Nikishin-Stein Theorem

Graduate Analysis Seminar
October 5, October 12, 2011
University of Wisconsin-Madison

In the following we shall consider only finite measure space (X, m). Without loss of generality, we always assume $m(X) = 1$. All the spaces are over field $\mathbb{K}$, where $\mathbb{K}$ can be $\mathbb{R}$ or $\mathbb{C}$.

Let B be a Banach space. A function $\varphi : X \to B$ is called simple if $\varphi(X)$ is finite and for any $y \in \varphi(X)$, $\varphi^{-1}(y)$ is measurable. A function $f : X \to B$ is called (strongly) measurable if there exists a sequence of simple functions φ_k such that $\varphi_k(x) \to f(x)$ for all x in X.

Let $Z := \{f : X \to B \text{ measurable}, \exists N, m(N) = 0, \text{ s.t. } f(x) = 0, \forall x \in X - N\}$, define $L^0_B(m) := \{f : X \to B \text{ measurable}\}/Z$. When $B = \mathbb{K}$, we denote $L^0_{\mathbb{K}}(m)$ by $L^0(m)$. Let $f_k, f \in L^0_B(m)$, we say that $f_k \to f$ in measure if $\|f(\cdot) - f_k(\cdot)\|_B \to 0$ in measure. There exists a unique metrizable topology on $L^0_B(m)$ making $L^0_B(m)$ into a topological vector space (TVS) s.t. $f_k \to f$ if and only if $f_k \to f$ in measure. Moreover, $L^0_B(m)$ is complete w.r.t. this topology, i.e. $L^0_B(m)$ is a F-space. However, $L^0_B(m)$ is not locally convex.

A function $f = (f_i)_{i=0}^{\infty} : X \to \mathbb{K}^{\mathbb{N}}$ is called componentwise measurable if $\forall i \in \mathbb{N}$, $f_i : X \to \mathbb{K}$ is measurable. Let $Z := \{f : X \to \mathbb{K}^{\mathbb{N}} \text{componentwise measurable}, \exists N, m(N) = 0, \text{ s.t. } f(x) = 0, \forall x \in X - N\}$, $L^0(m, \mathbb{K}^{\mathbb{N}}) := \{f : X \to \mathbb{K}^{\mathbb{N}} \text{componentwise measurable}\}/Z$,

$$L^0(m, l^{\infty}) := \{f = (f_i)_{i=0}^{\infty} \in L^0(m, \mathbb{K}^{\mathbb{N}}), \|f(x)\|_{l^{\infty}} = \sup_i |f_i(x)| < \infty, a.e.\}$$

Let $f_k, f \in L^0(m, l^{\infty})$, we say that $f_k \to f$ in measure if $\|f(\cdot) - f_k(\cdot)\|_{l^{\infty}} \to 0$ in measure. There exists a unique metrizable topology on $L^0(m, l^{\infty})$ making $L^0(m, l^{\infty})$ into a TVS s.t. $f_k \to f$ if and only if $f_k \to f$ in measure. Moreover, $L^0(m, l^{\infty})$ is complete w.r.t. this topology. Note that $L^0_{l^{\infty}}(m) \subset L^0(m, l^{\infty})$ and the subspace topology coincides the topology defined above. However, in general $L^0_{l^{\infty}}(m) \neq L^0(m, l^{\infty})$.

A set A is bounded in the TVS $L^0_B(m)$ (or $L^0(m, l^{\infty})$) if and only if there exists a function $C(\cdot) : (0, \infty) \to (0, \infty)$, $\lim_{\lambda \to \infty} C(\lambda) = 0$, s.t.

$$m(x \in X : \|f(x)\|_B > \lambda) \leq C(\lambda), \forall f \in A$$

Let V be a vector space. An operator $T : V \to L^0(m)$ is called sublinear if
i) $T(f) \geq 0, \forall f \in V$
ii) $T(\lambda f) = |\lambda| T(f), \forall \lambda \in \mathbb{K}, f \in V$
iii) $T(f + g) \leq T(f) + T(g), \forall f, g \in V$
Note that linear does not imply sublinear according to our definition. Our definition here is just for convenience.

An operator $T: V_1 \to V_2$ is called positive homogeneous if
$T(\alpha f) = \alpha T(f), \forall \alpha > 0, f \in V_1$

Let (Y, μ) be a measure space, V be a subspace of $L^0(Y, \mu)$. An operator $T: V \to L^0(m)$ is called positive if $|f| \leq g$ implies $|Tf| \leq |Tg|$.

Let V_1, V_2 be two TVSs, an operator $T : V_1 \to V_2$ is called bounded if T maps every bounded set to a bounded set. If V_1 is metrizable and T is linear, then T and continuous if and only if T is bounded.

