\bar{\phi}$ is called the lower-semicontinuous regularization of $\phi$.
1.4 EXAMPLE. Define $\phi: \mathbb{R} \rightarrow(-\infty, \infty]$ by

$$
\phi(x)=\left\{\begin{array}{ll}
0 & x \in(-1,1) \\
1 & x \in\{-1,1\} \\
\infty & x \notin[-1,1]
\end{array} \quad \text { Then } \quad \bar{\phi}(x)= \begin{cases}0 & x \in[-1,1] \\
\infty & x \notin[-1,1] .\end{cases}\right.
$$

### 1.2 Convex functions on $\mathbb{R}$

We first focus on the important special case in which the vector space $X$ is simply $\mathbb{R}$. In this case the set $C$ on which $\phi$ is finite is an interval. We may consider $\phi$ as defined on all of $\mathbb{R}$ by setting $\phi(x)=\infty$ for $x \notin C$. In what follows, when we speak of a real-valued convex function on $C \subset \mathbb{R}$, we shall always assume that $\phi(x)=\infty$ for $x \notin C$.
1.5 THEOREM. Let $\phi$ be a real valued convex function on an interval $C \subset \mathbb{R}$. Then $\phi$ is continuous on the interior of $C$. Moreover, for $a, b, c, d \in C$

$$
\begin{equation*}
b-a=d-c \quad \text { and } \quad a<c \quad \Rightarrow \quad \phi(b)-\phi(a) \leq \phi(d)-\phi(c) . \tag{1.2}
\end{equation*}
$$

That is, the increment of $\phi$ over an interval increases as the interval is translated to the right. If $\phi$ is strictly convex, the inequality in (1.2) is strict. Conversely, let $\phi$ be any function that is continuous and finite on an open interval $C$, and is such that (1.2) is valid for all $a, b, c, d \in C$. Then $\phi$ is convex.

Proof. We first prove (1.2) Note that $b, c \in[a, d]$, and hence for some $\lambda, \beta \in(0,1), b=\lambda a+(1-\lambda) d$ and $c=\beta a+(1-\beta) d$. Solving for $\lambda$ and $\beta$, we find

$$
\lambda=1-\beta=\frac{d-b}{d-a} .
$$

Since $\lambda=1-\beta, b=\lambda a+(1-\lambda) d$ and $c=\lambda d+(1-\lambda) a$. It now follows from the definition of convexity that

$$
\phi(b) \leq \lambda \phi(a)+(1-\lambda) \phi(d) \quad \text { and } \quad \phi(c)=\lambda \phi(d)+(1-\lambda) \phi(a) .
$$

Adding these two inequalities yields $\phi(b)+\phi(c) \leq \phi(a)+\phi(d)$, with strict inequality if $\phi$ is strictly convex, and this proves the first assertion, and evidently if $\phi$ is strictly convex, all of the inequalities are strict.

Fix any $a$ in the interior of $C$; we shall show that $\phi$ is continuous at $a$. For any $b \in C, b \neq a$, use a telescoping sum expansion to write

$$
\begin{equation*}
\phi(b)-\phi(a)=\sum_{j=1}^{n}\left[\phi\left(a+(b-a) \frac{j}{n}\right)-\phi\left(a+(b-a) \frac{j-1}{n}\right)\right] . \tag{1.3}
\end{equation*}
$$

Choose $\delta>0$ so that $a \pm \delta \in C$. By (1.2), for $b=a+\delta$, the first term in the sum is the least, and hence $\phi(a+\delta / n)-\phi(a) \leq \frac{\phi(a+\delta)-\phi(a)}{n}$ Likewise, for $b=a-\delta$, the same reasoning yields $\frac{\phi(a)-\phi(a-\delta)}{n} \leq \phi(a)-\phi(a-\delta / n)$. Again by (1.2), $\phi(a)-\phi(a-\delta) \leq \phi(a+\delta / n)-\phi(a)$. Altogether,

$$
\begin{equation*}
\frac{\phi(a)-\phi(a-\delta)}{n} \leq \phi(a)-\phi(a-\delta / n) \leq \phi(a+\delta / n)-\phi(a) \leq \frac{\phi(a+\delta)-\phi(a)}{n} . \tag{1.4}
\end{equation*}
$$

Taking $n \rightarrow \infty$, we conclude $\lim _{n \rightarrow \infty} \phi(a-\delta / n)=\phi(a)=\lim _{n \rightarrow \infty} \phi(a+\delta / n)$.
We now claim that $\lim _{x \downarrow a} \phi(x)=\phi(a)$. Suppose that $\lim _{\sup }^{x \downarrow a}$ $\phi(x)>\phi(a)$. Then for some $\epsilon>0$, there is an infinite sequence $\left\{t_{m}\right\}_{m \in \mathbb{N}}$ contained in $(a, a+\delta)$ such that $\lim _{m \rightarrow \infty} t_{m}=a$ and $\phi\left(t_{m}\right) \geq \phi(a)+\epsilon$ for all $m \in \mathbb{N}$. Choose $n$ so that $\phi(a+\delta / n) \leq \phi(a)+\epsilon / 2$. Then there exists $m \in \mathbb{N}$ such that $t_{m} \in(a, a+\delta / n)$ and thus $\lambda \in(0,1)$ such that $t_{m}=\lambda a+(1-\lambda)(a+\delta / n)$. Then $\phi\left(t_{m}\right) \leq$ $\lambda \phi(a)+(1-\lambda)(\phi(a)+\epsilon / 2)<\phi(a)+\epsilon / 2$. This contradiction shows that limsup $\operatorname{su}_{x \downarrow a} \phi(x) \leq \phi(a)$.

Next, suppose that $\lim _{\inf _{x \downarrow a}} \phi(x)<\phi(a)$. Then for some $\epsilon>0$, there is an infinite sequence $\left\{t_{m}\right\}_{m \in \mathbb{N}}$ contained in $(a, a+\delta)$ such that $\lim _{m \rightarrow \infty} t_{m}=a$ and $\phi\left(t_{m}\right) \leq \phi(a)-\epsilon$ for all $m \in \mathbb{N}$. Choose $n$ so that $\phi(a+\delta / n) \geq \phi(a)-\epsilon / 2$. Choose $n \in \mathbb{N}$ so that $\phi(a+\delta / n) \geq \phi(a)-\epsilon / 2$, and $a+\delta / n<t_{1}$. Choose $m$ so that $t_{m}<a+\delta / n$. Then $a+\delta / n \in\left(t_{m}, t_{1}\right)$ and there exists $\lambda \in(0,1)$ such that $a+\delta / n=\lambda t_{m}+(1-\lambda) t_{1}$. Then $\phi(a+\delta / n) \leq \lambda \phi\left(t_{m}\right)+(1-\lambda)\left(\phi\left(t_{m}\right)<\phi(a)-\epsilon / 2\right.$. This contradiction shows that $\lim ^{\inf }{ }_{x \downarrow a} \phi(x) \geq \phi(a)$.

Altogether, we have shown that $\phi$ is right continuous at $a$. We could repeat the same analysis to show that $\phi$ is also left continuous at $a$, but observe that the function $\psi(x):=\phi(2 a-x)$ is convex and finite on an open interval about $a .(\operatorname{Epi}(\psi)$ is just the reflection of $\operatorname{Epi}(\phi)$ about the vertical line $x=a)$. By what we have just proved, $\psi$ is right continuous at $a$. But then since $\phi$ is the reflection of $\phi$ about $x=a, \phi$ is left continuous at $x=a$. Altogether, the continuity of $\phi$ is proved.

For the converse, let $x<y \in C$ and $\lambda \in(0,1)$. We must show that when (1.2) is valid for all $a, b, c, d \in C$, then $\phi(\lambda x+(1-\lambda) y) \leq \lambda \phi(x)+(1-\lambda) \phi(y)$. By the continuity of $\phi$, it suffices to do this when $\lambda$ is a dyadic rational; i.e., $\lambda=k / 2^{n}$ for some $k, n \in \mathbb{N}$ with $k<2^{n}$. For $n=1$, note define $z=(x+y) / 2$ and note that

$$
\frac{1}{2} \phi(x)+\frac{1}{2} \phi(y)-\phi\left(\frac{x+y}{2}\right)=\frac{1}{2}(\phi(y)-\phi(z))-\frac{1}{2}(\phi(z)-\phi(x))>0
$$

because $y-z=z-x$. That is, when (1.2) is valid for all $a, b, c, d \in C$, then

$$
\begin{equation*}
\phi(\lambda x+(1-\lambda) y) \leq \lambda \phi(x)+(1-\lambda) \phi(y) \tag{1.5}
\end{equation*}
$$

for all $x, y \in C$ and $\lambda=1 / 2$.
This is the first step of an inductive proof that the same is true whenever $\lambda=j 2^{-m}, j, m \in \mathbb{N}$ and $j<2^{m}$. We suppose that this has been shown whenever $m<n$.

Now fix $\lambda=j / 2^{-n} \in(0,1)$. Let $k, \ell \in \mathbb{N}$ such that $k+\ell=j$ and $k, \ell \leq 2^{n-1}$. (If $j$ is even take $k=\ell=j / 2$, and if $j$ is odd, take $k$ to be the integer part of $j / 2$.) Then for all $x, y$,

$$
\begin{equation*}
\lambda x+(1-\lambda) y=\frac{j x+\left(2^{n}-j\right) y}{2^{n}}=\frac{1}{2}\left(\frac{k x+\left(2^{n-1}-k\right) y}{2^{n-1}}\right)+\frac{1}{2}\left(\frac{\ell x+\left(2^{n-1}-\ell\right) y}{2^{n-1}}\right) \tag{1.6}
\end{equation*}
$$

Since (1.5) is true for $\lambda=1 / 2$,

$$
\phi(\lambda x+(1-\lambda) y) \leq \frac{1}{2} \phi\left(\frac{k x+\left(2^{n-1}-k\right) y}{2^{n-1}}\right)+\frac{1}{2} \phi\left(\frac{\ell x+\left(2^{n-1}-\ell\right) y}{2^{n-1}}\right) .
$$

By the inductive hypothesis,

$$
\phi\left(\frac{k x+\left(2^{n-1}-k\right) y}{2^{n-1}}\right) \leq \frac{k}{2^{n-1}} \phi(x)+\frac{2^{n-1}-k}{2^{n-1}} \phi(y)
$$

and likewise with $\ell$ in place of $k$. Using these inequalities in (1.6) shows that (1.5) is valid for $\lambda=j 2^{-n}$, completing the inductive proof.

### 1.3 The subgradient of a convex function

1.6 THEOREM. Let $\phi$ be a real valued convex function on an interval $C \subset \mathbb{R}$. For all $a, b, c \in C$, $a<b<c$,

$$
\begin{equation*}
\frac{\phi(c)-\phi(a)}{c-a} \geq \frac{\phi(b)-\phi(a)}{b-a} \tag{1.7}
\end{equation*}
$$

If $\phi$ is strictly convex, the inequality in (1.7) is strict.
Proof. Because $\phi$ is continuous, it suffices to consider rational values of $b-a$ and $c-a$. Choosing a common denominator $n$, we can write $b-a=\frac{k}{n}$ and $c-a=\frac{m}{n}$ with $m>k$. Define a sequence $\left\{a_{j}\right\}$ by

$$
a_{j}=\phi\left(a+\frac{j}{n}\right)-\phi\left(a+\frac{j-1}{n}\right) .
$$

By (1.2), this is an increasing sequence, and hence,

$$
\frac{\phi(b)-\phi(a)}{b-a}=n \frac{\phi(a+k / n)-\phi(a)}{k}=n\left(\frac{1}{k} \sum_{j=1}^{k} a_{j}\right) .
$$

Likewise, $\frac{\phi(c)-\phi(a)}{\tilde{c}-a}=n\left(\frac{1}{m} \sum_{j=1}^{m} a_{j}\right)$. Since $a_{j}$ increases with $j, \frac{1}{k} \sum_{j=1}^{k} a_{j} \leq \frac{1}{m} \sum_{j=1}^{m} a_{j}$ for $m \geq k$. This proves (1.7). By theorem 1.5, if $\phi$ is strictly convex, then $a_{j+1}>a_{j}$ for all $j$, and then there is strict inequality in (1.7).

Let $\phi$ be convex on $\mathbb{R}$ and finite on $C$, For each $s \in C^{\circ}$, define

$$
\begin{equation*}
\sigma_{+}^{\phi}(s)=\lim _{h \rightarrow 0^{+}} \frac{\phi(s+h)-\phi(s)}{h} \quad \text { and } \quad \sigma_{-}^{\phi}(s)=\lim _{h \rightarrow 0^{+}} \frac{\phi(s)-\phi(s-h)}{h} . \tag{1.8}
\end{equation*}
$$

The limit defining $\sigma_{+}^{\phi}$ exists by (1.7), and then limit defining $\sigma_{-}^{\phi}$ also exists by (1.7), but applied to the convex functions $\phi(-s)$. We refer to $\sigma_{+}^{\phi}(s)$ as the right derivative of $\phi$ at $s$, and to $\sigma_{-}^{\phi}(s)$ as the left derivative of $\phi$ at $s$. If is clear from Theorem 1.5 that in general, $\sigma_{+}^{\phi}(s) \geq \sigma_{-}^{\phi}(s)$. Evidently $\phi$ is differentiable at $s$ if and only if $\sigma_{+}^{\phi}(s)=\sigma_{-}^{\phi}(s)$, so that in this case, $\phi^{\prime}(s)=\sigma_{+}^{\phi}(s)=\sigma_{-}^{\phi}(s)$ represents the slope of $\phi$ at $s$. Since the left derivative of $\phi$ at $s$ is minus the right derivative of $t \mapsto \phi(-t)$ at $t=-s$, we may economize in the formulation of the following theorem by referring only to right derivatives.
1.7 THEOREM (One-sided Derivatives). Let $\phi$ be a convex function on $\mathbb{R}$ that is finite on an interval $C$. for all $a, b \in C^{\circ}, a<b$.

$$
\begin{equation*}
\sigma_{+}^{\phi}(a) \leq \sigma_{-}^{\phi}(b) \tag{1.9}
\end{equation*}
$$

and for all $s \in\left[\sigma_{-}^{\phi}(a), \sigma_{+}^{\phi}(a)\right]$,

$$
\begin{equation*}
\phi(b) \leq \phi(a)+s(b-a) . \tag{1.10}
\end{equation*}
$$

Moreover, if $\phi$ is strictly convex, then both of these inequalities are strict. Finally, for all $a \in C^{\circ}$,

$$
\begin{equation*}
\sigma_{+}^{\phi}(a)=\inf _{b>a} \sigma_{+}^{\phi}(b) \quad \text { and } \quad \sigma_{-}^{\phi}(a)=\sup _{b<a} \sigma_{+}^{\phi}(b) \tag{1.11}
\end{equation*}
$$

In other words, $\sigma_{+}^{\phi}$ is a right-continuous non-decreasing function, and $\sigma_{-}^{\phi}$ is a left-continuous nondecreasing function.

