

NOTES 1: INTERPOLATION OF L^p SPACES

TERENCE TAO

In the previous two quarters, we have been focusing largely on the “soft” side of real analysis, which is primarily concerned with “qualitative” properties such as convergence, compactness, measurability, and so forth. In contrast, we will begin this quarter with more of an emphasis on the “hard” side of real analysis, in which we study estimates and upper and lower bounds of various quantities, such as norms of functions or operators. (Of course, the two sides of analysis are closely connected to each other; an understanding of both sides and their interrelationships, are needed in order to get the broadest and most complete perspective for this subject.)

One basic tool in hard analysis is that of *interpolation*, which allows one to start with a hypothesis of two (or more) “upper bound” estimates, e.g. $A_0 \leq B_0$ and $A_1 \leq B_1$, and conclude a family of intermediate estimates $A_\theta \leq B_\theta$ (or maybe $A_\theta \leq C_\theta B_\theta$, where C_θ is a constant) for any choice of parameter $0 < \theta < 1$. Of course, interpolation is not a magic wand; one needs various hypotheses (e.g. linearity, sublinearity, convexity, or complexifiability) on A_i, B_i in order for interpolation methods to be applicable. Nevertheless, these techniques are available for many important classes of problems, most notably that of establishing boundedness estimates such as $\|Tf\|_{L^q(Y,\nu)} \leq C\|f\|_{L^p(X,\mu)}$ for linear (or “linear-like”) operators T from one Lebesgue space $L^p(X,\mu)$ to another $L^q(Y,\nu)$. (Interpolation can also be performed for many other normed vector spaces than the Lebesgue spaces, but we will just focus on Lebesgue spaces in these notes to focus the discussion.) Using interpolation, it is possible to reduce the task of proving such estimates to that of proving various “endpoint” versions of these estimates. In some cases, each endpoint only faces a portion of the difficulty that the interpolated estimate did, and so by using interpolation one has split the task of proving the original estimate into two or more simpler subtasks. In other cases, one of the endpoint estimates is very easy, and the other one is significantly more difficult than the original estimate; thus interpolation does not really simplify the task of proving estimates in this case, but at least clarifies the relative difficulty between various estimates in a given family.

As is the case with many other tools in analysis, interpolation is not captured by a single “interpolation theorem”; instead, there are a family of such theorems, which can be broadly divided into two major categories, reflecting the two basic methods that underlie the principle of interpolation. The *real interpolation method* is based on a divide and conquer strategy: to understand how to obtain control on some expression such as $\|Tf\|_{L^q(Y,\nu)}$ for some operator T and some function f , one would divide f into two or more components, e.g. into components where f is large and where f is small, or where f is oscillating with high frequency or only varying with low frequency. Each component would be estimated using a carefully

chosen combination of the extreme estimates available; optimising over these choices and summing up (using whatever linearity-type properties on T are available), one would hope to get a good estimate on the original expression. The strengths of the real interpolation method are that the linearity hypotheses on T can be relaxed to weaker hypotheses, such as sublinearity or quasilinearity; also, the endpoint estimates are allowed to be of a weaker “type” than the interpolated estimates. On the other hand, the real interpolation often concedes a multiplicative constant in the final estimates obtained, and one is usually obligated to keep the operator T fixed throughout the interpolation process. The proofs of real interpolation theorems are also a little bit messy, though in many cases one can simply invoke a standard instance of such theorems (e.g. the Marcinkiewicz interpolation theorem) as a black box in applications.

The *complex interpolation method* instead proceeds by exploiting the powerful tools of complex analysis, in particular the *maximum modulus principle* and its relatives (such as the *Phragmén-Lindelöf principle*). The idea is to rewrite the estimate to be proven (e.g. $\|Tf\|_{L^q(Y,\nu)} \leq C\|f\|_{L^p(X,\mu)}$) in such a way that it can be embedded into a family of such estimates which depend holomorphically on a complex parameter s in some domain (e.g. the strip $\{\sigma + it : t \in \mathbf{R}, \sigma \in [0, 1]\}$). One then exploits things like the maximum modulus principle to bound an estimate corresponding to an interior point of this domain by the estimates on the boundary of this domain. The strengths of the complex interpolation method are that it typically gives cleaner constants than the real interpolation method, and also allows the underlying operator T to vary holomorphically with respect to the parameter s , which can significantly increase the flexibility of the interpolation technique. The proofs of these methods are also very short (if one takes the maximum modulus principle and its relatives as a black box), which make the method particularly amenable for generalisation to more intricate settings (e.g. multilinear operators, mixed Lebesgue norms, etc.). On the other hand, the somewhat rigid requirement of holomorphicity makes it much more difficult to apply this method to non-linear operators, such as sublinear or quasilinear operators; also, the interpolated estimate tends to be of the same “type” as the extreme ones, so that one does not enjoy the upgrading of weak type estimates to strong type estimates that the real interpolation method typically produces. Also, the complex method runs into some minor technical problems when target space $L^q(Y, \nu)$ ceases to be a Banach space (i.e. when $q < 1$) as this makes it more difficult to exploit duality.

Despite these differences, the real and complex methods tend to give broadly similar results in practice, especially if one is willing to ignore constant losses in the estimates or epsilon losses in the exponents.

The theory of both real and complex interpolation can be studied abstractly, in general normed or quasi-normed spaces; see e.g. *this book* for a detailed treatment. However in these notes we shall focus exclusively on interpolation for Lebesgue spaces L^p (and their cousins, such as the weak Lebesgue spaces $L^{p,\infty}$ and the Lorentz spaces $L^{p,r}$).

1. INTERPOLATION OF SCALARS

As discussed in the introduction, most of the interesting applications of interpolation occur when the technique is applied to operators T . However, in order to gain some intuition as to why interpolation works in the first place, let us first consider the significantly simpler (though rather trivial) case of interpolation in the case of scalars or functions.

We begin first with scalars. Suppose that A_0, B_0, A_1, B_1 are non-negative real numbers such that

$$A_0 \leq B_0 \tag{1}$$

and

$$A_1 \leq B_1. \tag{2}$$

Then clearly we will have

$$A_\theta \leq B_\theta \tag{3}$$

for every $0 \leq \theta \leq 1$, where we define

$$A_\theta := A_0^{1-\theta} A_1^\theta \tag{4}$$

and

$$B_\theta := B_0^{1-\theta} B_1^\theta; \tag{5}$$

indeed one simply raises (1) to the power $1 - \theta$, (2) to the power θ , and multiplies the two inequalities together. Thus for instance, when $\theta = 1/2$ one obtains the geometric mean of (1) and (2):

$$A_0^{1/2} A_1^{1/2} \leq B_0^{1/2} B_1^{1/2}.$$

One can view A_θ and B_θ as the unique log-linear functions of θ (i.e. $\log A_\theta, \log B_\theta$ are (affine-)linear functions of θ) which equal their boundary values A_0, A_1 and B_0, B_1 respectively as $\theta = 0, 1$.

Example 1. If $A_0 = AL^{1/p_0}$ and $A_1 = AL^{1/p_1}$ for some $A, L > 0$ and $0 < p_0, p_1 \leq \infty$, then the log-linear interpolant A_θ is given by $A_\theta = AL^{1/p_\theta}$, where $0 < p_\theta \leq \infty$ is the quantity such that $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

The deduction of (3) from (1), (2) is utterly trivial, but there are still some useful lessons to be drawn from it. For instance, let us take $A_0 = A_1 = A$ for simplicity, so we are interpolating two upper bounds $A \leq B_0, A \leq B_1$ on the same quantity A to give a new bound $A \leq B_\theta$. But actually we have a refinement available to this bound, namely

$$A_\theta \leq B_\theta \min \left(\frac{B_0}{B_1}, \frac{B_1}{B_0} \right)^\varepsilon \tag{6}$$

for any sufficiently small $\varepsilon > 0$ (indeed one can take any ε less than or equal to $\min(\theta, 1-\theta)$). Indeed one sees this simply by applying (3) with θ with $\theta - \varepsilon$ and $\theta + \varepsilon$ and taking minima. Thus we see that (3) is only sharp when the two original bounds B_0, B_1 are comparable; if instead we have $B_1 \sim 2^n B_0$ for some integer n , then (6) tells us that we can improve (3) by an exponentially decaying factor of $2^{-\varepsilon|n|}$. The geometric series formula tells us that such factors are absolutely summable, and so

in practice it is often a useful heuristic to pretend that the $n = O(1)$ cases dominate so strongly that the other cases can be viewed as negligible by comparison.

Also, one can trivially extend the deduction of (3) from (1), (2) as follows: if $\theta \rightarrow A_\theta$ is a function from $[0, 1]$ to \mathbf{R}^+ which is log-convex (thus $\theta \mapsto \log A_\theta$ is a convex function of θ , and (1), (2) hold for some $B_0, B_1 > 0$, then (3) holds for all intermediate θ also, where B_θ is of course defined by (5). Thus one can interpolate upper bounds on log-convex functions. However, one certainly cannot interpolate lower bounds: lower bounds on a log-convex function $\theta \rightarrow A_\theta$ at $\theta = 0$ and $\theta = 1$ yield no information about the value of, say, $A_{1/2}$. Similarly, one cannot extrapolate upper bounds on log-convex functions: an upper bound on, say, A_0 and $A_{1/2}$ does not give any information about A_1 . (However, an upper bound on A_0 coupled with a *lower bound* on $A_{1/2}$ gives a lower bound on A_1 ; this is the contrapositive of an interpolation statement.)