Let Q be linear space, $\|\cdot\| : Q \to [0, \infty)$ is called a quasi-norm if
i) $\|x\| = 0 \Rightarrow x = 0$
ii) $\|\lambda x\| = |\lambda| \|x\|, \forall \lambda \in \mathbb{K}, x \in Q$
iii) $\exists$ a constant $K \geq 1$, s.t. $\|x + y\| \leq K(\|x\| + \|y\|), \forall x, y \in Q$

Let $(Q, \|\cdot\|)$ be a quasi-normed space, then there exists a unique metrizable topology on Q making Q into a TVS such that $x_k \to x$ if and only if $\|x - x_k\| \to 0$. Under this topology, $\{x : \|x\| < 1/n\}, n = 1, 2, ...$ form a topological basis of Q; a subset $E \subset Q$ is bounded if and only if $\|E\| = \{\|x\|, x \in E\}$ is bounded. Note that in general quasi-norm is not continuous, i.e. $\|x - x_k\| \to 0$ may not imply $\|x_k\| \to \|x\|$. If Q is complete in this topology, we call $(Q, \|\cdot\|)$ a quasi-Banach space. Let Q_1, Q_2 be two quasi-normed spaces, then a positive homogeneous operator $T : Q_1 \to Q_2$ is bounded if and only if there exists a constant C such that $\|Tf\|_{Q_2} \leq C\|f\|_{Q_1}, \forall f \in Q_1$.

Let Q be a quasi-normed space, B be a Banach space, then a positive homogeneous operator $T : Q \to L^0_B(m)$ (or $L^0(m, l^\infty)$) is bounded if and only if there exists a function $C(\cdot) : (0, \infty) \to (0, \infty)$, $\lim_{\lambda \to \infty} C(\lambda) = 0$, s.t.

$$m(x \in X : \|Tf(x)\|_B > \lambda) \leq C(\lambda), \forall \|f\|_Q \leq 1$$

If $T : Q \to L^0(m)$ is sublinear, then T is bounded if and only if T is continuous.

A bonus for the functional analytic argument above is the following, which is immediate by applying closed graph theorem.

Theorem I (**Banach Continuity Principle**) : Let Q be a quasi-Banach space, $T_i : Q \to L^0(m), i \in \mathbb{N}$, be a sequence of bounded linear operators. Let $\boldsymbol{T} : Q \to L^0(m, \mathbb{K}^{\mathbb{N}})$, $\boldsymbol{T}f := (T_i f)_{i=0}^\infty$. If $\boldsymbol{T}(Q) \subset L^0(m, l^\infty)$, i.e.

$$T^* f(x) := \|\boldsymbol{T}f(x)\|_{l^\infty} = \sup_i |T_i f(x)| < \infty \text{ a.e.}$$

then $\boldsymbol{T} : Q \to L^0(m, l^\infty)$ is bounded, i.e. $\exists C(\cdot) : (0, \infty) \to (0, \infty)$, $\lim_{\lambda \to \infty} C(\lambda) = 0$, s.t.

$$m(x \in X : T^* f(x) > \lambda) \leq C(\lambda), \forall \|f\|_Q \leq 1$$

Now we develop a little more language. Let $0 < p < \infty$, define

$$L^p(m) := \{f \in L^0(m), \|f\|_p = (\int_X |f|^p dm)^{1/p} < \infty\}$$

then $(L^p(m), \|\cdot\|_p)$ becomes a quasi-Banach space with quasi-norm constant $1\vee(2^{\frac{1}{p}-1})$. Note that $L^p(m)$ $(0<p<1)$ is in general not locally convex. Define

$$L^{p,\infty}(m) := \{f \in L^0(m), \|f\|_{p,\infty} = \sup_{\lambda>0} \lambda m(|f|>\lambda)^{1/p} < \infty\}$$

then $L^{p,\infty}(m)$ becomes a quasi-Banach space satisfying with quasi-norm constant $2\vee 2^{\frac{1}{p}}$. Let $w \in L^0(m), w>0$, then w induces a measure wm on X with $(wm)(A) := \int_A wdm$. Since $w>0$, we have $L^0(wm) = L^0(m)$. Similarly we have quasi-Banach spaces $L^p(wm)$ and $L^{p,\infty}(wm)$ for $0<p<\infty$, which are all contained in $L^0(m)$.