Proof. By Theorem 1.5, for all $0<h<b-a, \phi(b)-\phi(b-h) \leq \phi(a+h)-\phi(a)$. Dividing by $h$ and taking the limit $h \downarrow 0$ yields (1.9).

Next, by the telescoping sum identity (1.3) and Theorem 1.5 to obtain, for $b>a$,

$$
\begin{equation*}
\phi(b)-\phi(a) \geq n(\phi(a+(b-a) / n)-\phi(a)) . \tag{1.12}
\end{equation*}
$$

Multiplying by $1=(b-a) /(b-a)$, yields

$$
\phi(b) \geq \phi(a)+(b-a)\left[\frac{\phi(a+(b-a) / n)-\phi(a)}{(b-a) / n}\right]
$$

Taking the limit $n \rightarrow \infty$ yields $\phi(b) \geq \phi(a)+\sigma_{+}^{\phi}(a)(b-a)$.
For $b<a$, (1.3) and Theorem 1.5 yield

$$
\begin{equation*}
\phi(a)-\phi(b) \leq n(\phi(a)-\phi(a-(a-b) / n)) . \tag{1.13}
\end{equation*}
$$

Multiplying and dividing by $-1=(b-a) /(a-b)$, yileds

$$
\phi(b) \leq \phi(a)+(b-a)\left[\frac{\phi(a)-\phi(a-(a-b) / n)}{(a-b) / n}\right]
$$

Taking the limit $n \rightarrow \infty$ yields $\phi(b) \geq \phi(a)+\sigma_{-}^{\phi}(a)(b-a)$. Therefore, for any $s \in\left[\sigma_{-}^{\phi}(a), \sigma_{+}^{\phi}(a)\right]$, (1.10) is valid.

Next, for $a \in C^{\circ}$, choose $\epsilon>0$, and then $h>0$ so that $\sigma_{+}^{\phi}(a)+\epsilon \geq(\phi(a+h)-\phi(a)) / h$. By the continuity of $\phi$, there is a $\delta>0$ so that $(\phi(a+h)-\phi(a)) / h+\epsilon \geq(\phi(a+\delta+h)-\phi(a+\delta)) / h$. Therefore,

$$
\sigma_{+}^{\phi}(a)+2 \epsilon \geq(\phi(a+\delta+h)-\phi(a+\delta)) / h \geq \sigma_{+}^{\phi}(a+\delta)
$$

The fact that for all $\delta>0, \sigma_{+}^{\phi}(a) \leq \sigma_{+}^{\phi}\left(a_{\delta}\right)$ is an immediate consequence of (1.9). This proves the first identity in (1.11). The second is proved in the same manner.

Theorem 1.7 has the following interpretation: Let $\phi$ be convex and finite on an open interval $C \subset \mathbb{R}$. Then for all $x_{0} \in C$, and all $s \in\left[\sigma_{-}^{\phi}(a), \sigma_{+}^{\phi}(a)\right]$, the affine function $h(x)=s\left(x-x_{0}\right)+\phi\left(x_{0}\right)$ satisfies $h(x) \leq \phi(x)$ for all $x$, with $h\left(x_{0}\right)=\phi\left(x_{0}\right)$ : The graph of $h$, a line, lies below the graph of $\phi$ but touches it at the point $\left(x_{0}, \phi\left(x_{0}\right)\right)$. Such a line is called a supporting line for the graph of $\phi$.

It is evident that the maximum $\phi_{1} \vee \phi_{2}(x)=\max \left\{\phi_{1}(x), \phi_{2}(x)\right\}$ of two convex functions $\phi_{1}, \phi_{2}$ is convex, and more generally, if $\mathcal{C}$ is any set of convex functions, then

$$
\bigvee_{\phi \in \mathcal{C}} \phi(x)=\sup \{\phi(x): \phi \in \mathcal{C}\}
$$

is convex.
Affine functions are evidently convex. For any function $\psi$ on an open interval $C$, convex or not, define $\mathcal{A}_{\psi}$ to be the set of affine functions $h$ such that $h(x) \leq \psi(x)$ for all $x \in C$. Define the function

$$
\operatorname{conv}(\psi)(x)=\sup \left\{\phi(x): \phi \in \mathcal{A}_{\psi}\right\}
$$

By what has been explained just above, $\operatorname{conv}(\psi)$ is a convex function satisfying $\operatorname{conv}(\psi)(x) \leq \psi(x)$ for all $x$. It is the largest such function: If $\phi$ is a convex function such that $\phi(x) \leq \psi(x)$, then at every $x_{0}$, there is a supporting line to the graph of $\phi$ at $x_{0}$, and hence the linear function $h$ whose graph is this supporting line belongs to $\mathcal{A}_{\psi}$. This implies that $\phi\left(x_{0}\right) \leq \operatorname{conv}(\psi)\left(x_{0}\right)$. Moreover, the supremum of any set of continuous functions is lower semi-continuous, so that $\operatorname{conv}(\psi)\left(x_{0}\right)$ is also lower-semicontinuous. The function $\operatorname{conv}(\psi)$ is called the convex envelope of $\psi$. Evidently every proper, lower-semicontinuous convex function is its own convex envelope. This simple fact has an important consequence.
1.8 THEOREM (Jensen's Inequality). Let $(\Omega, \mathcal{M}, \mu)$ be a measure space with $\mu(\Omega)=1$. Then for all real valued convex functions $\phi$ on $\mathbb{R}$, and all measurable function $f$, the negative part of $\phi(|f|)$ is integrable, so that $\int_{\Omega} \phi(|f|) \mathrm{d} \mu$ is well defined, though it may be infinite. Moreover,

$$
\begin{equation*}
\phi\left(\int_{\Omega}|f| \mathrm{d} \mu\right) \leq \int_{\Omega} \phi(|f|) \mathrm{d} \mu . \tag{1.14}
\end{equation*}
$$

and if $\phi$ is strictly convex there is equality if and only if

$$
\begin{equation*}
|f(x)|=\int_{\Omega}|f| \mathrm{d} \mu \tag{1.15}
\end{equation*}
$$

for almost every $x$.
Proof. Let $a:=\int_{\Omega}|f| \mathrm{d} \mu$. By (1.10), $\phi(|f(x)|) \geq \phi(a)+\sigma_{+}^{\phi}(a)(|f(x)|-a)$. Integrating this pointwise inequality yields (1.14). If $\phi$ is strictly convex, this pointwise inequality is strict wherever $|f(x)| \neq a$, and hence the inequality (1.14) is strict unless $|f(x)|=a$ almost everywhere.

### 1.4 The Legendre Transform

Let $\phi$ be a convex function on a closed interval $[a, b]$ where $a=-\infty$ and $b=\infty$ are allowed, and $\phi$ is finite on $(a, b)$. Define

$$
\begin{equation*}
c:=\lim _{x \downarrow a} \sigma_{-}^{\phi}(x) \quad \text { and } \quad d:=\lim _{x \uparrow b} \sigma_{+}^{\phi}(x) . \tag{1.16}
\end{equation*}
$$

By Theorem 1.7, both limits exist, though we may have $c=-\infty$ or $d=\infty$ (or both). Fix $y \in \mathbb{R}$ and consider the function $x \mapsto x y-\phi(x)$.

Suppose that $y \in(c, d)$. Define

$$
\begin{equation*}
x_{0}(y)=\sup \left\{x: \sigma_{-}^{\phi}(x)<y\right\} \quad \text { and } \quad x_{1}(y)=\inf \left\{x: \sigma_{+}^{\phi}(x)>y\right\} \tag{1.17}
\end{equation*}
$$

and note that $-\infty<x_{0}(y) \leq x_{1}(y)<\infty$. By (1.11),

$$
\begin{equation*}
\sigma_{-}^{\phi}\left(x_{0}(y)\right) \leq y \quad \text { and } \quad \sigma_{+}^{\phi}\left(x_{1}(y)\right) \geq y \tag{1.18}
\end{equation*}
$$

Thus, for all $x>x_{1}(y)$,

$$
\begin{aligned}
y x-\phi(x) & \leq y x_{1}(y)+y\left(x-x_{1}(y)\right)-\phi\left(x_{1}(y)\right)-\sigma_{+}^{\phi}\left(x_{1}(y)\right)\left(x-x_{1}(y)\right) \\
& =y x_{1}(y)-\phi\left(x_{1}(y)\right)+\left[y-\sigma_{+}^{\phi}\left(x_{1}(y)\right)\right]\left(x-x_{1}(y)\right) .
\end{aligned}
$$

By (1.18), $\left[y-\sigma_{+}^{\phi}\left(x_{1}(y)\right)\right]\left(x-x_{1}(y)\right) \leq 0$, and therefore $y x-\phi(x) \leq y x_{1}(y)-\phi\left(x_{1}(y)\right)$.
Similar reasoning shows that for all $x<x_{0}(y), y x-\phi(x) \leq y x_{0}(y)-\phi\left(x_{0}(y)\right.$. Therefore,

$$
\begin{equation*}
\sup _{x \in \mathbb{R}}\{y x-\phi(x)\}=\max _{x \in\left[x_{0}(y), x_{1}(y)\right]}\{y x-\phi(x)\}<\infty, \tag{1.19}
\end{equation*}
$$

using the fact that $\phi$ is continuous. Define a function $\phi^{*}$ on $\mathbb{R}$ with values in $(-\infty, \infty]$ by

$$
\begin{equation*}
\phi^{*}(y)=\sup _{x \in \mathbb{R}}\{y x-\phi(x)\} . \tag{1.20}
\end{equation*}
$$

As a supremum of affine functions, $\phi^{*}$ is convex and lower-semicontinuous, and by (1.19), $\phi^{*}$ is finite on the interval $(c, d)$. Now observe that if $\phi$ is finite only at a single point $x_{0}$, and $\sup _{x \in \mathbb{R}}\{y x-\phi(x)\}=y x_{0}-\phi\left(x_{0}\right)$ for all $y \in \mathbb{R}$.
1.9 DEFINITION (Legendre Transform). Let $\phi$ be convex and finite on a non-empty interval $C$. The Legendre transform $\phi^{*}$ of $\phi$ is the function defined by (1.20), which is proper, convex and lower-semicontinuous. A pair of proper lower-semicontinuous functions $\phi$ and $\psi$ such that $\psi=\phi^{*}$ and hence $\phi=\psi^{*}$ is called a dual pair of convex functions. (Note that, by definition, if $\phi$ and $\psi$ are a dual pair of convex functions, they are not only convex, but also proper and lower-semicontinuous.)
1.10 EXAMPLE. For $1<p<\infty$, define $\phi_{p}(x):=p^{-1}|x|^{p}$. Note that $\phi_{p}$ is continuously differentiable with $\phi_{p}^{\prime}(x)=|x|^{p-1} \operatorname{sgn}(x)$ which is strictly monotone increasing. Hence $\lim _{x \uparrow \infty} \sigma_{+}^{\phi}(x)=\lim _{x \uparrow \infty} \phi_{p}^{\prime}(x)=\infty$, and $\lim _{x \downarrow-\infty} \sigma_{-}^{\phi}(x)=\lim _{x \downarrow-\infty} \phi_{p}^{\prime}(x)=-\infty$. Thus, $\phi_{p}$ is strictly convex, and $\phi_{p}^{*}$ will be defined on all of $\mathbb{R}$. To compute it, note that for all Then for all $y \in \mathbb{R}$, the function $x \mapsto x y-\phi(x)$ is continuously differentiable, and its derivative is $y-|x|^{p-1} \operatorname{sgn}(x)$, and hence the unique maximum occurs where $x$ has the same sign as $y$, and $|x|=|y|^{1 /(p-1)}$. Evaluating $x y-\phi_{p}(x)$ as this $x$, we find

$$
\phi_{p}^{*}(y)=|y|^{1+1 /(p-1)}-\frac{1}{p}|y|^{p /(p-1)}=\frac{p-1}{p}|y|^{p /(p-1)} .
$$

Hence if we define $q=p /(p-1)$, we have that

$$
\begin{equation*}
\frac{1}{p}+\frac{1}{q}=1 \quad \text { and } \quad \phi_{p}^{*}=\phi_{q} \tag{1.21}
\end{equation*}
$$

Since the relation (1.21) is symmetric in $p$ and $q,\left(\phi_{p}^{*}\right)^{*}=\phi_{p}$. We shall soon see that this is no coincidence.
1.11 EXAMPLE. Now consider the two limiting functions $\phi_{1}(x)=\lim _{p \downarrow 1} \phi_{p}(x)=|x|$, and $\phi_{\infty}(x)=\lim _{p \uparrow 1} \phi_{p}(x)$, so that

$$
\phi_{\infty}(x):= \begin{cases}0 & x \in[-1,1]  \tag{1.22}\\ \infty & x \notin[-1,1] .\end{cases}
$$

Both $\phi_{1}$ and $\phi_{\infty}$ are lower-semicontinuous convex functions. (In fact, $\phi_{1}$ is even continuous.) Fix $y \in \mathbb{R}$, and note that $y x-\phi(x)=(y \operatorname{sgn}(x)-1)|x| \leq(|y|-1)|x|$, with equality when $\operatorname{sgn}(x)=\operatorname{sgn}(y)$. It follows that $\sup _{x \in \mathbb{R}}\left\{y x-\phi_{1}(x)\right\}=\phi_{\infty}(y)$.