Exercise 2. Show that the sum $f+g$, product fg , or pointwise maximum $\max(f, g)$ of two log-convex functions $f, g : [0, 1] \rightarrow \mathbf{R}^+$ is log-convex.

Remark 3. Every non-negative log-convex function $\theta \mapsto A_\theta$ is convex, thus in particular $A_\theta \leq (1 - \theta)A_0 + \theta A_1$ for all $0 \leq \theta \leq 1$ (note that this generalises the arithmetic mean-geometric mean inequality). Of course, the converse statement is not true.

Now we turn to the complex version of the interpolation of log-convex functions, a result known as *Lindelöf's theorem*:

Theorem 4. (*Lindelöf's theorem*) Let $s \mapsto f(s)$ be a holomorphic function on the strip $S := \{\sigma + it : 0 \leq \sigma \leq 1; t \in \mathbf{R}\}$, which obeys the bound

$$|f(\sigma + it)| \leq A \exp(\exp((\pi - \delta)|t|)) \quad (7)$$

for all $\sigma + it \in S$ and some constants $A, \delta > 0$. Suppose also that $|f(0 + it)| \leq B_0$ and $|f(1 + it)| \leq B_1$ for all $t \in \mathbf{R}$. Then we have $|f(\theta + it)| \leq B_\theta$ for all $0 \leq \theta \leq 1$ and $t \in \mathbf{R}$, where B_θ is of course defined by (5).

Remark 5. The hypothesis (7) is a qualitative hypothesis rather than a quantitative one, since the exact values of A, δ do not show up in the conclusion. It is quite a mild condition; any function of exponential growth in t , or even with such super-exponential growth as $O(|t|^{|t|})$ or $O(e^{|t|^{O(1)}})$, will obey (7). The principle however fails without this hypothesis, as one can see for instance by considering the holomorphic function $f(s) := \exp(-i \exp(\pi i s))$.

Proof. Observe that the function $s \mapsto B_0^{1-s} B_1^s$ is holomorphic and non-zero on S , and has magnitude exactly B_θ on the line $\operatorname{Re}(s) = \theta$ for each $0 \leq \theta \leq 1$. Thus, by dividing f by this function (which worsens the qualitative bound (7) slightly) we may reduce to the case when $B_\theta = 1$ for all $0 \leq \theta \leq 1$.

Suppose we temporarily assume that $f(\sigma + it) \rightarrow 0$ as $|\sigma + it| \rightarrow \infty$. Then by the maximum modulus principle (applied to a sufficiently large rectangular portion of the strip), it must then attain a maximum on one of the two sides of the strip. But $|f| \leq 1$ on these two sides, and so $|f| \leq 1$ on the interior as well.

To remove the assumption that f goes to zero at infinity, we use the *trick of giving ourselves an epsilon of room*. Namely, we multiply $f(s)$ by the holomorphic function $g_\varepsilon(s) := \exp(\varepsilon i \exp(i[(\pi - \delta/2)s + \delta/4]))$ for some $\varepsilon > 0$. A little complex arithmetic shows that the function $f(s)g_\varepsilon(s)g_\varepsilon(1-s)$ goes to zero at infinity in S (the $g_\varepsilon(s)$ factor decays fast enough to damp out the growth of f as $\text{Im}(s) \rightarrow -\infty$, while the $g_\varepsilon(1-s)$ damps out the growth as $\text{Im}(s) \rightarrow +\infty$), and is bounded in magnitude by 1 on both sides of the strip S . Applying the previous case to this function, then taking limits as $\varepsilon \rightarrow 0$, we obtain the claim. \square

Exercise 6. *With the notation and hypotheses of Theorem 4, show that the function $\sigma \mapsto \sup_{t \in \mathbf{R}} |f(\sigma + it)|$ is log-convex on $[0, 1]$.*

Exercise 7. *(Hadamard three-circles theorem) Let f be a holomorphic function on an annulus $\{z \in \mathbf{C} : R_1 \leq |z| \leq R_2\}$. Show that the function $r \mapsto \sup_{\theta \in [0, 2\pi]} |f(re^{i\theta})|$ is log-convex on $[R_1, R_2]$.*

Exercise 8. *(Phragmén-Lindelöf principle) Let f be as in Theorem 4, but suppose that we have the bounds $f(0 + it) \leq C(1 + |t|)^{a_0}$ and $f(1 + it) \leq C(1 + |t|)^{a_1}$ for all $t \in \mathbf{R}$ and some exponents $a_0, a_1 \in \mathbf{R}$ and a constant $C > 0$. Show that one has $f(\sigma + it) \leq C'(1 + |t|)^{(1-\sigma)a_0 + \sigma a_1}$ for all $\sigma + it \in S$ and some constant C' (which is allowed to depend on the constants A, δ in (7)). (Hint: it is convenient to work first in a half-strip such as $\{\sigma + it \in S : t \geq T\}$ for some large T . Then multiply f by something like $\exp(-((1-z)a_0 + za_1)\log(-iz))$ for some suitable branch of the logarithm and apply a variant of Theorem 4 for the half-strip. A more refined estimate in this regard is due to Rademacher.) This particular version of the principle gives the convexity bound for Dirichlet series such as the Riemann zeta function. Bounds which exploit the deeper properties of these functions to improve upon the convexity bound are known as subconvexity bounds and are of major importance in analytic number theory, which is of course well outside the scope of this course.*

2. INTERPOLATION OF FUNCTIONS

We now turn to the interpolation in function spaces, focusing particularly on the Lebesgue spaces $L^p(X)$ and the weak Lebesgue spaces $L^{p,\infty}(X)$. Here, $X = (X, \mathcal{X}, \mu)$ is a fixed measure space. It will not matter much whether we deal with real or complex spaces; for sake of concreteness we work with complex spaces. Then for $0 < p < \infty$, recall (see 245B Notes 3) that $L^p(X)$ is the space of all functions $f : X \rightarrow \mathbf{C}$ whose L^p norm

$$\|f\|_{L^p(X)} := \left(\int_X |f|^p d\mu \right)^{1/p}$$

is finite, modulo almost everywhere equivalence. The space $L^\infty(X)$ is defined similarly, but where $\|f\|_{L^\infty(X)}$ is the essential supremum of $|f|$ on X .

A simple test case in which to understand the L^p norms better is that of a *step function* $f = A1_E$, where A is a non-negative number and E a set of finite measure. Then one has $\|f\|_{L^p(X)} = A\mu(E)^{1/p}$ for $0 < p \leq \infty$. Observe that this is a log-convex function of $1/p$. This is a general phenomenon:

Lemma 9. (*Log-convexity of L^p norms*) Let that $0 < p_0 < p_1 \leq \infty$ and $f \in L^{p_0}(X) \cap L^{p_1}(X)$. Then $f \in L^p(X)$ for all $p_0 \leq p \leq p_1$, and furthermore we have

$$\|f\|_{L^{p_\theta}(X)} \leq \|f\|_{L^{p_0}(X)}^{1-\theta} \|f\|_{L^{p_1}(X)}^\theta$$

for all $0 \leq \theta \leq 1$, where the exponent p_θ is defined by $1/p_\theta := (1-\theta)/p_0 + \theta/p_1$. In particular, we see that the function $1/p \mapsto \|f\|_{L^p(X)}$ is log-convex whenever the right-hand side is finite (and is in fact log-convex for all $0 \leq 1/p < \infty$, if one extends the definition of log-convexity to functions that can take the value $+\infty$). In other words, we can interpolate any two bounds $\|f\|_{L^{p_0}(X)} \leq B_0$ and $\|f\|_{L^{p_1}(X)} \leq B_1$ to obtain $\|f\|_{L^{p_\theta}(X)} \leq B_\theta$ for all $0 \leq \theta \leq 1$.

Let us give several proofs of this lemma. We will focus on the case $p_1 < \infty$; the endpoint case $p_1 = \infty$ can be proven directly, or by modifying the arguments below, or by using an appropriate limiting argument, and we leave the details to the reader.

The first proof is to use *Hölder's inequality*

$$\|f\|_{L^{p_\theta}(X)}^{p_\theta} = \int_X |f|^{(1-\theta)p_\theta} |f|^{\theta p_\theta} d\mu \leq \| |f|^{(1-\theta)p_\theta} \|_{L^{p_0/(1-\theta)p_\theta}} \| |f|^{\theta p_\theta} \|_{L^{p_1/(\theta p_\theta)}}$$

when p_1 is finite (with some minor modifications in the case $p_1 = \infty$).

Another (closely related) proof proceeds by using the log-convexity inequality

$$|f(x)|^{p_\theta} \leq (1-\alpha)|f(x)|^{p_0} + \alpha|f(x)|^{p_1}$$

for all x , where $0 < \alpha < 1$ is the quantity such that $p_\theta = (1-\alpha)p_0 + \alpha p_1$. If one integrates this inequality in x , one already obtains the claim in the normalised case when $\|f\|_{L^{p_0}(X)} = \|f\|_{L^{p_1}(X)} = 1$. To obtain the general case, one can multiply the function f and the measure μ by appropriately chosen constants to obtain the above normalisation; we leave the details as an exercise to the reader. (The case when $\|f\|_{L^{p_0}(X)}$ or $\|f\|_{L^{p_1}(X)}$ vanishes is of course easy to handle separately.)