Let Q be a quasi-Banach space with quasi-norm constant K, $0<p<\infty$, define

$$l^p(Q) := \{q = (q_i)_{i=0}^\infty, q_i \in Q, \|q\|_{l^p(Q)} = (\sum_{i=0}^\infty \|q_i\|_Q^p)^{1/p} < \infty\}$$

then $l^p(Q)$ is a quasi-Banach space with quasi-norm constant $[1\vee(2^{\frac{1}{p}-1})]K$. Let $T: Q \to L^0(m)$ be an operator, denote

$$\boldsymbol{T}: l^p(Q) \to L^0(m, \mathbb{K}^{\mathbb{N}}), (f_i)_{i=0}^\infty \mapsto (Tf_i)_{i=0}^\infty$$

Denote by $l_0^p(Q) := \{q = (q_i)_{i=0}^\infty \in l^p(Q), \exists N, \text{s.t. } q_i = 0, \forall i \geq N\}$ which is dense in $l^p(Q)$.

For any measurable $E \subset X$, we have a measure space $(E, m|_E)$ and the restriction operator $R_E: L^0(m) \to L^0(E, m|_E), f \mapsto f|_E$.

Let Q_0 be the quasi-Banach space $L^q(m)$ or $L^{q,\infty}(m)$ with $0<q<\infty, 0<g\in L^0(m)$, then the map $M_g: Q_0 \to L^0(m), f \mapsto gf$ is continuous. Let Q be a another quasi-Banach space, $T_0: Q \to Q_0$ be continuous, then $T = M_gT_o: Q \to L^0(m)$ is continuous. In this case we say that T factors through Q_0. Notice that T factors through Q_0 if and only if there exists $g \in L^0(m), g>0$ s.t. $M_gT: Q \to Q_0$ continuous. We can make it fancier by saying that T is conformal to a continuous map from Q to Q_0.

The aim of this note is to prove the converse, i.e. roughly, every continuous map from Q to $L^0(m)$ is conformal to a continuous map from Q to some Q_0. For this we need some criteria.

Theorem II (**Equivalence of Boundedness**) : Let Q be a quasi-normed space, and let $T: Q \to L^0(m)$ be a positive homogeneous operator, $0<q<\infty$, then the following are equivalent:

a) (conformal boundedness) $gT: Q \to L^{q,\infty}(m)$ and is bounded, for some $0<g\in L^0(m)$

b) (weighted boundedness) $T: Q \to L^{q,\infty}(wm)$ and is bounded, for some $0<w\in L^0(m)$

c) (almost boundedness) For any $\epsilon>0$, there exists $E_\epsilon \subset X$ with $m(X-E_\epsilon)<\epsilon$, such that $R_{E_\epsilon}T: Q \to L^{q,\infty}(E_\epsilon, m|_{E_\epsilon})$ and is bounded

d) (vector-valued boundedness) $\boldsymbol{T}: l^q(Q) \to L^0(m, l^\infty)$ and is bounded

Proof. c) $\Rightarrow$ b): Let $\epsilon = \frac{1}{n}, n = 1,2,\ldots.$ Let E_n be the corresponding subset with $m(X-E_n) < \frac{1}{n}$ and $C_n>0$, such that

$$m\{x \in E_n : |T(f)| > \lambda\} \leq \frac{C_n^q}{\lambda^q}\|f\|_Q^q, \forall\lambda>0, f\in Q$$

We may assume $E_n \subset E_{n+1}$ (otherwise take the union). Now define ($E_0 := \emptyset$)

$$w := \sum_{n=1}^{\infty} \frac{1}{2^n C_n^q} \chi_{E_n - E_{n-1}}$$

Since $m(X - \cup_{n=1}^{\infty} E_n) = 0, w > 0$ a.e. Now $\forall \lambda > 0, f \in Q$,

$$\begin{aligned}
\int_{\{|Tf|>\lambda\}} wdm &= \int_{(E_n - E_{n-1}) \cap \{|Tf|>\lambda\}} wdm \\
&\leq \sum_{n=1}^{\infty} \frac{1}{2^n C_n^q} m\{x \in E_n : |T(f)(x)| > \lambda\} \\
&\leq \sum_{n=1}^{\infty} \frac{1}{2^n C_n^q} \frac{C_n^q}{\lambda^q} \|f\|_Q^q \\
&= \frac{1}{\lambda^q} \|f\|_Q^q
\end{aligned}$$

This shows $T : Q \to L^{q,\infty}(wm)$ and is bounded.

b) $\Rightarrow$ a): Assume $T : Q \to L^{q,\infty}(wm)$ and is bounded, i.e. $\exists C > 0$, such that

$$\int_{\{|Tf|>\lambda\}} wdm \leq \frac{C^q}{\lambda^q} \|f\|_Q^q, \forall \lambda > 0, f \in Q$$