By the Fenchel-Moreau Theorem, it follows that $\phi_{\infty}^{*}=\phi_{1}$, but this is also easy to check from the definition: Fix $x \in \mathbb{R}$, and note that $x y-\phi_{\infty}(y)=x y$ for $|y| \leq 1$ and $x y-\phi_{\infty}(y)=-\infty$ for $|y|>1$. Hence

$$
\phi_{\infty}^{*}(x)=\sup \{x y:|y| \leq 1\}=|x| .
$$

1.12 DEFINITION (Subgradient). Let $\phi$ be a convex function on $\mathbb{R}$ that is finite on an interval $C$. For $x \in C$, the subgradient of $\phi$ at $x, \partial \phi(x)$, is the set of numbers $s$ such that

$$
\phi(y) \geq \phi(x)+s(y-x)
$$

for all $y \in \mathbb{R}$. In other words, $s \in \partial \phi(x)$ if and only if the line with slope $s$ that passes through $(x, \phi(x))$ is a supporting line for $\phi$. If $\phi(x)=\infty$, we define $\partial \phi(x)=\emptyset$. For $A \subset \mathbb{R}$, define

$$
\begin{equation*}
\partial \phi(A)=\bigcup_{x \in A} \partial \phi(x) \tag{1.23}
\end{equation*}
$$

1.13 LEMMA. Let $\phi$ be a convex function on $\mathbb{R}$ that is finite on in a interval $C$. Then for all $x \in C^{o}, \partial \phi(x)=\left[\sigma_{-}^{\phi}(x), \sigma_{+}^{\phi}(x)\right]$. In particular, for $x \in C^{o}, \partial \phi(x) \neq \emptyset$.

Proof. This follows immediately from (1.10).
1.14 THEOREM (Young's Inequality). Let $\phi$ be a proper convex function, and let $\phi^{*}$ be the Legendre transform of $\phi$. Then for all $x, y \in \mathbb{R}$,

$$
\begin{equation*}
x y \leq \phi(x)+\phi^{*}(y) . \tag{1.24}
\end{equation*}
$$

Moreover, there is equality in (1.24) if and only if $y \in \partial \phi(x)$, and in this case, $x \in \partial \phi^{*}(y)$.
Proof. The inequality (1.24) is an immediate consequence of the definition (1.20). Suppose for some $x_{0}, y_{0} \in \mathbb{R}, x_{0} y_{0}=\phi\left(x_{0}\right)+\phi^{*}\left(y_{0}\right)$, so that necessarily $\phi^{*}\left(y_{0}\right)<\infty$. Then by (1.24), for all $x \in \mathbb{R}$,

$$
\phi(x)+\phi^{*}\left(y_{0}\right)-x y_{0} \geq \phi\left(x_{0}\right)+\phi^{*}\left(y_{0}\right)-x_{0} y_{0},
$$

Therefore, $\phi(x) \geq \phi\left(x_{0}\right)+y_{0}\left(x-x_{0}\right)$, which shows that $y_{0} \in \partial \phi\left(x_{0}\right)$. Conversely, if $y_{0} \in \partial \phi\left(x_{0}\right)$, then by definition, $\phi(x) \geq \phi\left(x_{0}\right)+y_{0}\left(x-x_{0}\right)$ for all $x$, and hence

$$
x_{0} y_{0} \geq \phi\left(x_{0}\right)+y_{0} x-\phi(x) .
$$

Taking the supremum over $x$, we find $x_{0} y_{0} \geq \phi\left(x_{0}\right)+\phi^{*}\left(y_{0}\right)$. Together with (1.24), this proves that $x_{0} y_{0}=\phi\left(x_{0}\right)+\phi^{*}\left(y_{0}\right)$.

Likewise, when $x_{0} y_{0}=\phi\left(x_{0}\right)+\phi^{*}\left(y_{0}\right)$, for all $y \in \mathbb{R}$,

$$
\phi\left(x_{0}\right)+\phi^{*}(y)-x_{0} y \geq \phi\left(x_{0}\right)+\phi^{*}\left(y_{0}\right)-x_{0} y_{0},
$$

and $\phi\left(x_{0}\right)<\infty$, so that $\phi^{*}(y) \geq \phi^{*}\left(y_{0}\right)+x_{0}\left(y-y_{0}\right)$. Therefore, $x_{0} \in \partial \phi^{*}\left(y_{0}\right)$.
Suppose that $\phi$ and $\phi^{*}$ are both continuously differentiable. For instance, this is the case when $\phi(x)=p^{-1}|x|^{p}, p \in(1, \infty)$, so that $\phi^{*}(y)=q^{-1}|y|^{q}, q=p /(p-1)$. Then for all $x, y, \partial \phi(x)=\left\{\phi^{\prime}(x)\right\}$ and $\partial \phi^{*}(y)=\left\{\left(\phi^{*}\right)^{\prime}(y)\right\}$. Then the statement about cases of equality in Young's inequality says that

$$
\left[\phi^{*}\right]^{\prime}\left(\phi^{\prime}(x)\right)=x \quad \text { and } \quad \phi^{\prime}\left[\left(\phi^{*}\right)^{\prime}(y)\right]=y .
$$

In other words, the functions $\phi^{\prime}$ and $\left(\phi^{*}\right)^{\prime}$ are inverse to one another. For the dual pair in Example 1.10 , this can be checked by simple computations..
1.15 THEOREM (Fenchel-Moreau Theorem). Let $\phi$ be a proper, lower-semicontinuous function on $\mathbb{R}$ and let $\phi^{*}$ be its Legendre transform. Let $\phi^{* *}$ be the Legendre transform of $\phi^{*}$. Then

$$
\begin{equation*}
\phi^{* *}=\phi . \tag{1.25}
\end{equation*}
$$

Proof. Let $(a, b)$ be the interior of the interval on which $\phi$ is finite. Let $(c, d)$ be be given by (1.16). Let $x \in(a, b)$ and $y \in \partial \phi(x)$. Then $\phi(x)+\phi^{*}(y)=x y$, and $x \in \partial \phi^{*}(y)$. Then by Young's inequality applied to the pair $\phi^{*}$ and $\phi^{* *}, x y=\phi^{*}(y)+\phi^{* *}(x)$, and

$$
\phi(x)+\phi^{*}(y)=x y=\phi^{*}(y)+\phi^{* *}(x) .
$$

This proves that at all points of $(a, b), \phi(x)=\phi^{*}(x)$. Since both functions are lower-semicontinuous, they both agree on $[a, b]$.

## $2 L^{p}$ norms and their close relatives

### 2.1 The $L^{p}$ norm and Hölder's Inequality

2.1 DEFINITION $\left(L^{p}(\Omega, \mathcal{M}, \mu)\right)$. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space. For $1 \leq p<\infty$, $L^{p}(\Omega, \mathcal{M}, \mu)$ consists of the equivalence classes, identified under equivalence almost everywhere, of measurable functions $f$ such that $|f|^{p}$ is integrable. The $L^{p}$ norm is the function $f \mapsto\|f\|_{p}$ where

$$
\begin{equation*}
\|f\|_{p}:=\left(\int_{\Omega}|f|^{p} \mathrm{~d} \mu\right)^{1 / p} \tag{2.1}
\end{equation*}
$$

As we have already observed, the function $t \mapsto t^{p}$ is convex on $[0, \infty)$, and hence for $f, g \in$ $L^{p}(\Omega, \mathcal{M}, \mu), x \in \Omega$,

$$
\left|\frac{|f(x)|+|g(x)|}{2}\right|^{p} \leq \frac{1}{2}|f(x)|^{p}+\frac{1}{2}|g(x)|^{p} .
$$

This shows that $|f+g|^{p}$ is integrable whenever $|f|^{p}$ and $|g|^{p}$ are, and thus $L^{p}(\Omega, \mathcal{M}, \mu)$ is closed under vector addition. It is also evidently closed under scalar multiplication, and thus is a vector space over $\mathbb{C}$. The next inequality will allow us to show that the $L^{p}$ norm is, in fact, a norm on this vector space.
2.2 THEOREM (Hölder's Inequality). Let $(\Omega, \mathcal{M}, \mu)$ be a measure space. Let $1<p, q<\infty$, with $\frac{1}{p}+\frac{1}{q}=1$. Let $f$ and $g$ be functions on $(\Omega, \mathcal{M}, \mu)$ such that $|f|^{q}$ and $|g|^{p}$ are integrable. Then $f g$ is integrable, and

$$
\begin{equation*}
\int_{\Omega}|f g| \mathrm{d} \mu \leq\|f\|_{q}\|g\|_{p} \tag{2.2}
\end{equation*}
$$

There is equality in (2.2) if and only if for almost every $x$,

$$
\begin{equation*}
\|g\|_{p}^{p}|f(x)|^{q}=\|f\|_{q}^{q}|g(x)|^{p} . \tag{2.3}
\end{equation*}
$$

Proof. If either $\int_{\Omega}|f|^{q} \mathrm{~d} \mu=0$ or $\int_{\Omega}|g|^{p} \mathrm{~d} \mu=0$, then (2.2) is true for trivial reasons. Therefore, suppose that both integrals are strictly positive.

Apply Young's inequality with the dual pair $\phi_{p}$ and $\phi_{q}=\phi_{p}^{*}$ from Example 1.10. By (1.21), for any $a>0$, and all $x \in \Omega$,

$$
\begin{equation*}
|f(x)||g(x)|=(a|f(x)|)\left(\frac{1}{a}|g(x)|\right) \leq a^{q} \frac{1}{q}|f(x)|^{q}+a^{-p} \frac{1}{p}|g(x)|^{p} \tag{2.4}
\end{equation*}
$$

with equality only in case

$$
\begin{equation*}
a|f(x)|=\phi_{p}^{\prime}\left(a^{-1}|g(x)|\right)=a^{1-p}|g(x)|^{p-1} . \tag{2.5}
\end{equation*}
$$

Integrating both sides of (2.4),

$$
\begin{equation*}
\int_{\Omega}|f g| \mathrm{d} \mu \leq a^{q}\left(\frac{1}{q} \int_{\Omega}|f|^{q} \mathrm{~d} \mu\right)+a^{-p}\left(\frac{1}{p} \int_{\Omega}|g|^{p} \mathrm{~d} \mu\right)=a^{q} \frac{1}{q}\|f\|_{q}^{q}+a^{-p} \frac{1}{p}\|g\|_{p}^{p} . \tag{2.6}
\end{equation*}
$$

We now choose the value of $a$ so as to make the right hand side as small as possible. A simple calculus exercise shows that the best choice is $a=\|f\|_{q}^{-1 / p}\|g\|_{p}^{1 / q}$. With this choice of $a$, (2.6) becomes (2.2), and (2.5) becomes $\|g\|_{p}^{p / q}|f(x)|=\|f\|_{q}|g(x)|^{p-1}$, and raising both sides to the $q$ th power, we obtain (2.3).

Consider any non-zero $f \in L^{p}(\Omega, \mathcal{M}, \mu), 1<p<\infty$, and let $q=p /(p-1)$. By Hölder's inequality, for all $h \in L^{q}(\Omega, \mathcal{M}, \mu)$ with $\|h\|_{q}=1$,

$$
\Re\left(\int_{\Omega} \bar{h} f \mathrm{~d} \mu\right) \leq \int_{\Omega}|h||f| \mathrm{d} \mu \leq\|f\|_{p}
$$

and there is equality in the second inequality if and only if $\|f\|_{p}^{p}|h(x)|^{q}=\|h\|_{q}^{q}|f(x)|^{p}=|f(x)|^{p}$ for almost every $x$. In this case, $|h(x)|=\|f\|_{p}^{1-p}|f(x)|^{p-1}$. There will be equality in the first inequality if and only if $\Re(\overline{h(x)} f(x))=|h(x)||f(x)|$ almost everywhere. This forces that

$$
g(x):=\|f\|_{p}^{1-p}|f(x)|^{p-1} \operatorname{sgn}(f(x)),
$$

almost everywhere. (The signum function, $z \mapsto \operatorname{sgn}(z)$ on $\mathbb{C}$ is defined by $z \in \mathbb{C}$, define $\operatorname{sgn}(z)=$ $z /|z|$ for $z \neq 0$, and $\operatorname{sgn}(0)=0$.)
2.3 DEFINITION. For $1<p<\infty, q=p /(p-1)$ define the function $D_{p}$ mapping $L^{p}(\Omega, \mathcal{M}, \mu) \backslash\{0\}$ to the unit sphere in $L^{q}(\Omega, \mathcal{M}, \mu)$ by

$$
\begin{equation*}
D_{p}(f)=\|f\|_{p}^{1-p}|f|^{p-1} \operatorname{sgn}(f) . \tag{2.7}
\end{equation*}
$$

$f \mapsto D_{p}(f)$ is called the gradient map for reasons that will become clear.
Altogether, we have proved:
2.4 THEOREM. Let $1<p<\infty, q=p /(p-1)$. For all $f \in L^{p}(\Omega, \mathcal{M}, \mu)$,

$$
\begin{equation*}
\|f\|_{p}=\sup \left\{\Re\left(\int_{\Omega} \bar{h} f \mathrm{~d} \mu\right): h \in L^{p}(\Omega, \mathcal{M}, \mu),\|h\|_{q}=1\right\} \tag{2.8}
\end{equation*}
$$

Moreover, the supremum in (2.8) is a maximum, and when $f \neq 0$, the unique maximizer is $h=$ $D_{p}(f)$.
2.5 THEOREM (Minkowski's Inequality). Let $1<p<\infty, q=p /(p-1)$. For all $f, g \in$ $L^{p}(\Omega, \mathcal{M}, \mu)$,

$$
\begin{equation*}
\|f+g\|_{p} \leq\|f\|_{p}+\|g\|_{p} \tag{2.9}
\end{equation*}
$$

and there is equality if and only if either $f=0$, or else $g$ is a non-negative multiple of $f$.
Proof. We may suppose that neither $f=0$ nor $g=0$. By Theorem 2.4,

$$
\|f+g\|_{p}=\int_{\Omega} \overline{D_{p}(f+g)}(f+g) \mathrm{d} \mu=\int_{\Omega} \overline{D_{p}(f+g)} f \mathrm{~d} \mu+\int_{\Omega} \overline{D_{p}(f+g)} g \mathrm{~d} \mu \leq\|f\|_{p}+\|g\|_{p} .
$$

There is equality if and only if $D_{p}(f+g)=D_{p}(f)=D_{p}(g)$. The second equality forces $\operatorname{sgn}(f)=$ $\operatorname{sgn}(g)$ and $|f|^{p-1}=a|g|^{p-1}$ for some $a>0$, and hence $|f|=a^{1 /(p-1)}|g|$.