A third approach is more in the spirit of the real interpolation method, avoiding the use of convexity arguments. As in the second proof, we can reduce to the normalised case $\|f\|_{L^{p_0}(X)} = \|f\|_{L^{p_1}(X)} = 1$. We then split $f = f1_{|f| \leq 1} + f1_{|f| > 1}$, where $1_{|f| \leq 1}$ is the indicator function to the set $\{x : |f(x)| \leq 1\}$, and similarly for $1_{|f| > 1}$. Observe that

$$\|f1_{|f| \leq 1}\|_{L^{p_\theta}(X)}^{p_\theta} = \int_{|f| \leq 1} |f|^{p_\theta} d\mu \leq \int_X |f|^{p_0} d\mu = 1$$

and similarly

$$\|f1_{|f| > 1}\|_{L^{p_\theta}(X)}^{p_\theta} = \int_{|f| > 1} |f|^{p_\theta} d\mu \leq \int_X |f|^{p_1} d\mu = 1$$

and so by the quasi-triangle inequality (or triangle inequality, when $p_\theta \geq 1$)

$$\|f\|_{L^{p_\theta}(X)} \leq C$$

for some constant C depending on p_θ . Note, by the way, that this argument gives the inclusions

$$L^{p_0}(X) \cap L^{p_1}(X) \subset L^{p_\theta}(X) \subset L^{p_0}(X) + L^{p_1}(X). \quad (8)$$

This is off by a constant factor by what we want. But one can eliminate this constant by using the *tensor power trick*. Indeed, if one replaces X with a Cartesian power X^M (with the product σ -algebra \mathcal{X}^M and product measure μ^M), and replace f by the tensor power $f^{\otimes M} : (x_1, \dots, x_M) \mapsto f(x_1) \dots f(x_M)$, we see from many applications of the Fubini-Tonelli theorem that

$$\|f^{\otimes M}\|_{L^p(X^M)} = \|f\|_{L^p(X)}^M$$

for all p . In particular, $f^{\otimes M}$ obeys the same normalisation hypotheses as f , and thus by applying the previous inequality to $f^{\otimes M}$, we obtain

$$\|f\|_{L^{p_\theta}(X)}^M \leq C$$

for every M , where it is key to note that the constant C on the right is independent of M . Taking M^{th} roots and then sending $M \rightarrow \infty$, we obtain the claim.

Finally, we give a fourth proof in the spirit of the complex interpolation method. By replacing f by $|f|$ we may assume f is non-negative. By expressing non-negative measurable functions as the monotone limit of simple functions and using the monotone convergence theorem, we may assume that f is a simple function, which is then necessarily of finite measure support from the L^p finiteness hypotheses. Now consider the function $s \mapsto \int_X |f|^{(1-s)p_0 + sp_1} d\mu$. Expanding f out in terms of step functions we see that this is an analytic function of f which grows at most exponentially in s ; also, by the triangle inequality this function has magnitude at most $\int_X |f|^{p_0}$ when $s = 0 + it$ and magnitude $\int_X |f|^{p_1}$ when $s = 1 + it$. Applying Theorem 4 and specialising to the value of s for which $(1-s)p_0 + sp_1 = p_\theta$ we obtain the claim.

Exercise 10. If $0 < \theta < 1$, show that equality holds in Lemma 9 if and only if $|f|$ is a step function.

Now we consider variants of interpolation in which the “strong” L^p spaces are replaced by their “weak” counterparts $L^{p,\infty}$. Given a measurable function $f : X \rightarrow \mathbf{C}$, we define the *distribution function* $\lambda_f : \mathbf{R}^+ \rightarrow [0, +\infty]$ by the formula

$$\lambda_f(t) := \mu(\{x \in X : |f(x)| \geq t\}) = \int_X 1_{|f| \geq t} d\mu.$$

This distribution function is closely connected to the L^p norms. Indeed, from the calculus identity

$$|f(x)|^p = p \int_0^\infty 1_{|f(x)| \geq t} t^{p-1} dt$$

and the Fubini-Tonelli theorem, we obtain the formula

$$\|f\|_{L^p(X)}^p = p \int_0^\infty \lambda_f(t) t^{p-1} dt \quad (9)$$

for all $0 < p < \infty$, thus the L^p norms are essentially moments of the distribution function. The L^∞ norm is of course related to the distribution function by the formula

$$\|f\|_{L^\infty(X)} = \inf\{t \geq 0 : \lambda_f(t) = 0\}.$$

Exercise 11. Show that we have the relationship

$$\|f\|_{L^p(X)}^p \sim_p \sum_{n \in \mathbf{Z}} \lambda_f(2^n) 2^{np}$$

for any measurable $f : X \rightarrow \mathbf{C}$ and $0 < p < \infty$, where we use $X \sim_p Y$ to denote a pair of inequalities of the form $c_p Y \leq X \leq C_p Y$ for some constants $c_p, C_p > 0$ depending only on p . (Hint: $\lambda_f(t)$ is non-increasing in t .) Thus we can relate the L^p norms of f to the dyadic values $\lambda_f(2^n)$ of the distribution function; indeed, for any $0 < p \leq \infty$, $\|f\|_{L^p(X)}$ is comparable (up to constant factors depending on p) to the $\ell^p(\mathbf{Z})$ norm of the sequence $n \mapsto 2^n \lambda_f(2^n)^{1/p}$.

Another relationship between the L^p norms and the distribution function is given by observing that

$$\|f\|_{L^p(X)}^p = \int_X |f|^p d\mu \geq \int_{|f| \geq t} t^p d\mu = t^p \lambda_f(t)$$

for any $t > 0$, leading to *Chebyshev's inequality*

$$\lambda_f(t) \leq \frac{1}{t^p} \|f\|_{L^p(X)}^p.$$

(The $p = 1$ version of this inequality is also known as *Markov's inequality*. In probability theory, Chebyshev's inequality is often specialised to the case $p = 2$, and with f replaced by a normalised function $f - \mathbf{E} f$. Note that, as with many other Cyrillic names, there are also a large number of alternative spellings of Chebyshev in the Roman alphabet.)

Chebyshev's inequality motivates one to define the *weak L^p norm* $\|f\|_{L^{p,\infty}(X)}$ of a measurable function $f : X \rightarrow \mathbf{C}$ for $0 < p < \infty$ by the formula

$$\|f\|_{L^{p,\infty}(X)} := \sup_{t > 0} t \lambda_f(t)^{1/p},$$

thus Chebyshev's inequality can be expressed succinctly as

$$\|f\|_{L^{p,\infty}(X)} \leq \|f\|_{L^p(X)}.$$

It is also natural to adopt the convention that $\|f\|_{L^{\infty,\infty}(X)} = \|f\|_{L^\infty(X)}$. If $f, g : X \rightarrow \mathbf{C}$ are two functions, we have the inclusion

$$\{|f + g| \geq t\} \subset \{|f| \geq t/2\} \cup \{|g| \geq t/2\}$$

and hence

$$\lambda_{f+g}(t) \leq \lambda_f(t/2) + \lambda_g(t/2);$$

this easily leads to the quasi-triangle inequality

$$\|f + g\|_{L^{p,\infty}(X)} \lesssim_p \|f\|_{L^{p,\infty}(X)} + \|g\|_{L^{p,\infty}(X)}$$

where we use $X \lesssim_p Y$ as shorthand for the inequality $X \leq C_p Y$ for some constant C_p depending only on p (it can be a different constant at each use of the \lesssim_p notation). [Note: in analytic number theory, it is more customary to use \ll_p instead of \lesssim_p , following Vinogradov. However, in analysis \ll is sometimes used instead to denote “much smaller than”, e.g. $X \ll Y$ denotes the assertion $X \leq cY$ for some sufficiently small constant c .]

Let $L^{p,\infty}(X)$ be the space of all $f : X \rightarrow \mathbf{C}$ which have finite $L^{p,\infty}(X)$, modulo almost everywhere equivalence; this space is also known as weak $L^p(X)$. The quasi-triangle inequality soon implies that $L^{p,\infty}(X)$ is a quasi-normed vector space with the $L^{p,\infty}(X)$ quasi-norm, and Chebyshev's inequality asserts that $L^{p,\infty}(X)$ contains $L^p(X)$ as a subspace (though the L^p norm is not a restriction of the $L^{p,\infty}(X)$ norm).

Example 12. If $X = \mathbf{R}^n$ with the usual measure, and $0 < p < \infty$, then the function $f(x) := |x|^{-n/p}$ is in weak L^p , but not strong L^p . It is also not in strong or weak L^q for any other q . But the “local” component $|x|^{-n/p} 1_{|x| \leq 1}$ of f is in strong and weak L^q for all $q < p$, and the “global” component $|x|^{-n/p} 1_{|x| > 1}$ of f is in strong and weak L^q for all $q > p$.

Exercise 13. For any $0 < p, q \leq \infty$ and $f : X \rightarrow \mathbf{C}$, define the (dyadic) Lorentz norm $\|f\|_{L^{p,q}(X)}$ to be $\ell^q(\mathbf{Z})$ norm of the sequence $n \mapsto 2^n \lambda_f(2^n)^{1/p}$, and define the Lorentz space $L^{p,q}(X)$ be the space of functions f with $\|f\|_{L^{p,q}(X)}$ finite, modulo almost everywhere equivalence. Show that $L^{p,q}(X)$ is a quasi-normed space, which is equivalent to $L^{p,\infty}(X)$ when $q = \infty$ and to $L^p(X)$ when $q = p$. Lorentz spaces arise naturally in more refined applications of the real interpolation method, and are useful in certain “endpoint” estimates that fail for Lebesgue spaces, but which can be rescued by using Lorentz spaces instead. However, we will not pursue these applications in detail here.