We may assume $w \leq 1$ (otherwise let $w = w \wedge 1$). For $n = 1, 2, ...,$ let

$$E_n = \{x \in X : \frac{1}{n+1} < w(x) \leq \frac{1}{n}\}$$

$$g = \frac{1}{2^n} \chi_{E_n}$$

Then $g > 0$. Now we have

$$\begin{aligned}
m\{g|Tf| > \lambda\} &= \sum_{n=1}^{\infty} m\{x \in E_n : |T(f)(x)| > 2^n \lambda\} \\
&\leq \sum_{n=1}^{\infty} \int_{E_n \cap \{|Tf|>2^n\lambda\}} (n+1) wdm \\
&\leq \sum_{n=1}^{\infty} (n+1) \int_{\{|Tf|>2^n\lambda\}} wdm \\
&\leq \sum_{n=1}^{\infty} (n+1) \frac{C^q}{2^{nq}\lambda^q} \|f\|_Q^q \\
&= C_q \frac{C^q}{\lambda^q} \|f\|_Q^q
\end{aligned}$$

This shows $gT : Q \to L^{q,\infty}(m)$ and is bounded.

a) ⇒ d): We show that there exists $C(\cdot):(0,\infty)\to(0,\infty)$, $\lim\limits_{\lambda\to\infty} C(\lambda)=0$, such that

$$m\{x\in X:\sup_i |T(f_i)(x)|>\lambda\}\le C(\lambda), \forall (f_i)_{i=1}^\infty \in l^q(Q), \sum_{i=1}^\infty \|f_i\|_Q^q \le 1$$

Assume $g>0, gT:Q\to L^{q,\infty}(m)$ is bounded with constant $C>0$, then

$$\begin{aligned}
&m\{x\in X:\sup_i |T(f_i)(x)|>\lambda\}\\
=\,&m\{x\in X:\sup_i |g(x)T(f_i)(x)|>g(x)\lambda\}\\
\le\,&m\{x:g(x)\le\lambda^{-1/2}\}\cap\{x\in X:\sup_i |g(x)T(f_i)(x)|>g(x)\lambda\}\\
&+m\{x:g(x)>\lambda^{-1/2}\}\cap\{x\in X:\sup_i |g(x)T(f_i)(x)|>g(x)\lambda\}\\
\le\,&m\{x:g(x)\le\lambda^{-1/2}\}+m\{x\in X:\sup_i |g(x)T(f_i)(x)|>\lambda^{1/2}\}\\
\le\,&m\{x:g(x)\le\lambda^{-1/2}\}+m(\cup_{i=1}^\infty\{x\in X:|g(x)T(f_i)(x)|>\lambda^{1/2}\})\\
\le\,&m\{x:g(x)\le\lambda^{-1/2}\}+\sum_{i=1}^\infty m\{x\in X:|g(x)T(f_i)(x)|>\lambda^{1/2}\}\\
\le\,&m\{x:g(x)\le\lambda^{-1/2}\}+\sum_{i=1}^\infty \frac{C^q}{\lambda^{q/2}}\|f_i\|_Q^q\\
\le\,&m\{x:g(x)\le\lambda^{-1/2}\}+\frac{C^q}{\lambda^{q/2}}
\end{aligned}$$

Note that $C(\lambda):=m\{x:g(x)\le\lambda^{-1/2}\}+\frac{C^q}{\lambda^{q/2}}$ satisfies the properties we want.

d) ⇒ c): Fix $\epsilon>0$, we are going to find an E_ϵ as in c). Let $C(\cdot)$ be the function in d) (see the proof above), let $\Lambda>0$ such that $C(\Lambda)<\epsilon$. Consider $F\subset X$ satisfying the following:

$$m(F)>0 \text{ and } \exists f\in Q, \|f\|_Q\le 1, \text{ s.t. } |T(f)(x)|>\frac{\Lambda}{m(F)^{1/p}}, \text{ a.e. } x\in F$$

Let $\mathcal{F}_0=\{F\subset X: F \text{ satisfies the above property}\}$ (we may assume $\mathcal{F}_0\neq\emptyset$, similarly assume $\mathcal{F}_i\neq\emptyset$ in the following). Choose $F_1\in\mathcal{F}_0$ such that

$$m(F_1)>\frac{1}{2}\sup_{F\in\mathcal{F}_0} m(F)$$

Denote by f_1 the associated element in Q. Let $\mathcal{F}_1=\{F\in\mathcal{F}_0: F\cap F_1=\emptyset\}$. Choose $F_2\in\mathcal{F}_1$, such that

$$m(F_2)>\frac{1}{2}\sup_{F\in F_1} m(F)$$

Denote by f_2 the associated element in Q. Let $\mathcal{F}_2=\{F\in\mathcal{F}_0: F\cap F_1=F\cap F_2=\emptyset\}$, etc.