It is easy to prove that for $1<p<\infty, L^{p}(\Omega, \mathcal{M}, \mu)$ is a complete metric space, and hence a Banach space. In fact the proof is very much like the one we have already given for completeness of $L^{2}(\Omega, \mathcal{M}, \mu)$, and it extends to a much wider class of norms that we now introduce.

### 2.2 Orlicz norms

Throughout this section, let $(\Omega, \mathcal{M}, \mu)$ be a given measure space, and define $L^{0}(\Omega, \mathcal{M}, \mu)$ to be the vector space of measurable complex valued functions $f$ on $\Omega$, identified under almost-everywhere equivalence.
2.6 DEFINITION. An Orlicz function $\phi$ is a convex lower-semicontinuous function on $\mathbb{R}$ such that is symmetric $(\phi(-x)=\phi(x)$ for all $x$ ), with $\phi(0)=0, \phi(1)<\infty$ with $1 \in \partial \phi(1)$, and $\lim _{x \uparrow \infty} \phi(x)=\infty$. For any Orlicz function $\phi$, define

$$
B_{\phi}=\left\{f \in L^{0}(\Omega, \mathcal{M}, \mu): \int_{\Omega} \phi(|f|) \mathrm{d} \mu \leq \phi(1)\right\}
$$

Define $L_{\phi}$ to be the subspace of $L^{0}(\Omega, \mathcal{M}, \mu)$ spanned by $B_{\phi}$.
Let $\phi$ be an Orlicz function. The $\phi$ is non-negative. Indeed since $\phi$ is convex and even, $\phi(0) \leq \frac{1}{2}(\phi(x)+\phi(-x))=\phi(x)$, Hence 0 is a minimizer for any symmetric convex function, and since $\phi$ is a Orlicz function, $\phi(0)=0$. Evidently the $x$-axis is a supporting line for the graph of $\phi$, and so $0 \in \partial \phi(0)$.

Let $\phi^{*}$ be the Legendre transform of $\phi$. We claim that $\phi^{*}$ is also an Orlicz function. First, $\phi^{*}$ is evidently proper, symmetric and lower semicontinuous. Since $0 \in \partial \phi(0)$, the conditions for equality in Young's inequality give us $0=\phi(0)+\phi^{*}(0)$, so that $\phi^{*}(0)=0$. Also, since $1 \in \partial \phi(1)$, the conditions for equality in Young's inequality give $1=\phi(1)+\phi^{*}(1)$, so that $\phi^{*}(1)$ is not only finite, but $\phi^{*}(1) \in[0,1]$, and $1 \in \partial \phi^{*}(1)$. Finally, since $1 \in \partial \phi^{*}(1), \phi^{*}(y) \geq \phi^{*}(1)+(y-1)$, and hence $\lim _{y \uparrow \infty} \phi^{*}(y)=\infty$. Summarizing, we have:
2.7 LEMMA. Let $\phi$ be an Orlicz function, and let $\phi^{*}$ be its Legendre transform. Then $\phi^{*}$ is also an Orlicz function and

$$
\begin{equation*}
\phi(1)+\phi^{*}(1)=1 . \tag{2.10}
\end{equation*}
$$

2.8 EXAMPLE. Let $1 \leq p<\infty$, and let $\phi_{p}(x)=p^{-1}|x|^{p}$, which is easily seen to be an Orlicz function since it is continuously differentiable at $x=1$, and $\phi_{p}^{\prime}(1)=1$, showing that $\partial \phi_{p}(1)=\{1\}$. A measurable function $f$ belongs to $B_{\phi_{p}}$ if and only if

$$
\frac{1}{p} \int_{\Omega}|f|^{p} \mathrm{~d} \mu \leq \frac{1}{p}
$$

and hence $f \in B_{\phi_{p}}$ is and only if $\|f\|_{p} \leq 1$. Thus, $B_{\phi_{p}}$ is precisely the closed unit ball in $L^{p}(\Omega, \mathcal{M}, \mu)$. Therefore, $L_{\phi_{p}}=L^{p}(\Omega, \mathcal{M}, \mu)$. As we have seen in Example 1.11, $\phi_{1}^{*}=\phi_{\infty}$ where $\phi_{\infty}$ is defined in (1.22). Then $\phi_{\infty}(1)=0$, and hence

$$
\int_{\Omega} \phi_{\infty}(|f|) \mathrm{d} \mu \leq \phi_{\infty}(1)
$$

if and only if $\phi_{\infty}(|f|)=0$ almost everywhere, and this is the case if an only if the essential supremum $\|f\|_{\infty}$ of $f$ belongs to $[0,1]$. The subspace of $L^{0}(\Omega, \mathcal{M}, \mu)$ of functions with $\|f\|_{\infty}<\infty$ defines the space $L^{\infty}(\Omega, \mathcal{M}, \mu)$, and hence $L^{\infty}(\Omega, \mathcal{M}, \mu)=L_{\phi_{\infty}}$. In particular, each $L^{p}$ space, $1 \leq p \leq \infty$ is an Orlicz space. We now introduce Orlicz norms which will turn out to be the $L^{p}$ norms on the $L^{p}$ spaces.
2.9 LEMMA. Let $\phi$ be any Orlicz function function. Then $B_{\phi}$ is a balanced convex subset of $L^{0}(\Omega, \mathcal{M}, \mu)$, and $\cap_{r>0} r B_{\phi}=\{0\}$.

Since $B_{\phi}$ is absorbing in $L_{\phi}$ by definition, the function $f \mapsto\|f\|_{\phi}$ in $L_{\phi}$ given by

$$
\begin{equation*}
\|f\|_{\phi}=\inf \left\{r>0: f \in r B_{\phi}\right\} \tag{2.11}
\end{equation*}
$$

is a norm on $L_{\phi}$ called the Luxembourg norm on $L_{\phi}$.
2.10 EXAMPLE. Continuing with Example 2.8, it is clear that $\|\cdot\|_{\phi_{p}}=\|\cdot\|_{p}$.
2.11 LEMMA. Let $\phi$ be a Young's function. For all $f \in L_{\phi}$,

$$
\begin{equation*}
\int_{\Omega} \phi\left(\frac{f}{\|f\|_{\phi}}\right) \mathrm{d} \mu \leq \phi(1) . \tag{2.12}
\end{equation*}
$$

and for all non-negative measurable functions $f$ and $g$,

$$
\begin{equation*}
f \leq g \quad \Rightarrow \quad\|f\|_{\phi} \leq\|g\|_{\phi}, \tag{2.13}
\end{equation*}
$$

Proof. The definition (2.11) is equivalent to

$$
\begin{equation*}
\|f\|_{\phi}=\inf \left\{t>0: \int_{\Omega} \phi\left(\frac{|f|}{t}\right) \mathrm{d} \mu \leq \phi(1)\right\} \tag{2.14}
\end{equation*}
$$

Hence for all $n \in \mathbb{N}, \int_{\Omega} \phi\left(\frac{|f|}{\|f\|_{\phi}+1 / n}\right) \mathrm{d} \mu \leq \phi(1)$, and then (2.12) follows from Fatou's Lemma and the lower-semicontinuity of $\phi$.

Next, since $\phi$ is monotone on $[0, \infty)$, for all $t>0, \int_{\Omega} \phi\left(\frac{|f|}{t}\right) \mathrm{d} \mu \leq \int_{\Omega} \phi\left(\frac{|g|}{t}\right) \mathrm{d} \mu$. From this and (2.14), we obtain (2.13).

The next theorem shows that when $\phi$ is an Orlicz function, $\left(L_{\phi},\|\cdot\|_{\phi}\right)$ is complete, and hence is a Banach space. By what has been shown in the examples above, this extends the Riesz-Fischer Theorem from $L^{2}$ to $L^{p}$ for all $1 \leq p \leq \infty$.
2.12 THEOREM (Completeness of Orlicz spaces). Let $\phi$ be a Young's function, and let ( $\Omega, \mathcal{M}, \mu$ ) be a measures space, Let $\left(L_{\phi},\|\cdot\|_{\phi}\right)$ be the Orlicz space associated to $\phi$ and $(\Omega, \mathcal{M}, \mu)$. Then $\left(L_{\phi},\|\cdot\|_{\phi}\right)$ is a Banach space, and from every Cauchy sequence $\left\{f_{n}\right\}_{n \in \mathbb{N}}$ in $L_{\phi}$, one may extract a subsequence that converges almost everywhere.

Proof. For each $k$, pick $n_{k}$ so that $j, \ell \geq n_{k} \Rightarrow\left\|f_{j}-f_{\ell}\right\|_{\phi} \leq 2^{-k}$. Without loss of generality, we may suppose that $n_{k+1}>n_{k}$ for each $k$. For $N \in \mathbb{N}$, define the function $F_{N}(x)$ by

$$
F_{N}(x)=\left|f_{n_{1}}(x)\right|+\sum_{k=1}^{N}\left|f_{n_{k+1}}(x)-f_{n_{k}}(x)\right| .
$$

By Minkowski's inequality,

$$
\left\|F_{N}\right\|_{\phi} \leq\left\|f_{n_{1}}\right\|_{\phi}+\sum_{k=1}^{N}\left\|f_{n_{k+1}}-f_{n_{k}}\right\|_{\phi} \leq\left\|f_{n_{1}}\right\|_{\phi}+\sum_{k=1}^{N} 2^{-k} \leq\left\|f_{n_{1}}\right\|_{\phi}+1
$$

Let $c=\left\|f_{n_{1}}\right\|_{\phi}+1$. By (2.14), $\int_{\Omega} \phi\left(\frac{F_{N}(x)}{c}\right) \mathrm{d} x \leq \phi(1)$. Define $F(x)=\lim _{N \rightarrow \infty} F_{N}(x)$. Since $\phi$ is lower-semicontinuous, $\phi\left(\frac{F(x)}{c}\right) \leq \liminf _{N \rightarrow \infty} \phi\left(\frac{F_{N}(x)}{c}\right)$, and then by Fatou's Lemma,

$$
\int_{\Omega} \phi\left(\frac{F(x)}{c}\right) \mathrm{d} x \leq \liminf _{N \rightarrow \infty} \int_{\Omega} \phi\left(\frac{F_{N}(x)}{c}\right) \mathrm{d} x \leq \phi(1) .
$$

Hence $\|F\|_{\phi} \leq c$, and since $\lim _{t \uparrow \infty} \phi(t)=\infty, F$ is finite almost everywhere.
Next define $f$ by

$$
f(x):=\lim _{k \rightarrow \infty} f_{n_{k}}=f_{n_{1}}+\sum_{k=1}^{\infty}\left(f_{n_{k+1}}-f_{n_{k}}\right)
$$

where the sum on the right converges absolutely wherever $F$ is finite; that is, almost everywhere. Since $|f| \leq F, f \in L_{\phi}$ by Lemma 2.11.

By the definition of the subsequence $\left\{f_{n_{j}}\right\}_{j \in \mathbb{N}}$, for all $k \in \mathbb{N}, \int_{\Omega} \phi\left(\frac{\left|f_{n_{\ell}}-f_{n_{k}}\right|}{2^{-k}}\right) \mathrm{d} \mu \leq \phi(1)$. Again since $\phi$ is lower-semicontinuous, $\phi\left(\frac{\left|f-f_{n_{k}}\right|}{2^{-k}}\right) \leq \liminf _{\ell \rightarrow \infty} \phi\left(\frac{\left|f_{n_{\ell}}-f_{n_{k}}\right|}{2^{-k}}\right)$, and by Fatou's Lemma once more, $\left\|f-f_{n_{k}}\right\|_{\phi} \leq 2^{-k}$.

$$
\lim _{k \rightarrow \infty}\left\|f_{n_{k}}-f\right\|_{\phi}=0
$$

Since the sequence is Cauchy, we have this convergence along the whole sequence as well.

### 2.3 Duality in Orlicz spaces

Let $\phi$ be an Orlicz function, and let $\phi^{*}$ be its Legendre transform. Any such pair of Orlicz functions is called a dual pair of Orlicz functions. Associated to such a pair of Orlicz functions is the pair of normed spaces $\left(L_{\phi},\|\cdot\|_{\phi}\right)$ and $\left(L_{\phi^{*}},\|\cdot\|_{\phi^{*}}\right)$.