Exercise 14. Let X be a finite set with counting measure, and let $f : X \rightarrow \mathbf{C}$ be a function. For any $0 < p < \infty$, show that

$$\|f\|_{L^{p,\infty}(X)} \leq \|f\|_{L^p(X)} \lesssim_p \log(1 + |X|) \|f\|_{L^{p,\infty}(X)}.$$

(Hint: to prove the second inequality, normalise $\|f\|_{L^{p,\infty}(X)} = 1$, and then manually dispose of the regions of X where f is too large or too small.) Thus, in some sense, weak L^p and strong L^p are equivalent “up to logarithmic factors”.

One can interpolate weak L^p bounds just as one can strong L^p bounds: if $\|f\|_{L^{p_0,\infty}(X)} \leq B_0$ and $\|f\|_{L^{p_1,\infty}(X)} \leq B_1$, then

$$\|f\|_{L^{p_\theta,\infty}(X)} \leq B_\theta \tag{10}$$

for all $0 \leq \theta \leq 1$. Indeed, from the hypotheses we have

$$\lambda_f(t) \leq \frac{B_0^{p_0}}{t^{p_0}}$$

and

$$\lambda_f(t) \leq \frac{B_1^{p_1}}{t^{p_1}}$$

for all $t > 0$, and hence by scalar interpolation (using an interpolation parameter $0 < \alpha < 1$ defined by $p_\theta = (1 - \alpha)p_0 + \alpha p_1$, and after doing some algebra) we have

$$\lambda_f(t) \leq \frac{B_\theta^{p_\theta}}{t^{p_\theta}} \tag{11}$$

for all $0 < \theta < 1$.

As remarked in the previous section, we can improve upon (11); indeed, if we define t_0 to be the unique value of t where $B_0^{p_0}/t^{p_0}$ and $B_1^{p_1}/t^{p_1}$ are equal, then we have

$$\lambda_f(t) \leq \frac{B_\theta^{p_\theta}}{t^{p_\theta}} \min(t/t_0, t_0/t)^\varepsilon$$

for some $\varepsilon > 0$ depending on p_0, p_1, θ . Inserting this improved bound into (9) we see that we can improve the weak-type bound (10) to a strong-type bound

$$\|f\|_{L^{p_\theta}(X)} \leq C_{p_0, p_1, \theta} B_\theta \quad (12)$$

for some constant $C_{p_0, p_1, \theta}$. Note that one cannot use the tensor power trick this time to eliminate the constant $C_{p_0, p_1, \theta}$ as the weak L^p norms do not behave well with respect to tensor product. Indeed, the constant $C_{p_0, p_1, \theta}$ must diverge to infinity in the limit $\theta \rightarrow 0$ if $p_0 \neq \infty$, otherwise it would imply that the L^{p_0} norm is controlled by the $L^{p_0, \infty}$ norm, which is false by Example 12; similarly one must have a divergence as $\theta \rightarrow 1$ if $p_1 \neq \infty$.

Exercise 15. Let $0 < p_0 < p_1 \leq \infty$ and $0 < \theta < 1$. Refine the inclusions in (8) to

$$\begin{aligned} L^{p_0}(X) \cap L^{p_1}(X) &\subset L^{p_0, \infty}(X) \cap L^{p_1, \infty}(X) \subset L^{p_\theta}(X) \subset \\ &\subset L^{p_\theta, \infty}(X) \subset L^{p_0}(X) + L^{p_1}(X) \subset L^{p_0, \infty}(X) + L^{p_1, \infty}(X). \end{aligned}$$

Define the *strong type diagram* of a function $f : X \rightarrow \mathbf{C}$ to be the set of all $1/p$ for which f lies in strong L^p , and the *weak type diagram* to be the set of all $1/p$ for which f lies in weak L^p . Then both the strong and weak type diagrams are connected subsets of $[0, +\infty)$, and the strong type diagram is contained in the weak type diagram, and contains in turn the interior of the weak type diagram. By experimenting with linear combinations of the examples in Example 12 we see that this is basically everything one can say about the strong and weak type diagrams, without further information on f or X .

Exercise 16. Let $f : X \rightarrow \mathbf{C}$ be a measurable function which is finite almost everywhere. Show that there exists a unique non-increasing left-continuous function $f^* : \mathbf{R}^+ \rightarrow [0, +\infty]$ such that $\lambda_{f^*}(t) = \lambda_f(t)$ for all $t > 0$, and in particular $\|f\|_{L^p(X)} = \|f^*\|_{L^p(\mathbf{R}^+)}$ for all $0 < p \leq \infty$, and $\|f\|_{L^{p, \infty}(X)} = \|f^*\|_{L^{p, \infty}(\mathbf{R}^+)}$. (Hint: first look for the formula that describes $f^*(x)$ for some $x > 0$ in terms of $\lambda_f(t)$.) The function f^* is known as the non-increasing rearrangement of f , and the spaces $L^p(X)$ and $L^{p, \infty}(X)$ are examples of rearrangement-invariant spaces. There are a class of useful rearrangement inequalities that relate f to its rearrangements, and which can be used to clarify the structure of rearrangement-invariant spaces, but we will not pursue this topic here.

Exercise 17. Let (X, \mathcal{X}, μ) be a σ -finite measure space, let $1 < p < \infty$, and $f : X \rightarrow \mathbf{C}$ be a measurable function. Show that the following are equivalent:

- f lies in $L^{p, \infty}(X)$, thus $\|f\|_{L^{p, \infty}(X)} \leq C$ for some finite C .
- There exists a constant C' such that $|\int_X f 1_E d\mu| \leq C' \mu(E)^{1/p'}$ for all sets E of finite measure.

Furthermore show that the best constants C, C' in the above statements are equivalent up to multiplicative constants depending on p , thus $C \sim_p C'$. Conclude that the modified weak $L^{p,\infty}(X)$ norm $\|f\|_{\tilde{L}^{p,\infty}(X)} := \sup_E \mu(E)^{-1/p'} |\int_X f 1_E d\mu|$, where E ranges over all sets of positive finite measure, is a genuine norm on $L^{p,\infty}(X)$ which is equivalent to the $L^{p,\infty}(X)$ quasinorm.

Exercise 18. Let $n > 1$ be an integer. Find a probability space (X, \mathcal{X}, μ) and functions $f_1, \dots, f_n : X \rightarrow \mathbf{R}$ with $\|f_j\|_{L^{1,\infty}(X)} \leq 1$ for $j = 1, \dots, n$ such that $\|\sum_{j=1}^n f_j\|_{L^{1,\infty}(X)} \geq cn \log n$ for some absolute constant $c > 0$. (Hint: exploit the logarithmic divergence of the harmonic series $\sum_{j=1}^\infty \frac{1}{j}$.) Conclude that there exists a probability space X such that the $L^{1,\infty}(X)$ quasi-norm is not equivalent to an actual norm.

Exercise 19. Let (X, \mathcal{X}, μ) be a σ -finite measure space, let $0 < p < \infty$, and $f : X \rightarrow \mathbf{C}$ be a measurable function. Show that the following are equivalent:

- f lies in $L^{p,\infty}(X)$.
- There exists a constant C such that for every set E of finite measure, there exists a subset E' with $\mu(E') \geq \frac{1}{2}\mu(E)$ such that $|\int_X f 1_{E'} d\mu| \leq C\mu(E)^{1/p'}$.

Exercise 20. Let (X, \mathcal{X}, μ) be a measure space of finite measure, and $f : X \rightarrow \mathbf{C}$ be a measurable function. Show that the following two statements are equivalent:

- There exists a constant $C > 0$ such that $\|f\|_{L^p(X)} \leq Cp$ for all $1 \leq p < \infty$.
- There exists a constant $c > 0$ such that $\int_X e^{c|f|} d\mu < \infty$.

3. INTERPOLATION OF OPERATORS

We turn at last to the central topic of these notes, which is interpolation of operators T between functions on two fixed measure spaces $X = (X, \mathcal{X}, \mu)$ and $Y = (Y, \mathcal{Y}, \nu)$. To avoid some (very minor) technicalities we will make the mild assumption throughout that X and Y are both σ -finite, although much of the theory here extends to the non- σ -finite setting.

A typical situation is that of a linear operator T which maps one $L^{p_0}(X)$ space to another $L^{q_0}(Y)$, and also maps $L^{p_1}(X)$ to $L^{q_1}(Y)$ for some exponents $0 < p_0, p_1, q_0, q_1 \leq \infty$; thus (by linearity) T will map the larger vector space $L^{p_0}(X) + L^{p_1}(X)$ to $L^{q_0}(Y) + L^{q_1}(Y)$, and one has some estimates of the form

$$\|Tf\|_{L^{q_0}(Y)} \leq B_0 \|f\|_{L^{p_0}(X)} \quad (13)$$

and

$$\|Tf\|_{L^{q_1}(Y)} \leq B_1 \|f\|_{L^{p_1}(X)} \quad (14)$$

for all $f \in L^{p_0}(X), f \in L^{p_1}(X)$ respectively, and some $B_0, B_1 > 0$. We would like to then interpolate to say something about how T maps $L^{p_\theta}(X)$ to $L^{q_\theta}(Y)$.

The complex interpolation method gives a satisfactory result as long as the exponents allow one to use duality methods, a result known as the *Riesz-Thorin theorem*:

Theorem 21. (*Riesz-Thorin theorem*) Let $0 < p_0, p_1 \leq \infty$ and $1 \leq q_0, q_1 \leq \infty$. Let $T : L^{p_0}(X) + L^{p_1}(X) \rightarrow L^{q_0}(Y) + L^{q_1}(Y)$ be a linear operator obeying the bounds

(13), (14) for all $f \in L^{p_0}(X), f \in L^{p_1}(X)$ respectively, and some $B_0, B_1 > 0$. Then we have

$$\|Tf\|_{L^{q_\theta}(Y)} \leq B_\theta \|f\|_{L^{p_\theta}(X)}$$

for all $0 < \theta < 1$ and $f \in L^{p_\theta}(X)$, where $1/p_\theta := (1-\theta)/p_0 + \theta/p_1$, $1/q_\theta := (1-\theta)/q_0 + \theta/q_1$, and $B_\theta := B_0^{1-\theta} B_1^\theta$.