Since the F_i's are disjoint, $\sum_{i=1}^\infty m(F_i)\le 1$. In particular, $\lim_{i\to\infty} m(F_i)=0$. Thus by the construction there is no $F\in\mathcal{F}_0$ such that F does not intersect any F_i. Now let

$$E^c=\cup_{i=1}^\infty F_i$$

We claim that $\forall f \in Q, \|f\|_Q \leq 1, \lambda > 0$, we have

$$m\{x \in E : |T(f)(x)| > \lambda\} \leq \frac{\Lambda^q}{\lambda^q}$$

And hence $R_E T : Q \to L^{q,\infty}(E, m|_E)$ and is bounded. In fact, assume

$$m\{x \in E : |T(f)(x)| > \lambda\} > \frac{\Lambda^q}{\lambda^q}$$

for some $\lambda > 0, \|f\|_Q \leq 1$, then $F := \{x \in E : |T(f)(x)| > \lambda\} \in \mathcal{F}_0$. But F does not intersect any F_i, contradiction.

It remains to show that $m(E^c) < \epsilon$. Let $c_i = m(F_i)^{1/q}$, then

$$\sup_i |T(c_i f_i)(x)| > \Lambda \text{ a.e. } x \in E^c$$

Moreover, $\sum_{i=1}^{\infty} \|c_i f_i\|_Q^q \leq \sum_{i=1}^{\infty} m(F_i) \leq 1$. Thus

$$m(E^c) \leq m\{x \in X : \sup_i |T(c_i f_i)(x)| > \Lambda\} \leq C(\Lambda) < \epsilon. \quad \square$$

We will see later that when T is positive, d) can be easily verified. In the general case, in order to verify d) we will use the randomization trick.

The randomization trick is very useful in proving vector-valued (more precisely l^p-valued) inequalities. The idea is to express the l^p norm (which is essentially a sum of norms) by the norm of a random series (which is the norm of a sum), and hence we can make use of the linearity of the operator (or "almost independence" of the operators).

Let $(\Omega, \mathbb{P})$ be a probability space. A sequence of random variables $\{\epsilon_i\}_{i=0}^{\infty}$ is called a Rademacher sequence if

i) The ϵ_i's are independent

ii) $\epsilon_i \in \{-1, 1\}$ with $\mathbb{P}(\epsilon_i = 1) = \mathbb{P}(\epsilon_i = -1) = \frac{1}{2}$

Let $\{\epsilon_i\}_{i=0}^{\infty}$ be a Rademacher sequence, then $\{\epsilon_i\}_{i=0}^{\infty}$ forms an orthonormal system in $L^2(\Omega)$, thus we have an isometric embedding

$$l^2 \to L^2(\Omega), \alpha = \{\alpha_i\}_{i=0}^{\infty} \mapsto \epsilon \cdot \alpha := \sum_i \epsilon_i \alpha_i$$

We call $\epsilon \cdot \alpha$ a Rademacher series (by Kolmogorov's maximal inequality, it also converges almost surely). What is surprising is that $\epsilon \cdot \alpha$ actually lies in $L^p(\Omega)$ for all $0 < p < \infty$ (in fact $|\epsilon \cdot \alpha|$ is exponentially integrable), and $\|\epsilon \cdot \alpha\|_{L^p(\Omega)} \approx \|\alpha\|_{l^2}$ with the implicit constants depending only on p, that is,

Theorem III (**Khinchin's inequality, 1923**) : For all $0 < p < \infty$, there exist constants $A_p, B_p > 0$, such that

$$A_p \|\alpha\|_{l^2} \leq \|\epsilon \cdot \alpha\|_{L^p(\Omega)} \leq B_p \|\alpha\|_{l^2}, \forall \alpha \in l^2$$

Notice that $\|\epsilon \cdot \alpha\|_{L^p(\Omega)} = \mathbb{E}\left[|\sum_i \epsilon_i \alpha_i|^p\right]^{1/p}$ and $\|\alpha\|_{l^2} = \left[\sum_i |\alpha_i|^2\right]^{1/2}$, Khinchin's inequality reveals that, although samplewise $|\sum_i \epsilon_i \alpha_i|$ and $\left[\sum_i |\alpha_i|^2\right]^{1/2}$ are not comparable since sometimes the random signs may cause significant cancellation or blowup, on average they are indeed comparable. From a practical point of view, Khinchin's inequality helps us get rid of (or conversely, obtain) the termwise absolute value signs by randomizing the the sign of α_i.