What is the relation between $L_{\phi}^{*}$ and $L_{\phi^{*}}$ ? The first step towards answering this question provided by the generalized Hölder inequality:
2.13 THEOREM (Generalized Hölder's Inequality). Let $\phi, \phi^{*}$ be any dual pair of Orlicz functions. Then for all functions $f \in L_{\phi}$ and $g \in L_{\phi^{*}}$, fg is measurable and

$$
\begin{equation*}
\int_{\Omega}|f g| \mathrm{d} \mu \leq\|f\|_{\phi}\|g\|_{\phi^{*}} . \tag{2.15}
\end{equation*}
$$

There is equality in (2.15) if and only if

$$
\begin{gather*}
\int_{\Omega} \phi\left(\frac{|f|}{\|f\|_{\phi}}\right) \mathrm{d} \mu=\phi(1) \quad \text { and } \quad \int_{\Omega} \phi^{*}\left(\frac{|g|}{\|g\|_{\phi^{*}}}\right) \mathrm{d} \mu=\phi^{*}(1),  \tag{2.16}\\
\frac{|g(x)|}{\|g\|_{\phi^{*}}} \in \partial \phi\left(\frac{|f(x)|}{\|f\|_{\phi}}\right) \tag{2.17}
\end{gather*}
$$

for almost every $x$.
The following lemma will be useful here and elsewhere:
2.14 LEMMA. Let $\phi$ be a Young's function. Then for all $a, b>0$, sis the function $\phi_{a, b}$ defined by $\phi_{a, b}(s)=b \phi(a s)$. Then

$$
\phi_{a, b}^{*}(t)=b \phi^{*}\left(\frac{t}{a b}\right) .
$$

Proof.

$$
\phi_{a, b}^{*}(t)=\sup _{s \geq 0}\{t s-b \phi(a s)\}=\sup _{s \geq 0}\left\{b \frac{t}{a b} a s-b \phi(a s)\right\}=b \phi^{*}\left(\frac{t}{a b}\right) .
$$

Proof of Theorem 2.13. By Young's inequality applied to $\phi_{a, b}$ with $a, b>0$, we have from Lemma 2.14 that

$$
s t \leq b \phi(a s)+b \phi^{*}\left(\frac{t}{a b}\right)
$$

for all $s, t \geq 0$, and there is equality only if $t \in \partial \phi_{a, b}(s)$ Then for any $f \in L_{\phi}$ and $g \in L_{\phi^{*}}$,

$$
\int_{\Omega}|f g| \mathrm{d} \mu \leq b \int_{\Omega} \phi(|a f|) \mathrm{d} \mu+b \int_{\Omega} \phi^{*}\left(\frac{|g|}{a b}\right) \mathrm{d} \mu
$$

By Lemma 2.11, $\int_{\Omega} \phi\left(\frac{|f|}{\|f\|_{\phi}}\right) \mathrm{d} \mu \leq \phi(1)$ and $\int_{\Omega} \phi^{*}\left(\frac{|g|}{\|g\|_{\phi^{*}}}\right) \mathrm{d} \mu \leq \phi^{*}(1)$. Choosing $a=\frac{1}{\|f\|_{\phi}}$ and $b=\|f\|_{\phi}\|g\|_{\phi^{*}}$, we obtain

$$
\int_{\Omega}|f g| \mathrm{d} \mu \leq\|f\|_{\phi}\|g\|_{\phi^{*}}\left(\phi(1)+\phi^{*}(1)\right)=\|f\|_{\phi}\|g\|_{\phi^{*}},
$$

where we have used the fact that $1=\phi(1)+\phi^{*}(1)$ since $1 \in \partial \phi(1)$. There is equality if and only if (2.16) is valid and equality holds in the application of Young's inequality at almost every $x$, which means that $|g(x)| \in \partial \phi_{a, b}(|f(x)|)$. Then since $s \in \partial \phi_{a, b}(t)$ if and only if $\frac{s}{a b} \in \partial \phi(a t)$, there is equality in (2.15) if and only if (2.17) is valid.

Suppose that $\phi$ and $\phi^{*}$ are a dual pair of Orlicz functions. For all $g \in L_{\phi^{*}}$, define the linear functional $L_{g}$ on $L_{\phi}$ by

$$
\begin{equation*}
L_{g}(f)=\int_{\Omega} \bar{g} f \mathrm{~d} \mu \tag{2.18}
\end{equation*}
$$

which is well-defined by Theorem 2.13, and in fact, by Theorem 2.13,

$$
\left|L_{g}(f)\right| \leq\|f\|_{\phi}\|g\|_{\phi^{*}}
$$

Therefore, $L_{g} \in L_{\phi}^{*}$, and

$$
\left\|L_{g}\right\|_{L_{\phi}^{*}} \leq\|g\|_{\phi^{*}} .
$$

Thus, the mapping $g \mapsto L_{g}$ is a linear contraction from $L_{\phi^{*}}$ into $L_{\phi}^{*}$. It is not hard to show that is is injective, at least when $\mu$ has the property that every measurable set with positive measure contains a measurable set with finite positive measure. One might hope that this map would also be surjective under these same mild conditions. In that case, by the Open Mapping Theorem, it would be a Banach space isomorphism, and thus we would identify $L_{\phi}^{*}$ with $L_{\phi^{*}}$. But then the same argument would identify $L_{\phi^{*}}^{*}$ with $L_{\phi^{* *}}=L_{\phi}$, and we would have that the natural injection of $L_{\phi}$ into $L_{\phi}^{* *}$ would be a Banach space isomorphism, i.e., that $L_{\phi}$ would be reflexive. This is not true in general: It fails for $\phi_{1}$ and $\phi_{\infty}$, but it is the case for $\phi_{p}, 1<p<\infty$. The key to this and other issues lies in the notion of uniform convexity, to which we now turn.

## 3 Uniform convexity and uniform smoothness

### 3.1 Uniform convexity

The unit ball of any normed vector space $V$ is convex, though it need not be strictly convex, which would mean that for all unit vectors $u$ and $v$ in $V$,

$$
\begin{equation*}
\|u-v\|>0 \rightarrow\|(u+v) / 2\|<1 . \tag{3.1}
\end{equation*}
$$

Indeed, strict convexity fails for $L^{1}(M, \mathcal{M}, \mu)$ and $L^{\infty}(M, \mathcal{M}, \mu)$, even for a two-point measure space.

To see this in $L^{1}(M, \mathcal{M}, \mu)$, take any two non-negative unit vectors $u(x)$ and $v(x)$. Then of course

$$
\left\|\frac{u+v}{2}\right\|=\frac{1}{2} \int_{M}(u(x)+v(x)) \mathrm{d} \mu=\frac{1}{2} \int_{M} u(x) \mathrm{d} \mu+\frac{1}{2} \int_{M} v(x) \mathrm{d} \mu=1 .
$$

To see this in $L^{\infty}(M, \mathcal{M}, \mu)$, take any two $u(x)$ and $v(x)$ to be the indicator functions of two measurable sets with different, non-zero measure. Then $u(x)$ and $v(x)$ are both unit vectors in $L^{\infty}(M, \mathcal{M}, \mu)$, as is their average, $(u+v) / 2$.

In some normed spaces however, a uniform version of strict convexity holds, and this has significant consequences.
3.1 DEFINITION (Uniform convexity). Let $V$ be a vector space with norm $\|\cdot\|$. The modulus of convexity of $V$ is the function $\delta_{V}$ defined by

$$
\begin{equation*}
\delta_{V}(\epsilon)=\inf \left\{1-\left\|\frac{v+w}{2}\right\|:\|v-w\|<2 \epsilon\right\} \tag{3.2}
\end{equation*}
$$

for $0 \leq \epsilon \leq 1$. We say that $V$ is uniformly convex in case $\delta_{V}(\epsilon)>0$ for all $0<\epsilon<1$.
If there is no ambiguity as to which space $V$ is under consideration, we just write $\delta(\epsilon)$ in place of $\delta_{V}(\epsilon)$. Then, by definition, in case $V$ is uniformly convex, for all $0<\epsilon<1$ there is a $\delta(\epsilon)>0$ so that for all unit vectors $v, w$,

$$
\begin{equation*}
\|v-w\|>2 \epsilon \Rightarrow\left\|\frac{v+w}{2}\right\|<1-\delta(\epsilon), \tag{3.3}
\end{equation*}
$$

which is indeed a uniform version of (3.1). The logically equivalent implication

$$
\begin{equation*}
\left\|\frac{v+w}{2}\right\| \geq 1-\delta(\epsilon) \Rightarrow\|v-w\| \leq 2 \epsilon \tag{3.4}
\end{equation*}
$$

will be used frequently in what follows.
By what we have seen just above, neither $L^{1}(M, \mathcal{M}, \mu)$ nor $L^{\infty}(M, \mathcal{M}, \mu)$ is uniformly convex. It turns out, however, that for $1<p<\infty, L^{p}(M, \mathcal{M}, \mu)$ is uniformly convex. This is easy to show for $L^{2}(M, \mathcal{M}, \mu)$, and we begin with that:

For any $f$ and $g$ in we have the parallelogram identity $\|f-g\|_{2}^{2}+\|f+g\|_{2}^{2}=2\|f\|_{2}^{2}+2\|g\|_{2}^{2}$. Take $f$ and $g$ to be unit vectors. Divide through by 4 to obtain

$$
\left\|\frac{f+g}{2}\right\|_{2}^{2}+\left\|\frac{f-g}{2}\right\|_{2}^{2}=\frac{\|f\|_{2}^{2}+\|g\|_{2}^{2}}{2}=1 .
$$

Therefore, $\left\|\frac{f+g}{2}\right\|_{2}=\sqrt{1-\left\|\frac{f-g}{2}\right\|_{2}^{2}}$. For any number $a$ with $0<a<1, \sqrt{1-a}<1-a / 2$, and hence,

$$
\left\|\frac{f+g}{2}\right\|_{2} \leq 1-\frac{1}{2}\left\|\frac{f-g}{2}\right\|_{2}^{2}
$$

Since $f$ abd $g$ are arbitrary unit vecotors, this gives us the exact modulus of convexity for $L^{2}(M, \mathcal{M}, \mu)$, namely

$$
\begin{equation*}
\delta_{L^{2}}(\epsilon)=1-\sqrt{1-\epsilon^{2}} \geq \frac{1}{2} \epsilon^{2} . \tag{3.5}
\end{equation*}
$$

### 3.2 First applications of uniform convexity

3.2 THEOREM (Convergence of norms plus weak convergence yields strong convergence). Let $V$ be a uniformly convex normed space. Let $\left\{f_{n}\right\}$ be a weakly convergent sequence in $V$ with $1<p<\infty$, and let $f$ be its limit. Then $\left\{f_{n}\right\}$ is strongly convergent if and only if $\|f\|=\lim _{n \rightarrow \infty}\left\|f_{n}\right\|$.
Proof. If $\left\{f_{n}\right\}$ is strongly convergent, it must converge strongly to $f$, and we have already observed that it must be the case that $\|f\|=\lim _{n \rightarrow \infty}\left\|f_{n}\right\|$.

The converse is more subtle, and it is here that uniform convexity comes in. If $f=0$, the strong convergence is obvious, so we may suppose that this is not the case. Then, dividing through by $\|f\|$, we may assume that $\|f\|=1$. Since $\lim _{n \rightarrow \infty}\left\|f_{n}\right\|=\|f\|=1$, we may delete a finite number of terms from the sequence to arrange that $\left\|f_{n}\right\| \neq 0$ so any $n$.

Consider the sequence $\left\{g_{n}\right\}$ where

$$
g_{n}=\frac{f_{n} /\left\|f_{n}\right\|+f}{2} .
$$

Since $\lim _{n \rightarrow \infty}\left\|f_{n}\right\|=\|f\|=1,\left\{g_{n}\right\}$ also converges weakly to $f$. By the weak lower semicontinuity of the norms,

$$
\begin{equation*}
\liminf _{n \rightarrow \infty}\left\|g_{n}\right\| \geq\|f\|=1 \tag{3.6}
\end{equation*}
$$

But by Minkowski's inequality,

$$
\begin{equation*}
1=\frac{\left\|f_{n}\right\| /\left\|f_{n}\right\|+\|f\|}{2} \geq\left\|g_{n}\right\|, \tag{3.7}
\end{equation*}
$$

and then combining (3.6) and (3.7),

$$
\begin{equation*}
\lim _{n \rightarrow \infty}\left\|g_{n}\right\|=1 \tag{3.8}
\end{equation*}
$$

Now use the uniform convexity, and in particular (3.4):

$$
\left\|g_{n}\right\|=\left\|\frac{\left\|f_{n}\right\| /\left\|f_{n}\right\|+\|f\|}{2}\right\| \geq 1-\delta_{V}(\epsilon) \Rightarrow\left\|f_{n} /\right\| f_{n}\|-f\| \leq 2 \epsilon .
$$

This together with (3.8) shows that

$$
\lim _{n \rightarrow \infty}\left\|f_{n} /\right\| f_{n}\|-f\|=0
$$

But

$$
\left\|f_{n}-f\right\|=\left\|\left(f_{n}-f_{n} /\left\|f_{n}\right\|\right)+\left(f_{n} /\left\|f_{n}\right\|-f\right)\right\| \leq \frac{\left|\left\|f_{n}\right\|-1\right|}{\left\|f_{n}\right\|}+\left\|f_{n} /\right\| f_{n}\|-f\|
$$

Hence it follows that $\lim _{n \rightarrow \infty}\left\|f_{n}-f\right\|=0$.

Our next application is very important: It is the generalization of the Projection Lemma to general uniformly convex spaces.
3.3 THEOREM (Projection Lemma for uniformly convex spaces). Let $V$ be a uniformly convex Banach space, and let $K$ be a non-empty, closed convex set in $V$. Then there exists a unique element of minimal norm in $K$. That is, there exists an element $v \in K$ with

$$
\|v\|<\|w\|
$$

for all $w \in K$ with $w \neq v$. Moreover, $\left\{v_{n}\right\}_{n \in \mathbb{N}}$ is any sequence in $K$ such that $\lim _{n \rightarrow \infty}\left\|v_{n}\right\|=\|v\|$, $\lim _{n \rightarrow \infty}\left\|v_{n}-v\right\|=0$.

Proof. Let $D=\inf \{\|w\| \mid w \in K\}$. Let $\left\{v_{n}\right\}_{n \in \mathbb{N}}$ be any sequence in $K$ with $\lim _{n \rightarrow \infty}\left\|v_{n}\right\|=D$. If $D=0, \lim _{n \rightarrow \infty} v_{n}=0$. Since $K$ is closed, this means $0 \in K$, and this is our unique element of minimal norm. Therefore, assume that $D>0$.