Remark 22. When X is a point, this theorem essentially collapses to Lemma 9 (and when Y is a point, this is a dual formulation of that lemma); and when X and Y are both points; this collapses to interpolation of scalars.

Proof. If $p_0 = p_1$ then the claim follows from Lemma 9, so we may assume $p_0 \neq p_1$, which in particular forces p_θ to be finite. By symmetry we can take $p_0 < p_1$. By multiplying the measures μ and ν (or the operator T) by various constants, we can normalise $B_0 = B_1 = 1$ (the case when $B_0 = 0$ or $B_1 = 0$ is trivial). Thus we have $B_\theta = 1$ also.

By Hölder's inequality, the bound (13) implies that

$$\left| \int_Y (Tf)g \, d\nu \right| \leq \|f\|_{L^{p_0}(X)} \|g\|_{L^{q'_0}(Y)} \quad (15)$$

for all $f \in L^{p_0}(X)$ and $g \in L^{q'_0}(Y)$, where q'_0 is the dual exponent of q_0 . Similarly we have

$$\left| \int_Y (Tf)g \, d\nu \right| \leq \|f\|_{L^{p_1}(X)} \|g\|_{L^{q'_1}(Y)} \quad (16)$$

for all $f \in L^{p_1}(X)$ and $g \in L^{q'_1}(Y)$.

We now claim that

$$\left| \int_Y (Tf)g \, d\nu \right| \leq \|f\|_{L^{p_\theta}(X)} \|g\|_{L^{q'_\theta}(Y)} \quad (17)$$

for all f, g that are simple functions with finite measure support. To see this, we first normalise $\|f\|_{L^{p_\theta}(X)} = \|g\|_{L^{q'_\theta}(Y)} = 1$. Observe that we can write $f = |f|\text{sgn}(f)$, $g = |g|\text{sgn}(g)$ for some functions $\text{sgn}(f), \text{sgn}(g)$ of magnitude at most 1. If we then introduce the quantity

$$F(s) := \int_Y (T[|f|^{(1-s)p_\theta/p_0 + sp_\theta/p_1} \text{sgn}(f)])[|g|^{(1-s)q'_\theta/q'_0 + sq'_\theta/q'_1} \text{sgn}(g)] \, d\nu$$

(with the conventions that $q'_\theta/q'_0, q'_\theta/q'_1 = 1$ in the endpoint case $q'_0 = q'_1 = q'_\theta = \infty$) we see that F is a holomorphic function of s of at most exponential growth which equals $\int_Y (Tf)g \, d\nu$ when $s = \theta$. When instead $s = 0 + it$, an application of (15) shows that $|F(s)| \leq 1$; a similar claim obtains when $s = 1 + it$ using (16). The claim now follows from Theorem 4.

The estimate (17) has currently been established for simple functions f, g with finite measure support. But one can extend the claim to any $f \in L^{p_\theta}(X)$ (keeping g simple with finite measure support) by decomposing f into a bounded function and a function of finite measure support, approximating the former in $L^{p_\theta}(X) \cap L^{p_1}(X)$ by simple functions of finite measure support, and approximating the latter in $L^{p_\theta}(X) \cap L^{p_0}(X)$ by simple functions of finite measure support, and taking limits

using (15), (16) to justify the passage to the limit. One can then also allow arbitrary $g \in L^{q'_\theta}(Y)$ by using the monotone convergence theorem. The claim now follows from the duality between $L^{q_1}(Y)$ and $L^{q'_1}(Y)$. \square

Suppose one has a linear operator T that maps simple functions of finite measure support on X to measurable functions on Y (modulo almost everywhere equivalence). We say that such an operator is of strong type (p, q) if it can be extended in a continuous fashion to an operator on $L^p(X)$ to an operator on $L^q(Y)$; this is equivalent to having an estimate of the form $\|Tf\|_{L^q(Y)} \leq B\|f\|_{L^p(X)}$ for all simple functions f of finite measure support. (The extension is unique if p is finite or if X has finite measure, due to the density of simple functions of finite measure support in those cases. Annoyingly, uniqueness fails for L^∞ of an infinite measure space, though this turns out not to cause much difficulty in practice, as the conclusions of interpolation methods are usually for finite exponents p .) Define the *strong type diagram* to be the set of all $(1/p, 1/q)$ such that T is of strong type (p, q) . The Riesz-Thorin theorem tells us that if T is of strong type (p_0, q_0) and (p_1, q_1) with $0 < p_0, p_1 \leq \infty$ and $1 \leq q_0, q_1 \leq \infty$, then T is also of strong type (p_θ, q_θ) for all $0 < \theta < 1$; thus the strong type diagram contains the closed line segment connecting $(1/p_0, 1/q_0)$ with $(1/p_1, 1/q_1)$. Thus the strong type diagram of T is convex in $[0, +\infty) \times [0, 1]$ at least. (As we shall see later, it is in fact convex in all of $[0, +\infty)^2$.) Furthermore, on the intersection of the strong type diagram with $[0, 1] \times [0, +\infty)$, the operator norm $\|T\|_{L^p(X) \rightarrow L^q(Y)}$ is a log-convex function of $(1/p, 1/q)$.

Exercise 23. If $X = Y = [0, 1]$ with the usual measure, show that the strong type diagram of the identity operator is the triangle $\{(1/p, 1/q) \in [0, +\infty) \times [0, +\infty) : 1/p \leq 1/q\}$. If instead $X = Y = \mathbf{Z}$ with the usual counting measure, show that the strong type diagram of the identity operator is the triangle $\{(1/p, 1/q) \in [0, +\infty) \times [0, +\infty) : 1/p \geq 1/q\}$. What is the strong type diagram of the identity when $X = Y = \mathbf{R}$ with the usual measure?

Exercise 24. Let T (resp. T^*) be a linear operator from simple functions of finite measure support on Y (resp. X) to measurable functions on Y (resp. X) modulo a.e. equivalence that are absolutely integrable on finite measure sets. We say T, T^* are formally adjoint if we have $\int_Y (Tf)\bar{g} \, d\nu = \int_X f\overline{T^*g} \, d\mu$ for all simple functions f, g of finite measure support on X, Y respectively. If $1 \leq p, q \leq \infty$, show that T is of strong type (p, q) if and only if T^* is of strong type (q', p') . Thus, taking formal adjoints reflects the strong type diagram around the line of duality $1/p + 1/q = 1$, at least inside the Banach space region $[0, 1]^2$.

Remark 25. There is a powerful extension of the Riesz-Thorin theorem known as the Stein interpolation theorem, in which the single operator T is replaced by a family of operators T_s for $s \in S$ that vary holomorphically in s in the sense that $\int_Y (T_s 1_E) 1_F \, d\nu$ is a holomorphic function of s for any sets E, F of finite measure. Roughly speaking, the Stein interpolation theorem asserts that if T_{j+it} is of strong type (p_j, q_j) for $j = 0, 1$ with a bound growing at most exponentially in t , and T_s itself grows at most exponentially in t in some sense, then T_θ will be of strong type

(p_θ, q_θ) . A precise statement of the theorem and some applications can be found in Stein's book on harmonic analysis.

Now we turn to the real interpolation method. Instead of linear operators, it is now convenient to consider *sublinear operators* T mapping simple functions $f : X \rightarrow \mathbf{C}$ of finite measure support in X to $[0, +\infty]$ -valued measurable functions on Y (modulo almost everywhere equivalence, as usual), obeying the homogeneity relationship

$$|T(cf)| = |c||Tf|$$

and the pointwise bounds

$$|T(f+g)| \leq |Tf| + |Tg|$$

and

$$|Tf - Tg| \leq |T(f - g)|$$

for all $c \in \mathbf{C}$, and all simple functions f, g of finite measure support.

Every linear operator is sublinear; also, the absolute value $Tf := |Sf|$ of a linear (or sublinear) operator is also sublinear. More generally, any *maximal operator* of the form $Tf := \sup_{\alpha \in A} |S_\alpha f|$, where $(S_\alpha)_{\alpha \in A}$ is a family of sub-linear operators, is also a non-negative sublinear operator; note that one can also replace the supremum here by any other norm in α , e.g. one could take an ℓ^p norm $(\sum_{\alpha \in A} |S_\alpha f|^p)^{1/p}$ for any $1 \leq p \leq \infty$. (After $p = \infty$ and $p = 1$, a particularly common case is when $p = 2$, in which case T is known as a *square function*.)

The basic theory of sublinear operators is similar to that of linear operators in some respects. For instance, continuity is still equivalent to boundedness:

Exercise 26. Let T be a sublinear operator, and let $0 < p, q \leq \infty$. Assume that either p is finite, or X has finite measure. Then the following are equivalent:

- T can be extended to a continuous operator from $L^p(X)$ to $L^q(Y)$.
- There exists a constant $B > 0$ such that $\|Tf\|_{L^q(Y)} \leq B\|f\|_{L^p(X)}$ for all simple functions f of finite measure support.
- T can be extended to a operator from $L^p(X)$ to $L^q(Y)$ such that $\|Tf\|_{L^q(Y)} \leq B\|f\|_{L^p(X)}$ for all $f \in L^p(X)$ and some $B > 0$.