Remarks:

1) Khinchin's inequality is a special case of Marcinkiewicz-Zygmund inequality (1937) for sum of independent random variables, and more generally, Burkholder-Davis-Gundy inequality (1973) for martingales.

2) The best constants in the Khinchin's inequality are obtained in U. Haagerup, *The best constants in the Khintchine inequality*, Studia Math, 1981.

3) A Rademacher series of the projection operators associated with an unconditional basis can be uniformly bounded, see, for example, D. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, Ann. Prob., 1984.

4) For more introduction to random series, see, for example, J.-P. Kahane, *Some random series of functions.*

When $0 < p < q < \infty$,

$$l^p \subset l^q, L^q(\Omega) \subset L^p(\Omega)$$

So there exist constants $C_p, D_p > 0$, such that

$$\|\epsilon \cdot \alpha\|_{L^p(\Omega)} \leq C_p \|\alpha\|_{l^p}, \forall \alpha \in l^p, 0 < p \leq 2$$

$$\|\alpha\|_{l^p} \leq D_p \|\epsilon \cdot \alpha\|_{L^p(\Omega)}, \forall \alpha \in l^2, 2 \leq p < \infty$$

Moreover, we have

$$\mathbb{P}\left(\omega : \|\alpha\|_{l^\infty} \leq |\epsilon \cdot \alpha|\right) \geq \frac{1}{2}, \forall \alpha \in l^2$$

If we consider a Rademacher series of functions instead of scalars, then by Khinchin's inequality we obtain the following:

Theorem IV : Let $0 < p < \infty, q = p \wedge 2$, (Y, ν) be an arbitrary measure space, then there exists a constant $E_p > 0$, such that

$$\left\| \|\epsilon \cdot f\|_{L^p(\nu)} \right\|_{L^q(\Omega)} \leq E_p \|f\|_{l^q(L^p(\nu))}, \forall f \in l^q_0(L^p(\nu))$$

Proof. If $p \leq 2$, then $q = p$,

$$\begin{aligned}
\text{LHS}^p &= \int_\Omega \int_Y |\sum_i \epsilon_i(\omega) f_i(y)|^p \nu(dy)\mathbb{P}(d\omega) \\
&= \int_Y \int_\Omega |\sum_i \epsilon_i(\omega) f_i(y)|^p \mathbb{P}(d\omega)\nu(dy) \\
&\leq \int_Y C_p^p \sum_i |f_i(y)|^p \nu(dy) \\
&\leq C_p^p \sum_i \int_Y |f_i(y)|^p \nu(dy) \\
&= C_p^p \text{RHS}^p.
\end{aligned}$$

If $p \geq 2$, then $q = 2$,

$$\begin{aligned}
\text{LHS}^2 &= \int_\Omega [\int_Y |\sum_i \epsilon_i(\omega) f_i(y)|^p \nu(dy)]^{2/p} \mathbb{P}(d\omega) \\
&\leq [\int_\Omega \int_Y |\sum_i \epsilon_i(\omega) f_i(y)|^p \nu(dy)\mathbb{P}(d\omega)]^{2/p} \\
&\leq [\int_Y \int_\Omega |\sum_i \epsilon_i(\omega) f_i(y)|^p \mathbb{P}(d\omega)\nu(dy)]^{2/p} \\
&\leq [\int_Y B_p^p (\sum_i |f_i(y)|^2)^{p/2} \nu(dy)]^{2/p} \\
&= B_p^2 [\int_Y (\sum_i |f_i(y)|^2)^{p/2} \nu(dy)]^{2/p} \\
&\leq B_p^2 \sum_i [\int_Y |f_i(x)|^p \nu(dy)]^{2/p} \\
&= B_p^2 \text{RHS}^2. \quad \square
\end{aligned}$$

Now we can verify d) in Theorem II to obtain the following:

Theorem V (**Nikishin′s Factorization Theorem, 1970**) : Let (Y, ν) be an arbitrary measure space, $0 < p < \infty, T : L^p(Y, \nu) \to L^0(m)$ be a continuous linear or sublinear operator, then there exists $0 < w \in L^0(m)$, such that

$$T : L^p(Y, \nu) \to L^{q,\infty}(wm)$$

and is bounded, where $q = p \wedge 2$. Moreover, if T is positive, then we can take $q = p$.

Proof. By Theorem II, it suffices to show that $\exists C(\cdot) : (0, \infty) \to (0, \infty), \lim_{\lambda \to \infty} C(\lambda) = 0$, such that

$$m\{x \in X : \sup_i |T(f_i)(x)| > \lambda\} \leq C(\lambda), \forall (f_i)_{i=1}^\infty \in l_0^q(L^p(\nu)), \sum_i \|f_i\|_{L^p(\nu)}^q \leq 1$$

Note that we change l^p into l^p_0 here, which can be justified easily using limiting argument.