Normalize the $v_{n}$ to obtain unit vectors, as needed for the application of uniform convexity. Let $u_{n}=v_{n} /\left\|v_{n}\right\|$. For large $n$, we have that $u_{n} \approx v_{n} / D$ since $\lim _{k \rightarrow \infty}\left\|v_{k}\right\|=D$. Indeed, adding and subtracting,

$$
u_{n}=\frac{1}{D} v_{n}+\frac{D-\left\|v_{n}\right\|}{D\left\|v_{n}\right\|} v_{n} .
$$

Therefore, for any $m$ and $n$,

$$
\begin{aligned}
\frac{1}{D}\left\|\frac{v_{n}+v_{m}}{2}\right\| & =\left\|\frac{u_{n}+u_{m}}{2}-\frac{D-\left\|v_{n}\right\|}{2 D\left\|v_{n}\right\|} v_{n}-\frac{D-\left\|v_{m}\right\|}{2 D\left\|v_{m}\right\|} v_{m}\right\| \\
& \leq\left\|\frac{u_{n}+u_{m}}{2}\right\|+\frac{\left(\left\|v_{n}\right\|-D\right)+\left(\left\|v_{m}\right\|-D\right)}{2}
\end{aligned}
$$

Now by the convexity of $K,\left(v_{n}+v_{m}\right) / 2 \in K$ and hence $\left\|\left(v_{n}+v_{m}\right) / 2\right\| \geq D$. Therefore,

$$
\left\|\frac{u_{n}+u_{m}}{2}\right\| \geq 1-\frac{\left(\left\|v_{n}\right\|-D\right)+\left(\left\|v_{m}\right\|-D\right)}{2}
$$

Then by (3.4), for all $\epsilon>0,\left(\left\|v_{n}\right\|-D\right)+\left(\left\|v_{m}\right\|-D\right) \leq 2 \delta(\epsilon) \Rightarrow\left\|u_{n}-u_{m}\right\| \leq 2 \epsilon$. Therefore, since $\lim _{n \rightarrow \infty}\left\|v_{n}\right\|=D,\left\{u_{n}\right\}_{n \in \mathbb{N}}$ is a Cauchy sequence. Since $V$ is complete, $\left\{u_{n}\right\}_{n \in \mathbb{N}}$ converges in norm to $u \in V$, and this implies that $\left\{v_{n}\right\}_{n \in \mathbb{N}}$ converges in norm to $v:=D u$. Since $K$ is closed, $D u \in K$, and since $\|u\|=1,\|v\|=\|D u\|=D$. This proves the existence of an element $v$ of $K$ with minimal norm, and that $\lim _{n \rightarrow \infty} v_{n}=v$.

To prove the uniqueness, let $\tilde{v}$ also be in $K$ with $\|\tilde{v}\|=D$. Define $v_{n}=n$ for $n$ even and $v_{n}=\tilde{v}$ for $n$ odd. By what we proved above $\left\{v_{n}\right\}_{n \in \mathbb{N}}$ converges, and hence $\tilde{v}=v$.

Recall that for any normed space $V$, and any $v \in V$, there exists an $f \in V^{*}$ with $\|f\|_{*}=1$ and $f(v)=\|v\|$. This is a consequence of the Hahn-Banach Theorem. However, given $f \in V^{*}$, there may or may not be any unit vector $u$ in $V$ such that $f(u)=\|f\|_{*}$, as we have seen in the case of $V=\mathcal{C}([0,1])$ with the uniform norm. If $V$ is uniformly convex, things are much better.
3.4 THEOREM (Uniform Convexity and Unit Normal Vectors). Let $V$ be a uniformly convex Banach space, and let $f$ be any non-zero linear functional in $V^{*}$. Then there is a unique unit vector $v_{f} \in V$ so that

$$
f\left(v_{f}\right)=\|f\|_{*} .
$$

Moreover, the function $f \mapsto v_{f}$ from $V^{*}$ to $V$ is continuous $f \neq 0$, and in fact, for all non zero $f, g \in V^{*}$,

$$
\begin{equation*}
\|f-g\|_{*} \leq 2\|f\|_{*} \delta_{V}(\epsilon) \Rightarrow\left\|v_{f}-v_{g}\right\| \leq 2 \epsilon \tag{3.9}
\end{equation*}
$$

The vector whose existence is asserted by the theorem is called the unit normal vector at $f$ for reasons that will soon be explained.

Proof. Let $K$ be given by

$$
K=\left\{v \in V: f(v)=\|f\|_{*}\right\} .
$$

$K$ is closed, convex and non-empty. By the projection lemma, $K$ contains a unique element $v$ of minimal norm. Note that

$$
\|f\|_{*}=f(v) \leq\|f\|_{*}\|v\|
$$

so $\|v\| \geq 1$.
Now we prove an upper bound on $\|v\|$. For any $\epsilon$ with $0<\epsilon<\|f\|_{*}$, there is a unit vector $w$ with $|f(w)| \geq\|f\|_{*}-\epsilon$. Multiplying $w$ by a complex number of unit magnitude, we can assume that $f(w)=|f(w)|$. Now define $v=\frac{\|f\|_{*}}{f(w)} w$. Then $v \in K$, and since $w$ is a unit vector,

$$
\|v\|=\frac{\|f\|_{*}}{f(w)} \leq \frac{\|f\|_{*}}{\|f\|_{*}-\epsilon} .
$$

Since $v$ is the element of $K$ with minimal norm, we have $1 \leq\|v\| \leq \frac{\|f\|_{*}}{\|f\|_{*}-\epsilon}$ for all $\epsilon$ with $0<\epsilon<$ $\|f\|_{*}$. This means that $\|v\|=1$, and $v_{f}=v$ is the vector we seek.

Since $v_{f}$ is uniquely determined, the function $v \mapsto v_{f}$ is well defined. We now show that it is continuous.

Let $f$ and $g$ in $V^{*}$ be given, and let $v_{f}$ and $v_{g}$ be the corresponding unit vectors in $V$. Then

$$
\begin{align*}
\|f+g\|_{*}\left\|v_{f}+v_{g}\right\| & \geq \mathcal{R}\left((f+g)\left(v_{f}+v_{g}\right)\right) \\
& =\mathcal{R}\left(f\left(v_{f}\right)+f\left(v_{g}\right)+g\left(v_{f}\right)+g\left(v_{g}\right)\right) \\
& =2\left(\|f\|_{*}+\|g\|_{*}\right)+\mathcal{R}\left(f\left(v_{g}\right)+g\left(v_{f}\right)-f\left(v_{f}\right)-g\left(v_{g}\right)\right) \\
& =2\left(\|f\|_{*}+\|g\|_{*}\right)-\mathcal{R}\left((f-g)\left(v_{f}-v_{g}\right)\right) \\
& \geq 2\|f+g\|_{*}-\|f-g\|_{*}\left\|v_{f}-v_{g}\right\| . \tag{3.10}
\end{align*}
$$

Dividing through by $2\|f+g\|_{*}$ we get

$$
\left(\frac{\|f-g\|_{*}}{\|f+g\|_{*}}\right)\left\|\frac{v_{f}-v_{g}}{2}\right\| \geq 1-\left\|\frac{v_{f}+v_{g}}{2}\right\| .
$$

If $\left\|v_{f}-v_{g}\right\|>2 \epsilon, 1-\left\|\left(v_{f}+v_{g}\right) / 2\right\| \geq \delta(\epsilon)$, and since $\left\|\left(v_{f}-v_{g}\right) / 2\right\| \leq 1$ in any case,

$$
\left(\frac{\|f-g\|_{*}}{2\|f\|_{*}-\|f-g\|_{*}}\right) \geq \delta(\epsilon),
$$

or $\|f-g\|_{*} \geq \frac{\delta(\epsilon)}{1+\delta(\epsilon)} 2\|f\|_{*}$. Hence $\|f-g\|_{*}<\frac{\delta(\epsilon)}{1+\delta(\epsilon)} 2\|f\|_{*} \Rightarrow\left\|v_{f}-v_{g}\right\|<\epsilon$. Discarding the $\delta(\epsilon)$ in the denominator, we obtain (3.9), which quantifies the continuity of $f \mapsto v_{f}$ at all $f \neq 0$.

### 3.3 Uniform smoothness

Let $V$ be a normed space with norm $\|\cdot\|$. The derivative gives the "best linear approximation" to a function. We say that a functional $F$ on $V$ is Frechét differentiable at $u \in V$ in case there is a linear functional $\ell_{F, u} \in V^{*}$ so that

$$
F(u+v)-F(u)=\ell_{F, u}(v)+o(\|v\|)
$$

or, in other words, if

$$
\begin{equation*}
\lim _{v \rightarrow 0} \frac{\left|F(u+v)-F(u)-\ell_{F, u}(v)\right|}{\|v\|}=0 \tag{3.11}
\end{equation*}
$$

where the limit is taken in the norm sense.
There is another notion of differentiability, corresponding to the usual directional derivative. A functional $F$ is said to be Gateaux differentiable at $u \in V$ in case for there is a linear functional $\ell_{F, u} \in V^{*}$ so that for each $v \in V$,

$$
F(u+t v)-F(u)=t \ell_{F, u}(v)+o(t)
$$

or, in other words, if

$$
\begin{equation*}
\lim _{v \rightarrow 0} \frac{\left|F(u+t v)-F(u)-t \ell_{F, u}(v)\right|}{t}=0 . \tag{3.12}
\end{equation*}
$$

Clearly, if a functional $F$ is Frechét differentiable, then it is Gateaux differentiable, and the two derivatives coincide. However, there are functionals that are Gateaux differentiable, but not Frechét differentiable.

To check differentiability from the definition, you have to know the derivative $\ell_{F, u}$, which is somewhat inconvenient. There is, however, a necessary condition for differentiability that can be stated solely in terms of $F$ itself. If $F$ is Frechét differentiable at $u$, then for any $v$,

$$
F(u+v)-F(u)=\ell_{F, u}(v)+o(\|v\|)
$$

and

$$
F(u-v)-F(u)=-\ell_{F, u}(v)+o(\|v\|) .
$$

Summing, the terms involving $\ell_{F, u}$ cancel, and we have

$$
F(u+v)+F(u-v)-2 F(u)=o(\|v\|) .
$$

In particular, a necessary condition for Frechét differentiability the norm functional on a Banach space is that

$$
\left\|\frac{u+v}{2}\right\|+\left\|\frac{u-v}{2}\right\|-\|u\|=o(\|v\|) .
$$

This brings us to the following definition:
3.5 DEFINITION (Uniform Smoothness). Let $V$ be a Banach space with norm $\|\cdot\|$. The modulus of smoothness of $V$ is the function $\rho_{V}(\tau)$ defined by

$$
\begin{equation*}
\rho_{V}(\tau)=\sup \left\{\left\|\frac{u+\tau v}{2}\right\|+\left\|\frac{u-\tau v}{2}\right\|-1:\|u\|=\|v\|=1\right\} \tag{3.13}
\end{equation*}
$$

for each $\tau \geq 0$. Then $V$ is said to be uniformly smooth in case $\rho_{V}(\tau)=o(\tau)$, i.e., if

$$
\begin{equation*}
\lim _{\tau \rightarrow 0} \frac{\rho_{V}(\tau)}{\tau}=0 \tag{3.14}
\end{equation*}
$$

When there is no ambiguity as to which space $V$ is intended, we write $\rho(\tau)$ in place of $\rho_{V}(\tau)$.
It is easy to see that uniform smoothness fails for $L^{1}(M, \mathcal{M}, \mu)$ and $L^{\infty}(M, \mathcal{M} \mu)$, even for a two-point measure space, while $L^{2}(M, \mathcal{M}, \mu)$ is uniformly smooth. This is left as an exercise. In fact, it is a good exercise to compute the moduli of smoothness for these spaces. The results are:
(1) When $V=L^{1}(M, \mathcal{M}, \mu), \delta_{V}(\epsilon)=0$ and $\rho_{V}(\tau)=\tau$.
(2) When $V=L^{2}(M, \mathcal{M}, \mu), \delta_{V}(\epsilon)=1-\sqrt{1-\epsilon^{2}}$ and $\rho_{V}(\tau)=\sqrt{1+\tau^{2}}-1$.
(3) When $V=L^{\infty}(M, \mathcal{M}, \mu), \delta_{V}(\epsilon)=0$ and $\rho_{V}(\tau)=\tau$.

The next theorem gives the relation between uniform convexity and uniform smoothness. Before stating it, we introduce the notion of a dual pair of Banach spaces.
3.6 DEFINITION (Dual Pairs). A dual pair of Banach spaces is a pair of Banach spaces $V$ and $W$ with norms $\|\cdot\|_{V}$ and $\|\cdot\|_{W}$ respectively, and a bilinear form $\langle\cdot, \cdot\rangle$ on $V \times W$ so that for all $v \in V$,

$$
\begin{equation*}
\|v\|_{V}=\sup \left\{|\langle v, w\rangle|:\|w\|_{W} \leq 1\right\} \tag{3.15}
\end{equation*}
$$

and

$$
\begin{equation*}
\|w\|_{W}=\sup \left\{|\langle v, w\rangle|:\|v\|_{V} \leq 1\right\} \tag{3.16}
\end{equation*}
$$

The primary example is that in which $W=V^{*}$, and

$$
\langle v, w\rangle=w(v)
$$

Then (3.16) holds by the definition of the norm on $V^{*}$, while (3.15) holds by the Hahn-Banach Theorem, which asserts the existence of a $w \in V^{*}$ with $\|w\|=1$ and $w(v)=\|v\|$.

When $V$ and $W$ are a dual pair, there is a map from $V$ into $W^{*}$ which assigns to $v$ the linear functional

$$
f_{v}(\cdot)=\langle v, \cdot\rangle
$$

By (3.16), and the definition of the dual norm $\|\cdot\|_{*}$,

$$
\left\|f_{w}\right\|_{*}=\|w\|_{W}
$$

Hence the map $w \mapsto\langle\cdot, w\rangle$, which is clearly linear, is also an isometry.
However, it need not be the case that its image is all of $V^{*}$. In summary,

- When $V$ and $W$ are a dual pair, $W$ may be identified with a subset of $V^{*}$ through the isometric linear transformation

$$
w \mapsto\langle\cdot, v\rangle
$$

However, it is not necessarily the case that every $f \in V^{*}$ is in the range of this transformation.
We now prove that when $V$ and $W$ are a dual pair the moduli of smoothness and convexity of the one space can be determined from those of the other.
3.7 THEOREM (Lindenstrauss-Day Theorem). Let $V$ and $W$ be a dual pair of Banach spaces. Then $W$ is uniformly smooth if and only if $V$ is uniformly convex. Moreover,

$$
\begin{equation*}
\rho_{W}(\tau)=\sup _{0 \leq \epsilon \leq 1}\left\{\epsilon \tau-\delta_{V}(\epsilon)\right\} . \tag{3.17}
\end{equation*}
$$

Before giving the proof of Theorem 3.7, we give three simple but important applications.
3.8 THEOREM (Uniqueness and continuity of unit tangent functionals). If $V$ is a uniformly smooth Banach space, then for each non-zero $v \in V$, there exists a unique unit vector $f_{v} \in V^{*}$ such that $f_{v}(v)=\|v\|$. Moreover, the map $v \mapsto f_{v}$ is continuous in the norm topologies.