Show that the extension mentioned above is unique. Finally, show that the same equivalences hold if $L^q(Y)$ is replaced by $L^{q,\infty}(Y)$ throughout.

We say that T is of *strong type* (p, q) if any of the above equivalent statements (for $L^q(Y)$) hold, and of *weak type* (p, q) if any of the above equivalent statements (for $L^{q,\infty}(Y)$) hold. We say that a linear operator S is of strong or weak type (p, q) if its non-negative counterpart $|S|$ is; note that this is compatible with our previous definition of strong type for such operators. Also, Chebyshev's inequality tells us that strong type (p, q) implies weak type (p, q) .

We now give the real interpolation counterpart of the Riesz-Thorin theorem, namely the *Marcinkiewicz interpolation theorem*:

Theorem 27. (Marcinkiewicz interpolation theorem) Let $0 < p_0, p_1, q_0, q_1 \leq \infty$ and $0 < \theta < 1$ be such that $q_0 \neq q_1$, and $p_i \leq q_i$ for $i = 0, 1$. Let T be a sublinear

operator which is of weak type (p_0, q_0) and of weak type (p_1, q_1) . Then T is of strong type (p_θ, q_θ) .

Remark 28. Of course, the same claim applies to linear operators S by setting $T := |S|$. One can also extend the argument to quasilinear operators, in which the pointwise bound $|T(f+g)| \leq |Tf| + |Tg|$ is replaced by $|T(f+g)| \leq C(|Tf| + |Tg|)$ for some constant $C > 0$, but this generalisation only appears occasionally in applications. The conditions $p_0 \leq q_0, p_1 \leq q_1$ can be replaced by the variant condition $p_\theta \leq q_\theta$ (see Exercise 31, Exercise 33), but cannot be eliminated entirely: see Exercise 32. The precise hypotheses required on $p_0, p_1, q_0, q_1, p_\theta, q_\theta$ are rather technical and I recommend that they be ignored on a first reading.

Proof. For notational reasons it is convenient to take q_0, q_1 finite; however the arguments below can be modified without much difficulty to deal with the infinite case (or one can use a suitable limiting argument); we leave this to the interested reader.

By hypothesis, there exist constants $B_0, B_1 > 0$ such that

$$\lambda_{Tf}(t) \leq B_0^{q_0} \|f\|_{L^{p_0}(X)}^{q_0} / t^{q_0} \quad (18)$$

and

$$\lambda_{Tf}(t) \leq B_1^{q_1} \|f\|_{L^{p_1}(X)}^{q_1} / t^{q_1} \quad (19)$$

for all simple functions f of finite measure support, and all $t > 0$. Let us write $A \lesssim B$ to denote $A \leq C_{p_0, p_1, q_0, q_1, \theta, B_0, B_1} B$ for some constant $C_{p_0, p_1, q_0, q_1, \theta, B_0, B_1}$ depending on the indicated parameters. By (9), it will suffice to show that

$$\int_0^\infty \lambda_{Tf}(t) t^{q_\theta} \frac{dt}{t} \lesssim \|f\|_{L^{p_\theta}(X)}^{q_\theta}.$$

By homogeneity we can normalise $\|f\|_{L^{p_\theta}(X)} = 1$.

Actually, it will be more slightly convenient to work with the dyadic version of the above estimate, namely

$$\sum_{n \in \mathbf{Z}} \lambda_{Tf}(2^n) 2^{q_\theta n} \lesssim 1; \quad (20)$$

see Exercise 11. The hypothesis $\|f\|_{L^{p_\theta}(X)} = 1$ similarly implies that

$$\sum_{m \in \mathbf{Z}} \lambda_f(2^m) 2^{p_\theta m} \lesssim 1. \quad (21)$$

The basic idea is then to get enough control on the numbers $\lambda_{Tf}(2^n)$ in terms of the numbers $\lambda_f(2^m)$ that one can deduce (20) from (21).

When $p_0 = p_1$, the claim follows from direct substitution of (18), (19) (see also the discussion in the previous section about interpolating strong L^p bounds from weak ones), so let us assume $p_0 \neq p_1$; by symmetry we may take $p_0 < p_1$, and thus $p_0 < p_\theta < p_1$. In this case we cannot directly apply (18), (19) because we only control f in L^{p_θ} , not L^{p_0} or L^{p_1} . To get around this, we use the basic real interpolation trick of *decomposing* f into pieces. There are two basic choices for what decomposition to pick. On one hand, one could adopt a “minimalistic” approach and just decompose into two pieces

$$f = f_{\geq s} + f_{< s}$$

where $f_{\geq s} := f1_{|f| \geq s}$ and $f_{< s} := f1_{|f| < s}$, and the threshold s is a parameter (depending on n) to be optimised later. Or we could adopt a “maximalistic” approach and perform the dyadic decomposition

$$f = \sum_{m \in \mathbf{Z}} f_m$$

where $f_m = f1_{2^m \leq |f| < 2^{m+1}}$. (Note that only finitely many of the f_m are non-zero, as we are assuming f to be a simple function.) We will adopt the latter approach, in order to illustrate the dyadic decomposition method; the former approach also works, but we leave it as an exercise to the interested reader.

From sublinearity we have the pointwise estimate

$$Tf \leq \sum_m Tf_m$$

which implies that

$$\lambda_{Tf}(2^n) \leq \sum_m \lambda_{Tf_m}(c_{n,m}2^n)$$

whenever $c_{n,m}$ are positive constants such that $\sum_m c_{n,m} = 1$, but for which we are otherwise at liberty to choose. We will set aside the problem of deciding what the optimal choice of $c_{n,m}$ is for now, and continue with the proof.

From (18), (19), we have two bounds for the quantity $\lambda_{Tf_m}(c_{n,m}2^n)$, namely

$$\lambda_{Tf_m}(c_{n,m}2^n) \lesssim c_{n,m}^{-q_0} 2^{-nq_0} \|f_m\|_{L^{p_0}(X)}^{q_0}$$

and

$$\lambda_{Tf_m}(c_{n,m}2^n) \lesssim c_{n,m}^{-q_1} 2^{-nq_1} \|f_m\|_{L^{p_1}(X)}^{q_1}.$$

From construction of f_m we can bound

$$\|f_m\|_{L^{p_0}(X)} \lesssim 2^m \lambda_f(2^m)^{1/p_0}$$

and similarly for p_1 , and thus we have

$$\lambda_{Tf_m}(c_{n,m}2^n) \lesssim c_{n,m}^{-q_i} 2^{-nq_i} 2^{mq_i} \lambda_f(2^m)^{q_i/p_i}.$$

for $i = 0, 1$. To prove (20), it thus suffices to show that

$$\sum_n 2^{nq_\theta} \sum_m \min_{i=0,1} c_{n,m}^{-q_i} 2^{-nq_i} 2^{mq_i} \lambda_f(2^m)^{q_i/p_i} \lesssim 1.$$

It is convenient to introduce the quantities $a_m := \lambda_f(2^m) 2^{mp_\theta}$ appearing in (21), thus

$$\sum_m a_m \lesssim 1$$

and our task is to show that

$$\sum_n 2^{nq_\theta} \sum_m \min_{i=0,1} c_{n,m}^{-q_i} 2^{-nq_i} 2^{mq_i} 2^{-mq_i p_\theta / p_i} a_m^{q_i/p_i} \lesssim 1.$$

Since $p_i \leq q_i$, we have $a_m^{q_i/p_i} \lesssim a_m$, and so we are reduced to the purely numerical task of locating constants $c_{n,m}$ with $\sum_m c_{n,m} \leq 1$ for all n such that

$$\sum_n 2^{nq_\theta} \min_{i=0,1} c_{n,m}^{-q_i} 2^{-nq_i} 2^{mq_i} 2^{-mq_i p_\theta / p_i} \lesssim 1 \quad (22)$$

for all m .

We can simplify this expression a bit by collecting terms and making some substitutions. The points $(1/p_0, 1/q_0), (1/p_\theta, 1/q_\theta), (1/p_1, 1/q_1)$ are collinear, and we can capture this by writing

$$\frac{1}{p_i} = \frac{1}{p_\theta} + x_i; \quad \frac{1}{q_i} = \frac{1}{q_\theta} + \alpha x_i$$

for some $x_0 > 0 > x_1$ and some $\alpha \in \mathbf{R}$. We can then simplify the left-hand side of (22) to

$$\sum_n \min_{i=0,1} c_{n,m}^{-q_i} (2^{n\alpha q_\theta - mp_\theta})^{q_i x_i}.$$

Note that $q_0 x_0$ is positive and $q_1 x_1$ is negative. If we then pick $c_{n,m}$ to be a sufficiently small multiple of $2^{\beta|n\alpha q_\theta - mp_\theta|}$ where $\beta := \frac{1}{2} \min(x_0, x_1)$ (say), we obtain the claim by summing geometric series. \square

Remark 29. A closer inspection of the proof (or a rescaling argument to reduce to the normalised case $B_0 = B_1 = 1$, as in preceding sections) reveals that one establishes the estimate

$$\|Tf\|_{L^{q_\theta}(Y)} \leq C_{p_0, p_1, q_0, q_1, \theta, C} B_0^{1-\theta} B_1^\theta \|f\|_{L^{p_\theta}(X)}$$

for all simple functions f of finite measure support (or for all $f \in L^{p_\theta}(X)$, if one works with the continuous extension of T to such functions), and some constant $C_{p_0, p_1, q_0, q_1, \theta, C} > 0$. Thus the conclusion here is weaker by a multiplicative constant from that in the Riesz-Thorin theorem, but the hypotheses are weaker too (weak-type instead of strong-type). Indeed, we see that the constant $C_{p_0, p_1, q_0, q_1, \theta}$ must blow up as $\theta \rightarrow 0$ or $\theta \rightarrow 1$.