Since $T : L^p(Y, \mu) \to L^0(m)$ is bounded, there exists $c(\cdot) : (0, \infty) \to (0, \infty)$, $\lim_{\lambda\to\infty} c(\lambda) = 0$, such that

$$m\{x \in X : |T(f)(x)| > \lambda\} \leq c(\lambda), \forall f \in L^p(\nu), \|f\|_{L^p(\nu)} \leq 1$$

If T is positive, let $q = p$. Define $F := (\sum_i |f_i|^p)^{1/p}$, then $\|F\|_{L^p(\nu)} \leq 1$ and $|f_i| \leq F, \forall i$. Since T is positive, $|Tf_i| \leq |TF|, \forall i$. Thus

$$m\{x \in X : \sup_i |T(f_i)(x)| > \lambda\} \leq m\{x \in X : |T(F)(x)| > \lambda\} \leq c(\lambda)$$

In the general case, let $q = p \wedge 2$. Let $\{\epsilon_i\}_{i=1}^{\infty}$ be a Rademacher sequence. Fix $\lambda > 0$ and i, let $\epsilon'_j = (-1)^{\delta_{ij}}\epsilon_j$, then a.e. x, we have for all ω,

$$\begin{aligned}
|T(f_i)(x)| &= |T(\epsilon_i(\omega)f_i)(x)| \\
&= \frac{1}{2}|T(2\epsilon_i(\omega)f_i)(x)| \\
&= \frac{1}{2}|T(\sum_j \epsilon_j(\omega)f_j - \sum_j \epsilon'_j(\omega)f_j)(x)| \\
&\leq \frac{1}{2}|T(\sum_j \epsilon_j(\omega)f_j)(x)| + \frac{1}{2}|T(\sum_j \epsilon'_j(\omega)f_j)(x)|
\end{aligned}$$

If $|T(f_i)(x)| > \lambda$, then $\Omega = \{\omega : |T(\sum_j \epsilon_j(\omega)f_j)(x)| > \lambda\} \cup \{\omega : |T(\sum_j \epsilon'_j(\omega)f_j)(x)| > \lambda\}$. Hence

$$1 \leq \mathbb{P}\{\omega : |T(\sum_j \epsilon_j(\omega)f_j)(x)| > \lambda\} + \mathbb{P}\{\omega : |T(\sum_j \epsilon'_j(\omega)f_j)(x)| > \lambda\}$$

Since

$$\mathbb{P}\{\omega : |T(\sum_j \epsilon_j(\omega)f_j)(x)| > \lambda\} = \mathbb{P}\{\omega : |T(\sum_j \epsilon'_j(\omega)f_j)(x)| > \lambda\}$$

We get

$$\mathbb{P}\{\omega : |T(\sum_j \epsilon_j(\omega)f_j)(x)| > \lambda\} \geq \frac{1}{2}$$

Hence for a.e. x,

$$\chi_{\{x:|T(f_i)(x)|>\lambda\}} \leq 2\mathbb{P}\{\omega : |T(\sum_j \epsilon_j(\omega)f_j)(x)| > \lambda\}$$

This holds for every i, so for a.e. x,

$$\chi_{\{x:\sup_i |T(f_i)(x)|>\lambda\}} \leq 2\mathbb{P}\{\omega : |T(\sum_j \epsilon_j(\omega)f_j)(x)| > \lambda\}$$

Integrating over x, we get

$$\begin{aligned}
m\{x : \sup_i |T(f_i)(x)| > \lambda\} &\le 2 \int_X \int_\Omega \chi_{\{|T(\sum_j \epsilon_j(\omega) f_j)(x)| > \lambda\}}(\omega, x) \mathbb{P}(d\omega) m(dx) \\
&= 2 \int_\Omega m\{x : |T(\sum_j \epsilon_j(\omega) f_j)(x)| > \lambda\} \mathbb{P}(d\omega) \\
&\le 2\mathbb{P}\{\omega : \| \sum_j \epsilon_j(\omega) f_j \|_{L^p(\nu)} > \lambda^{1/2}\} \\
&\quad + 2 \int_{\{\| \sum_j \epsilon_j f_j \|_{L^p(\nu)} \le \lambda^{1/2}\}} m\{x : |T(\frac{\sum_j \epsilon_j(\omega) f_j}{\lambda^{1/2}})(x)| > \lambda^{1/2}\} \mathbb{P}(d\omega) \\
&\le \frac{2}{\lambda^{q/2}} \int_\Omega \| \sum_j \epsilon_j(\omega) f_j \|^q_{L^p(\nu)} \mathbb{P}(d\omega) + 2c(\lambda^{1/2}) \\
&\le \frac{2E_p^q}{\lambda^{q/2}} \sum_j \|f_j\|^q_{L^p(\nu)} + 2c(\lambda^{1/2}) \\
&\le \frac{2E_p^q}{\lambda^{q/2}} + 2c(\lambda^{1/2}). \quad \square
\end{aligned}$$

By averaging the conformality, we can now obtain Stein's maximal principle in the case of compact groups.