Proof. The Hahn-Banach Theorem tell us that the linear functional $f_{v}$ exists; the points to be shown are the uniqueness and the continuity. By the Lindenstrauss-Day Theorem, $V^{*}$ is uniformly convex. If $f_{v}$ and $g_{v}$ were two distinct unit vectors in $V^{*}$ such that $f_{v}(v)=g_{v}(v)=\|v\|$, then $\frac{1}{2}\left(f_{v}+g_{v}\right)(v)=\|v\|$ but $\left\|f_{v}+g_{v}\right\|_{*}<1$, which is impossible.
3.9 THEOREM (Differentiability of the Norm). Let $V$ be a uniformly smooth Banach space. Then the norm on $V$ is continuously Frechét differentiable at all $v \neq 0$ in $V$, and the derivative is given by $\mathcal{R}\left(f_{v}\right)$, where $f_{v}$ is the unique unit vector in $V^{*}$ with $f_{v}(v)=\|v\|$

Proof. Since $V^{*}$ is uniformly convex, for each $u \in V$, there exists a unique unit vector $f_{u} \in V^{*}$ so that $f_{u}(u)=\|u\|$. Hence,

$$
\|v+w\|=f_{v+w}(v+w) \leq \mathcal{R}\left(f_{v+w}(v)\right)+\mathcal{R}\left(f_{v+w}(w)\right) \leq\|v\|+\mathcal{R}\left(f_{v+w}(w)\right) .
$$

On the other hand,

$$
\|v+w\| \geq \mathcal{R}\left(f_{v}(v+w)\right)=\mathcal{R}\left(f_{v}(v)\right)+\mathcal{R}\left(f_{v}(w)\right)=\|v\|+\mathcal{R}\left(f_{v}(w)\right)
$$

Altogether,

$$
0 \leq\|v+w\|-\|v\|-\mathcal{R}\left(f_{v}(w)\right) \leq \mathcal{R}\left(f_{v+w}(w)\right)-\mathcal{R}\left(f_{v}(w)\right) \leq\left\|f_{v+w}-f_{v}\right\|_{*}\|w\|
$$

Hence $\left|\|v+w\|-\|v\|-\mathcal{R}\left(f_{v}(w)\right)\right| \leq\left\|f_{v+w}-f_{v}\right\|_{*}\|w\|=o(\|w\|)$ by Theorem 3.8.
Now let $V$ be a Banach space, and let $V^{*}$ be the dual space of linear functionals on $V$, and let $V^{* *}$ be the dual space of continuous linear functionals on $V^{*}$. We have seen that in case $V=L^{2}(M, \mathcal{S}, \mu), V^{*}$ can be identified with $V$, and so $V^{* *}$ can as well.

This is a rather special circumstance. However, it is frequently the case that $V^{* *}=V$. Let $V$ be a Banach space, and consider the isometric mapping $v \mapsto L_{v} \in V^{* *}$, where

$$
L_{v}(f)=f(v)
$$

for all $f \in V^{*}$. Recall that $V$ is called reflexive in case the image of this mapping is all of $V^{* *}$, which we express by writing $V=V^{* *}$.
3.10 THEOREM (Millman). A uniformly convex Banach space is reflexive.

Proof. By the Lindenstrauss-Day Theorem, $V^{* *}$ is uniformly convex. For any unit vector $L \in V^{* *}$, and any $\epsilon>0$, pick a unit vector $f \in V^{*}$ so that $L(f)>1-\epsilon$ which is possible by the definition of the $\|\cdot\|_{* *}$ norm. Since $V$ is uniformly convex, there is a unit vector $v \in V$ with $f(v)=1$. Let $L_{v}$ be the corresponding element of $V^{* *}$, given by $L_{v}(g)=g(v)$ for al $g \in V^{*}$. We then have

$$
\left\|\frac{L+L_{v}}{2}\right\|_{* *} \geq \frac{L(f)+L_{v}(f)}{2}=\left(\frac{L+L_{v}}{2}\right)(f) \geq 1-\frac{\epsilon}{2} .
$$

It follows that $\left\|\frac{L-L_{v}}{2}\right\|_{* *} \leq \sup _{0 \leq s \leq 1}\left\{s: \delta_{V^{* *}}(s)<\epsilon / 2\right\}$.
Proof of Theorem 3.7. We will use $f$ and $g$ to denote elements of $W$, and $u$ and $v$ to denote elements of $V$. We will leave subscripts off the norms as this convention makes it clear which norm is intended.

The first step of the proof is to show that

$$
\begin{equation*}
\rho_{W}(\tau)+\delta_{V}(\epsilon) \geq \tau \epsilon \tag{3.18}
\end{equation*}
$$

for all $\tau \geq 0$ and all $0 \leq \epsilon \leq 1$. To see this, fix any such $\tau$ and $\epsilon$. Take any $u$ and $v$ in $V$ with $\|u\|=\|v\|=1$ and $\|u-v\| \geq 2 \epsilon$.

Since $V$ and $W$ are a dual pair, for any $\eta>0$, there are unit vectors $f$ and $g$ in $W$ with

$$
\langle f,(u+v) / 2\rangle \geq\left\|\frac{u+v}{2}\right\|-\eta \quad \text { and } \quad\langle f,(u-v) / 2\rangle \geq\left\|\frac{u-v}{2}\right\|-\eta .
$$

Then

$$
\begin{align*}
\rho_{W}(\tau) & \geq\left\|\frac{f+\tau g}{2}\right\|+\left\|\frac{f-\tau g}{2}\right\|-1 \\
& \geq\langle(f+\tau g) / 2, v\rangle+\langle(f-\tau g) / 2, v\rangle-1 \\
& =\langle f,(u+v) / 2\rangle+\tau\langle g,(u-v) / 2\rangle-1 \\
& \geq\left\|\frac{u+v}{2}\right\|+\tau\left\|\frac{u-v}{2}\right\|-1-2 \eta \geq\left\|\frac{u+v}{2}\right\|+\tau \epsilon-1-2 \eta \tag{3.19}
\end{align*}
$$

Hence $\rho_{W}(\tau)+\left(1-\left\|\frac{u+v}{2}\right\|\right) \geq \tau \epsilon-2 \eta$ By the definition of $\delta_{V}$, and the fact that $\eta>0$ is arbitrary, this proves (3.21).

The second step is to prove an upper bound on $\rho_{W}$. To do this, fix any $\tau>0$ and any unit vectors $f$ and $g$ in $W$. Fix any $\eta>0$, and choose unit vectors $u_{\tau}$ and $v_{\tau}$ in $V$ with

$$
\begin{equation*}
\left\langle(f+\tau g), u_{\tau}\right\rangle \geq\|f+\tau g\|-\eta \quad \text { and } \quad\left\langle(f-\tau g), v_{\tau}\right\rangle \geq\|f-\tau g\|-\eta \tag{3.20}
\end{equation*}
$$

Then

$$
\begin{aligned}
\left\|\frac{f+\tau g}{2}\right\|+\left\|\frac{f-\tau g}{2}\right\| & \leq \frac{\left\langle(f+\tau g), u_{\tau}\right\rangle}{2}+\frac{\left\langle(f-\tau g), v_{\tau}\right\rangle}{2}+\eta \\
& =\frac{\left\langle f, u_{\tau}+v_{\tau}\right\rangle}{2}+\tau \frac{\left\langle g, u_{\tau}-v_{\tau}\right\rangle}{2}+\eta \leq\left\|\frac{u_{\tau}+v_{\tau}}{2}\right\|+\tau\left\|\frac{u_{\tau}-v_{\tau}}{2}\right\|+\eta
\end{aligned}
$$

Now define $\epsilon_{\tau}:=\left\|\frac{u_{\tau}-v_{\tau}}{2}\right\|$ so that $0 \leq \epsilon_{\tau} \leq 1$ and $\left\|\frac{u_{\tau}+v_{\tau}}{2}\right\| \leq 1-\delta_{V}\left(\epsilon_{\tau}\right)$. Therefore,

$$
\begin{equation*}
\left(\left\|\frac{f+\tau g}{2}\right\|+\left\|\frac{f-\tau g}{2}\right\|-1\right) \leq \tau \epsilon_{\tau}-\delta_{V}\left(\epsilon_{\tau}\right)+\eta . \tag{3.21}
\end{equation*}
$$

By the definition of $\rho_{W}, \rho_{W}(\tau) \leq \tau \epsilon_{\tau}-\delta_{V}\left(\epsilon_{\tau}\right)+\eta \leq \sup _{0 \leq \epsilon \leq 1}\left\{\epsilon \tau-\delta_{V}(\epsilon)\right\}+\eta$. Since $\eta>0$ is arbitrary, this proves that $\rho_{W}(\tau) \leq \sup _{0 \leq \epsilon \leq 1}\left\{\epsilon \tau-\delta_{V}(\epsilon)\right\}$, and together with the lower bound (3.21), this proves (3.17).

Now suppose that $W$ is uniformly smooth, and consider $\epsilon \in(0,1)$. By (??), or even (3.21), $\delta_{V}(\epsilon) \geq \sup _{\tau \geq 0}\left\{\epsilon \tau-\rho_{W}(\tau)\right\}$. Since $\rho_{W}(\tau)=o(\tau)$, there is a $\tau_{\epsilon}>0$ so that $\frac{\rho_{W}\left(\tau_{\epsilon}\right)}{\tau_{\epsilon}} \leq \frac{\epsilon}{2}$. Therefore,

$$
\delta_{V}(\epsilon) \geq \frac{\epsilon \tau_{\epsilon}}{2}>0,
$$

and since $\epsilon>0$ is arbitrary, $V$ is uniformly convex.
Now suppose that $V$ is uniformly convex. It follows from (3.17) that for any $\tau>0$

$$
\frac{\rho_{W}(\tau)}{\tau} \leq \sup _{0 \leq \epsilon \leq 1}\left\{\epsilon-\frac{\delta_{V}(\epsilon)}{\tau}\right\}
$$

Therefore, if we define $\epsilon_{\tau}=\inf \left\{\epsilon>0: \delta_{V}(\epsilon)>\tau\right\}$, then $\epsilon_{\tau}>0$, and for $\epsilon_{\tau}<\epsilon \leq 1, \epsilon-\delta_{V}(\epsilon) / \tau \leq 0$, and hence

$$
\sup _{0 \leq \epsilon \leq 1}\left\{\epsilon-\frac{\delta_{V}(\epsilon)}{\tau}\right\}=\sup _{0 \leq \epsilon \leq \epsilon_{\tau}}\left\{\epsilon-\frac{\delta_{V}(\epsilon)}{\tau}\right\} \leq \epsilon_{\tau}
$$

Since $\lim _{\tau \downarrow 0} \epsilon-\tau=0, W$ is uniformly smooth.

## 4 Uniform convexity and smoothness in $L^{p}$ spaces

In any Hilbert space, in particular $L^{2}(\Omega, \mathcal{M}, \mu)$, we have the parallelogram identity:

$$
\left\|\frac{f+g}{2}\right\|_{2}^{2}+\left\|\frac{f-g}{2}\right\|_{2}^{2}=\frac{\|f\|_{2}^{2}+\|g\|_{2}^{2}}{2}
$$

If $f$ and $g$ are unit vectors, this yields

$$
\left\|\frac{f+g}{2}\right\|_{2} \leq\left(1-\left\|\frac{f-g}{2}\right\|_{2}^{2}\right)^{1 / 2} \leq 1-\frac{1}{2}\left\|\frac{f-g}{2}\right\|_{2}^{2}
$$

and hence $\delta_{L^{2}}(\epsilon) \geq \frac{1}{2} \epsilon^{2}$.
There is a close analog of the paralleleogram law in $L^{p}(\Omega, \mathcal{M}, \mu), p>2$ : Recall that for counting measure, $\|f\|_{p} \geq\|f\|_{q}$ for $p<q$, while for any probability measure, $\|f\|_{p} \leq\|f\|_{q}$ for $p<q$.

Therefore, for all $a, b>0$,

$$
\left(\left|\frac{a+b}{2}\right|^{p}+\left|\frac{a-b}{2}\right|^{p}\right)^{1 / p} \leq\left(\left|\frac{a+b}{2}\right|^{2}+\left|\frac{a-b}{2}\right|^{2}\right)^{1 / 2}=\left(\frac{a^{2}+b^{2}}{2}\right)^{1 / 2} \leq\left(\frac{a^{p}+b^{p}}{2}\right)^{1 / p}
$$

Next, for all $z, w \in \mathbb{C}$, and all $p \geq 2$,

$$
|z+w|^{p}+|z-w|^{p} \leq\left\|z|+|w||^{p}+\right\| z|-|w||^{p} .
$$

Combining, we obtain, $\left|\frac{f(x)+g(x)}{2}\right|^{p}+\left|\frac{f(x)-g(x)}{2}\right|^{p} \leq \frac{|f(x)|^{p}+|g(x)|^{p}}{2}$. Integrating in $x$ yields Clarkson's inequality:

$$
\left\|\frac{f+g}{2}\right\|_{p}^{p}+\left\|\frac{f-g}{2}\right\|_{p}^{p} \leq \frac{\|f\|_{p}^{p}+\|g\|_{p}^{p}}{2} .
$$

If $f$ and $g$ are unit vectors, this becomes

$$
\left\|\frac{f+g}{2}\right\|_{p} \leq\left(1-\left\|\frac{f-g}{2}\right\|_{p}^{p}\right)^{1 / p} \leq 1-\frac{1}{p}\left\|\frac{f-g}{2}\right\|_{p}^{p}
$$

Thus we see that

$$
\delta_{L^{p}}(\epsilon) \geq \frac{1}{p} \epsilon^{p} .
$$

Uniform convexity for $1<p<2$ is more subtle, and the result is somewhat surprising: It turns out that for $1<p \leq 2$,

$$
\delta_{L^{p}}(\epsilon) \geq \frac{p-1}{2} \epsilon^{2} .
$$

Notice that the exponent is 2 , as in the Hilbert space case. However, as $p$ decreases towards 1 , the constant $(p-1) / 2$ decreases to zero. Both the exponent 2 , and the constant $(p-1) / 2$ are best possible, and both have significant implications. The result is can be obtained from a result of Hanner, who exactly computed $\delta_{L^{p}}(\epsilon)$ for all $1<p<\infty$. The final remark in his 1955 paper is (in slightly different notation) that

$$
\begin{equation*}
\delta_{L^{p}}(\epsilon)=\frac{p-1}{2} \epsilon^{2}+\mathcal{O}\left(\epsilon^{3}\right), \tag{4.1}
\end{equation*}
$$

which certinaly shows that $\delta_{L^{p}}(\epsilon)$ is bounded below by some multiple of $\epsilon^{2}$. The fact that the remainder term is positive and may be droped, yielding the asserted lower bound, may be folklore, but appears in work by Ball and Pisier in the 1990's. They used Hanner's exact computationof $\delta_{L^{p}}$ and controlled the sign of the remainder term in (4.1). However, the sharp bound may be proved directly, as we now explain.