The power of the Marcinkiewicz interpolation theorem, as compared to the Riesz-Thorin theorem, is that it allows one to weaken the hypotheses on T from strong type to weak type. Actually, it can be weakened further. We say that a non-negative sublinear operator T is *restricted weak-type* (p, q) for some $0 < p, q \leq \infty$ if there is a constant $B > 0$ such that

$$\|Tf\|_{L^{q, \infty}(Y)} \leq B \mu(E)^{1/p}$$

for all sets E of finite measure and all simple functions f with $|f| \leq 1_E$. Clearly restricted weak-type (p, q) is implied by weak-type (p, q) , and thus by strong-type (p, q) . (One can also define the notion of *restricted strong-type* (p, q) by replacing $L^{q, \infty}(Y)$ with $L^q(Y)$; this is between strong-type (p, q) and restricted weak-type (p, q) , but is incomparable to weak-type (p, q) .)

Exercise 30. Show that the Marcinkiewicz interpolation theorem continues to hold if the weak-type hypotheses are replaced by restricted weak-type hypothesis. (Hint: where were the weak-type hypotheses used in the proof?)

We thus see that the strong-type diagram of T contains the interior of the restricted weak-type or weak-type diagrams of T , at least in the triangular region $\{(1/p, 1/q) \in [0, +\infty)^2 : p \geq q\}$.

Exercise 31. Suppose that T is a sublinear operator of restricted weak-type (p_0, q_0) and (p_1, q_1) for some $0 < p_0, p_1, q_0, q_1 \leq \infty$. Show that T is of restricted weak-type (p_θ, q_θ) for any $0 < \theta < 1$, or in other words the restricted type diagram is convex in $[0, +\infty)^2$. (This is an easy result requiring only interpolation of scalars.) Conclude that the hypotheses $p_0 \leq q_0, p_1 \leq q_1$ in the Marcinkiewicz interpolation theorem can be replaced by the variant $p_\theta < q_\theta$.

Exercise 32. For any $\alpha \in \mathbf{R}$, let X_α be the natural numbers \mathbf{N} with the weighted counting measure $\sum_{n \in \mathbf{N}} 2^{\alpha n} \delta_n$, thus each point n has mass $2^{\alpha n}$. Show that if $\alpha > \beta > 0$, then the identity operator from X_α to X_β is of weak-type (p, q) but not strong-type (p, q) when $1 < p, q < \infty$ and $\alpha/p = \beta/q$. Conclude that the hypotheses $p_0 \leq q_0, p_1 \leq q_1$ cannot be dropped entirely.

Exercise 33. Suppose we are in the situation of the Marcinkiewicz interpolation theorem, with the hypotheses $p_0 \leq q_0, p_1 \leq q_1$ replaced by $p_0 \neq p_1$. Show that for all $0 < \theta < 1$ and $1 \leq r \leq \infty$ there exists a $B > 0$ such that

$$\|Tf\|_{L^{q_\theta, r}(Y)} \leq B\|f\|_{L^{p_\theta, r}(X)}$$

for all simple functions f of finite measure support, where the Lorentz norms $L^{p, q}$ were defined in Exercise 13. (Hint: repeat the proof of the Marcinkiewicz interpolation theorem, but partition the sum $\sum_{n, m}$ into regions of the form $\{n\alpha q_\theta - m p_\theta = k + O(1)\}$ for integer k . Obtain a bound for each summand which decreases geometrically as $k \rightarrow \pm\infty$.) Conclude that the hypotheses $p_0 \leq q_0, p_1 \leq q_1$ in the Marcinkiewicz interpolation theorem can be replaced by $p_\theta \leq q_\theta$. This Lorentz space version of the interpolation theorem is in some sense the “right” version of the theorem, but the Lorentz spaces are slightly more technical to deal with than the Lebesgue spaces, and the Lebesgue space version of Marcinkiewicz interpolation is largely sufficient for most applications.

Exercise 34. For $i = 1, 2$, let $X_i = (X_i, \mathcal{X}_i, \mu_i)$, $Y_i = (Y_i, \mathcal{Y}_i, \nu_i)$ be σ -finite measure spaces, and let T_i be a linear operator from simple functions of finite measure support on X_i to measurable functions on Y_i (modulo almost everywhere equivalence, as always). Let $X = X_1 \times X_2$, $Y = Y_1 \times Y_2$ be the product spaces (with product σ -algebra and product measure). Show that there exists a unique (modulo a.e. equivalence) linear operator T defined on linear combinations of indicator functions $1_{E_1 \times E_2}$ of product sets of sets $E_1 \subset X_1$, $E_2 \subset X_2$ of finite measure, such that

$$T1_{E_1 \times E_2}(y_1, y_2) := T_1 1_{E_1}(y_1) T_2 1_{E_2}(y_2)$$

for a.e. $(y_1, y_2) \in Y$; we refer to T as the tensor product of T_1 and T_2 and write $T = T_1 \otimes T_2$. Show that if T_1, T_2 are of strong-type (p, q) for some $1 \leq p, q < \infty$ with operator norms B_1, B_2 respectively, then T can be extended to a bounded linear operator on $L^p(X)$ to $L^q(Y)$ with operator norm exactly equal to $B_1 B_2$, thus

$$\|T_1 \otimes T_2\|_{L^p(X_1 \times X_2) \rightarrow L^q(Y_1 \times Y_2)} = \|T_1\|_{L^p(X_1) \rightarrow L^q(Y_1)} \|T_2\|_{L^p(X_2) \rightarrow L^q(Y_2)}.$$

(Hint: for the lower bound, show that $T_1 \otimes T_2(f_1 \otimes f_2) = (T_1 f_1) \otimes (T_2 f_2)$ for all simple functions f_1, f_2 . For the upper bound, express $T_1 \times T_2$ as the composition of two other operators $T_1 \otimes I_1$ and $I_2 \otimes T_2$ for some identity operators I_1, I_2 , and

establish operator norm bounds on these two operators separately.) Use this and the tensor power trick to deduce the Riesz-Thorin theorem (in the special case when $1 \leq p_i \leq q_i < \infty$ for $i = 0, 1$, and $q_0 \neq q_1$) from the Marcinkiewicz interpolation theorem. Thus one can (with some effort) avoid the use of complex variable methods to prove the Riesz-Thorin theorem, at least in some cases.

Exercise 35. (Hölder's inequality for Lorentz spaces) Let $f \in L^{p_1, r_1}(X)$ and $g \in L^{p_2, r_2}(X)$ for some $0 < p_1, p_2, r_1, r_2 \leq \infty$. Show that $fg \in L^{p_3, r_3}(X)$, where $1/p_3 = 1/p_1 + 1/p_2$ and $1/r_3 = 1/r_1 + 1/r_2$, with the estimate

$$\|fg\|_{L^{p_3, r_3}(X)} \leq C_{p_1, p_2, r_1, r_2} \|f\|_{L^{p_1, r_1}(X)} \|g\|_{L^{p_2, r_2}(X)}$$

for some constant C_{p_1, p_2, r_1, r_2} . (This estimate is due to O'Neil.)

Remark 36. Just as interpolation of functions can be clarified by using step functions $f = A1_E$ as a test case, it is instructive to use rank one operators such as

$$Tf := A\langle f, 1_E \rangle 1_F = A\left(\int_E f \, d\mu\right) 1_F$$

where $E \subset X, F \subset Y$ are finite measure sets, as test cases for the real and complex interpolation methods. (After understanding the rank one case, I then recommend looking at the rank two case, e.g. $Tf := A_1\langle f, 1_{E_1} \rangle 1_{F_1} + A_2\langle f, 1_{E_2} \rangle 1_{F_2}$, where E_2, F_2 could be very different in size from E_1, F_1 .)

4. SOME EXAMPLES OF INTERPOLATION

Now we apply the interpolation theorems to some classes of operators. An important such class is given by the *integral operators*

$$Tf(y) := \int_X K(x, y) f(x) \, d\mu(x)$$

from functions $f : X \rightarrow \mathbf{C}$ to functions $Tf : Y \rightarrow \mathbf{C}$, where $K : X \times Y \rightarrow \mathbf{C}$ is a fixed measurable function, known as the *kernel* of the integral operator T . Of course, this integral is not necessarily convergent, so we will also need to study the sublinear analogue

$$|T|f(y) := \int_X |K(x, y)| |f(x)| \, d\mu(x)$$

which is well-defined (though it may be infinite).

The following useful lemma gives us strong-type bounds on $|T|$ and hence T , assuming certain L^p type bounds on the rows and columns of K .

Lemma 37. (Schur's test) Let $K : X \times Y \rightarrow \mathbf{C}$ be a measurable function obeying the bounds

$$\|K(x, \cdot)\|_{L^{q_0}(Y)} \leq B_0$$

for almost every $x \in X$, and

$$\|K(\cdot, y)\|_{L^{p'_1}(X)} \leq B_1$$

for almost every $y \in Y$, where $1 \leq p_1, q_0 \leq \infty$ and $B_0, B_1 > 0$. Then for every $0 < \theta < 1$, $|T|$ and T are of strong-type (p_θ, q_θ) , with $Tf(y)$ well-defined for all $f \in L^{p_\theta}(X)$ and almost every $y \in Y$, and furthermore

$$\|Tf\|_{L^{q_\theta}(Y)} \leq B_\theta \|f\|_{L^{p_\theta}(X)}.$$

Here we adopt the convention that $p_0 := 1$ and $q_1 := \infty$, thus $q_\theta = q_0/(1 - \theta)$ and $p'_\theta = p'_1/\theta$.