Let (G, τ) be a topological group. G is called a compact group if G is compact and Hausdorff with respect to τ. In the following we always assume that G is a compact group.

A finite Borel measure m on G is called a Radon measure if $m(A) = \inf_{U \supset A} m(U)$ for all Borel set A, where U is taken over all the open sets containing A. A nonzero Radon measure m on G is called a left Haar measure if $m(xA) = m(A), \forall A, x \in G$. If a left Haar measure is also right invariant, we call it a Haar measure. Left Haar measure always exists and is unique up to a positive multiple. In the following we assume $m(G) = 1$.

Note that the outer regularity assumption here automatically implies inner regularity, and when the topology of G is second countable, any finite Borel measure is automatically a Radon measure. Note also that by uniqueness, any left Haar measure is automatically a Haar measure.

Let $x \in G$, x acts on $L^0(m)$ by $xf(\cdot) := f(x^{-1}\cdot)$. A subspace V of $L^0(m)$ is called translation invariant if V is invariant under the action of G. Let V be translation invariant, $T : V \to L^0(m)$ is called translation invariant if T commutes with the action of G. Note that $L^p(m)$ and $L^{p,\infty}(m)$ are translation invariant, for all $0 < p < \infty$.

Theorem VI (**Stein's Maximal Principle, 1961**) : Let G be a second countable compact group with Haar measure m, let $0 < p < \infty, T : L^p(m) \to L^0(m)$ be a continuous, linear (or sublinear), translation invariant operator, then

$$T : L^p(m) \to L^{q,\infty}(m)$$

and is bounded, where $q = p \wedge 2$. Moreover, when T is positive we can take $q = p$.

Proof. By Theorem V, there exists $w > 0$ such that $T : L^p(m) \to L^{q,\infty}(wm)$ and is bounded, i.e., there exists a constant $C > 0$, such that

$$\int_{\{|T(f)(\cdot)|>\lambda\}} w(x)m(dx) \leq \frac{C^q}{\lambda^q}\|f\|_p^q, \forall\lambda > 0, f \in L^p(m)$$

Now fix any $y \in G$. We show that $w(y\cdot)$ provides the same estimate. In fact,

$$\begin{aligned}
\int_{\{|T(f)(\cdot)|>\lambda\}} w(yx)m(dx) &= \int_G \chi_{\{|T(f)(\cdot)|>\lambda\}}(x)w(yx)m(dx) \\
&= \int_G \chi_{\{|T(f)(\cdot)|>\lambda\}}(y^{-1}x)w(x)m(dx) \\
&= \int_G \chi_{\{|T(f)(y^{-1}\cdot)|>\lambda\}}(x)w(x)m(dx) \\
&= \int_G \chi_{\{|T(yf)(\cdot)|>\lambda\}}(x)w(x)m(dx) \\
&\leq \frac{C^q}{\lambda^q}\|yf\|_p^q = \frac{C^q}{\lambda^q}\|f\|_p^q
\end{aligned}$$

Integrating over y, we get

$$\begin{aligned}
\int_G\int_G \chi_{\{|T(f)(\cdot)|>\lambda\}}(x)w(yx)m(dx)m(dy) &= \int_G \chi_{\{|T(f)(\cdot)|>\lambda\}}(x)(\int_G w(yx)m(dy))m(dx) \\
&= \int_G \chi_{\{|T(f)(\cdot)|>\lambda\}}(x)(\int_G w(y)m(dy))m(dx) \\
&= a\int_G \chi_{\{|T(f)(\cdot)|>\lambda\}}(x)m(dx) \\
&\leq \frac{C^q}{\lambda^q}\|f\|_p^q
\end{aligned}$$

Here $a = \int_G w(y)m(dy) > 0$. The last inequality shows

$$m\{|T(f)(\cdot)| > \lambda\} \leq \frac{C^q}{a\lambda^q}\|f\|_p^q. \quad \square$$

Remark: Here we assume that G is second countable in order to guarantee the change of order of integration. In fact it is still legitimate to do so in the general case, see, for example, Donald Cohn, *Measure Theory* (Theorem 7.6.7).

Reference: José García-Cuerva and J.-L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North-Holland Mathematics Studies, Volume 116, 1985.