Let $f$ and $g$ be simple functions of the form

$$
f(x)=\sum_{j=1}^{n} z_{j} 1_{A_{j}}(x) \quad \text { and } \quad g(x)=\sum_{j=1}^{n} w_{j} 1_{A_{j}}(x),
$$

where for each $j, z_{j} w_{j}^{*}$ is not real. This guarantees that $z_{j}+t w_{j} \neq 0$ for any real $t$, and thus for all $x \in \cup_{j=1}^{n} A_{j}$, and all $t \in \mathbb{R}, f(x)+t g(x) \neq 0$. Define

$$
Y(t)=\|f+t g\|_{p}^{p} \quad \text { and } \quad q=\frac{p}{2}
$$

so that $\|f+t g\|_{p}^{2}=Y^{1 / q}(t)$. Differentiating twice,

$$
\begin{aligned}
\frac{\mathrm{d}^{2}}{\mathrm{~d} t^{2}}\|f+t g\|_{p}^{2} & =\frac{1}{q}\left(\frac{1}{q}-1\right) Y^{1 / q-2}\left(Y^{\prime}\right)^{2}+\frac{1}{q} Y^{1 / q-1} Y^{\prime \prime} \\
& \geq \frac{1}{q} Y^{1 / q-1} Y^{\prime \prime}
\end{aligned}
$$

A simple calculation yields $Y^{\prime \prime}(t) \geq p(p-1) \int|f+t g|^{2 q-2}|g|^{2} \mathrm{~d} \mu$. To this we apply the reverse Hölder inequality, which says that for $0<r<1$ and $s=r /(r-1)$, whenever $a_{j} \geq 0$ for $j=1, \ldots, n$, and $b_{j}>0$ for $j=1, \ldots, n$,

$$
\sum_{j=1}^{n} a_{j} b_{j} \geq\left(\sum_{j=1}^{n} a_{j}^{r}\right)^{1 / r}\left(\sum_{j=1}^{n} b_{j}^{s}\right)^{1 / s}
$$

The result is that, for all $t, \frac{\mathrm{~d}^{2}}{\mathrm{~d} t^{2}}\|f+t g\|_{p}^{2} \geq 2(p-1)\|g\|_{p}^{2}$. Let $\psi^{\prime \prime}(t) \geq 2 c$ for all $t$, and define

$$
\varphi(t):=\psi(t)+c t(1-t) .
$$

Then $\varphi$ is convex, and thus

$$
\varphi(1 / 2) \leq \frac{\varphi(0)+\varphi(1)}{2}, \quad \text { that is, } \quad \psi(1 / 2)+\frac{c}{4} \leq \frac{\psi(0)+\psi(1)}{2}
$$

We conclude that with $f$ and $g$ as above,

$$
\|f+g / 2\|_{p}^{2}+\frac{p-1}{4}\|g\|_{p}^{2} \leq \frac{\|f\|_{p}^{2}+\|f+g\|_{p}^{2}}{2} .
$$

The simple function approximation is now easily removed.
Now let $u$ and $v$ be vectors in $L^{p}$ space, $1<p \leq 2$, and let $f=u$ and $g=v-u$. Then

$$
\left\|\frac{u+v}{2}\right\|_{p}^{2}+(p-1)\left\|\frac{u-v}{2}\right\|_{p}^{2} \leq \frac{\|u\|_{p}^{2}+\|v\|_{p}^{2}}{2}
$$

If $u$ and $v$ are unit vectors, the right hand side is 1 , and this implies

$$
\left\|\frac{u+v}{2}\right\|_{p} \leq 1-\frac{p-1}{2}\left\|\frac{u-v}{2}\right\|_{p}^{2}
$$

which proves that

$$
\delta_{L^{p}}(\epsilon) \geq \frac{p-1}{2} \epsilon^{2} .
$$

We now have the following result:
4.1 THEOREM (Uniform convexity of $\left.L^{p}, 1<p<\infty\right)$. For any measure space $(M, \mathcal{M}, \mu)$ and any $L^{p}(M, \mathcal{M}, \mu)$ is uniformly convex. For $1<p<2$, one has the bound

$$
\begin{equation*}
\delta_{L^{p}}(\epsilon) \geq \frac{p-1}{2} \epsilon^{2} \tag{4.2}
\end{equation*}
$$

while for $2<p<\infty$, one has the bound

$$
\begin{equation*}
\delta_{L^{p}}(\epsilon) \geq \frac{1}{p} \epsilon^{p} \tag{4.3}
\end{equation*}
$$

Proof. In the discussion just above, we have proved the bounds on uniform convexity, and the left halves of (4.2) and (4.3).

We are now give two proofs of the Riesz Representation Theorem for $L^{p}, 1<p<\infty$.
4.2 THEOREM (Riesz Representation Theorem for $\left.L^{p}, 1<p<\infty\right)$. Let ( $M, \mathcal{M}, \mu$ ) be any measure space. Let $1 \leq p \leq \infty$, and let $q=p /(1-p)$. Then the map from $L^{q}$ into $\left(L^{p}\right)^{*}$ given by $g \mapsto \varphi_{g}$ where

$$
\varphi_{g}(f)=\int_{M} f g \mathrm{~d} \mu, \quad f \in L^{p}
$$

is an isometry from $L^{q}$ into $\left(L^{p}\right)^{*}$.
Before giveing our two proofs, let us take stock of what has been dealt with, and what remains to be dealt with. We have already seen, as a consequence of Hölder's inequality, that for every $g \in L^{q},\left\|\varphi_{g}\right\|_{\left(L^{p}\right)^{*}}=\|g\|_{q}$, and hence $g \mapsto \varphi_{g}$ is an isometric map into ( $\left.L^{p}\right)^{*}$. It remains to be shown that this map is onto $\left(L^{p}\right)^{*}$. We now give two proofs of this.

First Proof of Theorem 4.2. Let $1<p<\infty$. Let $V$ be the range of the mapping $g \mapsto \varphi_{g}$ in $\left(L^{p}\right)^{*}$. Since the map is an isometry, and since $L^{q}$ is complete, $V$ is a closed subspace of $\left(L^{p}\right)^{*}$. If $V$ is a proper subspace of $\left(L^{p}\right)^{*}$, there exists a non-zero $\varphi \in\left(L^{p}\right)^{*} \backslash V$ and then by the Hahn-Banach Theorem, there is an $L \in\left(L^{p}\right)^{* *}$ such that

$$
\begin{equation*}
L(\varphi)=\|\varphi\|_{\left(L^{p}\right)^{*}} \neq 0, \tag{4.4}
\end{equation*}
$$

and $L\left(\varphi_{g}\right)=0$ for all $g \in L^{q}$.
However, by Millman's Theorem, since $L^{p}$ is uniformly convex, it is reflexive, and so there exists an $f \in L^{p}$ so that

$$
L(\psi)=\psi(f) \quad \text { for all } \quad \psi \in\left(L^{p}\right)^{*}
$$

Therefore, for all $g \in L^{q}$, since $\varphi_{g} \in V$,

$$
0=L\left(\varphi_{g}\right)=\varphi_{g}(f)=\int_{M} g f \mathrm{~d} \mu
$$

But since

$$
\|f\|_{p}=\sup \left\{\int_{M} g f \mathrm{~d} \mu:\|g\|_{q}=1\right\}
$$

it would follow that $\|f\|_{p}=0$, and hence $L=0$. This is contradicts (4.4), and hence $V$ is not a proper subspace of $\left(L^{p}\right)^{*}$.

Second Proof of Theorem 4.2. Let $1<p<\infty$. Since $L^{p}$ is uniformly convex, for each $\varphi \in\left(L^{p}\right)^{*}$, there exists a unique $f_{\varphi} \in L^{p}$ with $f_{\varphi} \in L^{p}$ and $\varphi\left(f_{\varphi}\right)=\|\varphi\|_{\left(L^{p}\right)^{*}}$. Then, for any $g \in L^{p}$, the function $t \mapsto \varphi\left(\frac{f_{\varphi}+t g}{\left\|f_{\varphi}+t g\right\|_{p}}\right)$ has a maximum at $t=0$.

If we assume for the moment that $t \mapsto\left\|f_{\varphi}+t g\right\|_{p}$ is differentiable at $t=0$, then

$$
0=\left.\frac{\mathrm{d}}{\mathrm{~d} t} \varphi\left(\frac{f_{\varphi}+t g}{\left\|f_{\varphi}+t g\right\|_{p}}\right)\right|_{t=0}=\mathcal{R} \varphi(g)-\left.\|\varphi\|_{\left(L^{p}\right)^{*}} \frac{\mathrm{~d}}{\mathrm{~d} t}\left\|f_{\varphi}+t g\right\|_{p}\right|_{t=0}
$$

By the convexity of $x \mapsto|x|^{p}$, for all $0<t<1$ and all $x \in M$,

$$
\left|f_{\varphi}(x)\right|^{p}-\left|f_{\varphi}(x)-g(x)\right|^{p} \leq \frac{\left|f_{\varphi}(x)+t g(x)\right|^{p}-\left|f_{\varphi}(x)\right|^{p}}{t} \leq\left|f_{\varphi}(x)+g(x)\right|^{p}-\left|f_{\varphi}(x)\right|^{p} .
$$

Then, since $\left|f_{\varphi}-g\right|^{p},\left|f_{\varphi}-g\right|^{p}$ and $\left|f_{\varphi}\right|^{p}$ are all integrable, The Dominated Convergence Theorem yields us

$$
\lim _{t \rightarrow 0} \int_{M} \frac{\left|f_{\varphi}(x)+t g(x)\right|^{p}-\left|f_{\varphi}(x)\right|^{p}}{t} \mathrm{~d} \mu=\int_{M} \lim _{t \rightarrow 0} \frac{\left|f_{\varphi}(x)+t g(x)\right|^{p}-\left|f_{\varphi}(x)\right|^{p}}{t} \mathrm{~d} \mu .
$$

Now one easily computes that for all $x$,

$$
\lim _{t \rightarrow 0} \frac{\left|f_{\varphi}(x)+t g(x)\right|^{p}-\left|f_{\varphi}(x)\right|^{p}}{t}=\mathcal{R}\left|f_{\varphi}\right|^{p-2} \overline{f_{\varphi}}(x) g(x) .
$$

Thus, for all $g \in L^{q}$, we have

$$
\mathcal{R} \varphi(g)=\|\varphi\|_{\left(L^{p}\right)^{*}} \int_{M} \mathcal{R}\left|f_{\varphi}\right|^{p-2} \overline{f_{\varphi}}(x) g(x) \mathrm{d} \mu
$$

Substituting $g$ by $i g$, we obtain the same result for the imaginary part, and hence

$$
\varphi(g)=\|\varphi\|_{\left(L^{p}\right)^{*}} \int_{M}\left|f_{\varphi}\right|^{p-2} \overline{f_{\varphi}}(x) g(x) \mathrm{d} \mu .
$$

It is now easily checked that $\left|f_{\varphi}\right|^{p-2} \overline{f_{\varphi}}$ is a unit vector in $L^{q}$, and hence $\varphi$ is in the range of our isometry into $\left(L^{p}\right)^{*}$. But since $\varphi$ is an arbitrary element of $\left(L^{p}\right)^{*}$, we see that our isometry is onto $\left(L^{p}\right)^{*}$.

Next, as a consequence of Theorem 4.1 the Lindenstrauss-Day Theorem ,we obtain the following result:
4.3 THEOREM (Uniform smoothness of $\left.L^{p}, 1<p<\infty\right)$. For any measure space $(M, \mathcal{M}, \mu)$ and any $L^{p}(M, \mathcal{M}, \mu)$ is uniformly smooth. For $1<p<2$,one has the bound

$$
\begin{equation*}
\rho_{L^{p}}(\tau) \leq \frac{1}{2(p-1)} \tau^{2}, \tag{4.5}
\end{equation*}
$$

while for $2<p<\infty$ and $q=p /(p-1)$, one has the bound

$$
\begin{equation*}
\rho_{L^{p}}(\tau) \leq \frac{1}{q} \tau^{2} . \tag{4.6}
\end{equation*}
$$

Proof. The uniform smoothness follows directly from Theorems 3.7 and 4.1. To obtain (4.5) use (3.17) to deduce

$$
\rho_{L^{p}}(\tau) \leq \sup _{0 \leq \epsilon \leq 1}\left\{\epsilon \tau-\delta_{L^{q}}(\epsilon)\right\}
$$

Then by (4.2) and a simple calculation, one obtains and (4.5). The proof of (4.6) is similar.


[^0]:    ${ }^{1}$ © 2017 by the author. This article may be reproduced, in its entirety, for non-commercial purposes.