Proof. The hypothesis $\|K(x, \cdot)\|_{L^{q_0}(Y)} \leq B_0$, combined with *Minkowski's integral inequality*, shows us that

$$\||T|f\|_{L^{q_0}(Y)} \leq B_0 \|f\|_{L^1(X)}$$

for all $f \in L^1(X)$; in particular, for such f , Tf is well-defined almost everywhere, and

$$\|Tf\|_{L^{q_0}(Y)} \leq B_0 \|f\|_{L^1(X)}.$$

Similarly, Hölder's inequality tells us that for $f \in L^{p_1}(X)$, Tf is well-defined everywhere, and

$$\|Tf\|_{L^\infty(Y)} \leq B_1 \|f\|_{L^{p_1}(X)}.$$

Applying the Riesz-Thorin theorem we conclude that

$$\|Tf\|_{L^{q_\theta}(Y)} \leq B_\theta \|f\|_{L^{p_\theta}(X)}$$

for all simple functions f with finite measure support; replacing K with $|K|$ we also see that

$$\||T|f\|_{L^{q_\theta}(Y)} \leq B_\theta \|f\|_{L^{p_\theta}(X)}$$

for all simple functions f with finite measure support, and thus (by monotone convergence) for all $f \in L^{p_\theta}(X)$. The claim then follows. \square

Example 38. Let $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ be a matrix such that the sum of the magnitudes of the entries in every row and column is at most B , i.e. $\sum_{i=1}^n |a_{ij}| \leq B$ for all j and $\sum_{j=1}^m |a_{ij}| \leq B$ for all i . Then one has the bound

$$\|Ax\|_{\ell_m^p} \leq B \|x\|_{\ell_n^p}$$

for all vectors $x \in \mathbf{C}^n$ and all $1 \leq p \leq \infty$. Note the extreme cases $p = 1$, $p = \infty$ can be seen directly; the remaining cases then follow from interpolation.

A useful special case arises when A is an S -sparse matrix, which means that at most S entries in any row or column are non-zero (e.g. permutation matrices are 1-sparse). We then conclude that the ℓ^p operator norm of A is at most $S \sup_{i,j} |a_{i,j}|$.

Exercise 39. Establish Schur's test by more direct means, taking advantage of the duality relationship

$$\|g\|_{L^p(Y)} := \sup\left\{\left|\int_Y gh\right| : \|h\|_{L^{p'}(Y)} \leq 1\right\}$$

for $1 \leq p \leq \infty$, as well as Young's inequality $xy \leq \frac{1}{r}x^r + \frac{1}{r'}x^{r'}$ for $1 < r < \infty$. (You may wish to first work out Example 38, say with $p = 2$, to figure out the logic.)

A useful corollary of Schur's test is *Young's convolution inequality* for the convolution $f * g$ of two functions $f : \mathbf{R}^n \rightarrow \mathbf{C}$, $g : \mathbf{R}^n \rightarrow \mathbf{C}$, defined as

$$f * g(x) := \int_{\mathbf{R}^n} f(y)g(x-y) dy$$

provided of course that the integrand is absolutely convergent.

Exercise 40. (*Young's inequality*) Let $1 \leq p, q, r \leq \infty$ be such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$. Show that if $f \in L^p(\mathbf{R}^n)$ and $g \in L^q(\mathbf{R}^n)$, then $f * g$ is well-defined almost everywhere and lies in $L^r(\mathbf{R}^n)$, and furthermore that

$$\|f * g\|_{L^r(\mathbf{R}^n)} \leq \|f\|_{L^p(\mathbf{R}^n)} \|g\|_{L^q(\mathbf{R}^n)}.$$

(Hint: Apply Schur's test to the kernel $K(x, y) := g(x - y)$.)

Remark 41. There is nothing special about \mathbf{R}^n here; one could in fact use any locally compact group G with a bi-invariant Haar measure. On the other hand, if one specialises to \mathbf{R}^n , then it is possible to improve Young's inequality slightly, to

$$\|f * g\|_{L^r(\mathbf{R}^n)} \leq (A_p A_q A_{r'})^{n/2} \|f\|_{L^p(\mathbf{R}^n)} \|g\|_{L^q(\mathbf{R}^n)},$$

where $A_p := p^{1/p}/(p')^{1/p'}$, a result of Beckner; the constant here is best possible, as can be seen by testing the inequality in the case when f, g are Gaussians.

Exercise 42. Let $1 \leq p \leq \infty$, and let $f \in L^p(\mathbf{R}^n)$, $g \in L^{p'}(\mathbf{R}^n)$. Young's inequality tells us that $f * g \in L^\infty(\mathbf{R}^n)$. Refine this further by showing that $f * g \in C_0(\mathbf{R}^n)$, i.e. $f * g$ is continuous and goes to zero at infinity. (Hint: first show this when $f, g \in C_c(\mathbf{R}^n)$, then use a limiting argument.)

We now give a variant of Schur's test that allows for weak estimates.

Lemma 43. (*Weak-type Schur's test*) Let $K : X \times Y \rightarrow \mathbf{C}$ be a measurable function obeying the bounds

$$\|K(x, \cdot)\|_{L^{q_0, \infty}(Y)} \leq B_0$$

for almost every $x \in X$, and

$$\|K(\cdot, y)\|_{L^{p'_1, \infty}(X)} \leq B_1$$

for almost every $y \in Y$, where $1 < p_1, q_0 < \infty$ and $B_0, B_1 > 0$ (note the endpoint exponents $1, \infty$ are now excluded). Then for every $0 < \theta < 1$, $|T|$ and T are of strong-type (p_θ, q_θ) , with $Tf(y)$ well-defined for all $f \in L^{p_\theta}(X)$ and almost every $y \in Y$, and furthermore

$$\|Tf\|_{L^{q_\theta}(Y)} \leq C_{p_1, q_0, \theta} B_\theta \|f\|_{L^{p_\theta}(X)}.$$

Here we again adopt the convention that $p_0 := 1$ and $q_1 := \infty$.

Proof. From Exercise 17 we see that

$$\int_Y |K(x, y)| 1_E(y) d\nu(y) \lesssim B_0 \mu(E)^{1/q'_0}$$

for any measurable $E \subset Y$, where we use $A \lesssim B$ to denote $A \leq C_{p_1, q_0, \theta} B$ for some $C_{p_1, q_0, \theta}$ depending on the indicated parameters. By the Fubini-Tonelli theorem, we conclude that

$$\int_Y |T|f(y)1_E(y) d\nu(y) \lesssim B_0\mu(E)^{1/q'_0} \|f\|_{L^1(X)}$$

for any $f \in L^1(X)$; by Exercise 17 again we conclude that

$$\| |T|f \|_{L^{q_0, \infty}(Y)} \lesssim B_0 \|f\|_{L^1(X)}$$

thus $|T|$ is of weak-type $(1, q_0)$. In a similar vein, from yet another application of Exercise 17 we see that

$$\| |T|f \|_{L^\infty(Y)} \lesssim B_1\mu(F)^{1/p_1}$$

whenever $0 \leq f \leq 1_F$ and $F \subset X$ has finite measure; thus $|T|$ is of restricted type (p_1, ∞) . Applying Exercise 30 we conclude that $|T|$ is of strong type (p_θ, q_θ) (with operator norm $\lesssim B_\theta$), and the claim follows. \square

This leads to a weak-type version of Young's inequality:

Exercise 44. (*Weak-type Young's inequality*) Let $1 < p, q, r < \infty$ be such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$. Show that if $f \in L^p(\mathbf{R}^n)$ and $g \in L^{q, \infty}(\mathbf{R}^n)$, then $f * g$ is well-defined almost everywhere and lies in $L^r(\mathbf{R}^n)$, and furthermore that

$$\|f * g\|_{L^r(\mathbf{R}^n)} \leq C_{p, q} \|f\|_{L^p(\mathbf{R}^n)} \|g\|_{L^{q, \infty}(\mathbf{R}^n)}.$$

for some constant $C_{p, q} > 0$.

Exercise 45. Refine the previous exercise by replacing $L^r(\mathbf{R}^n)$ with the Lorentz space $L^{r, p}(\mathbf{R}^n)$ throughout.

Recall that the function $1/|x|^\alpha$ will lie in $L^{n/\alpha, \infty}(\mathbf{R}^n)$ for $\alpha > 0$. We conclude

Corollary 46. (*Hardy-Littlewood-Sobolev fractional integration inequality*) Let $1 < p, r < \infty$ and $0 < \alpha < n$ be such that $\frac{1}{p} + \frac{\alpha}{n} = \frac{1}{r} + 1$. If $f \in L^p(\mathbf{R}^n)$, then the function $I_\alpha f$, defined as

$$I_\alpha f(x) := \int_{\mathbf{R}^n} \frac{f(y)}{|x-y|^\alpha} dy$$

is well-defined almost everywhere and lies in $L^r(\mathbf{R}^n)$, and furthermore that

$$\|I_\alpha f\|_{L^r(\mathbf{R}^n)} \leq C_{p, \alpha, n} \|f\|_{L^p(\mathbf{R}^n)}$$

for some constant $C_{p, \alpha, n} > 0$.

This inequality is of importance in the theory of Sobolev spaces, which we will discuss in a subsequent lecture.

Exercise 47. Show that Corollary 46 can fail at the endpoints $p = 1$, $r = \infty$, or $\alpha = n$.