LESSON 26 - STUDY GUIDE

Abstract. This is the last lesson of the course. We wrap up the results and techniques that we have developed in the previous lessons to show, first, that Poisson integrals $P_r * f$ of functions $f \in L^1(\mathbb{T})$ converge pointwise almost everywhere to f when $r \to 1^-$. Then, we use this result to prove that the harmonic conjugate $v = Q_r * f$ of the Poisson integral $u = P_r * f$ of $f \in L^1(\mathbb{T})$ also has pointwise limit almost everywhere when $r \to 1^-$. We define the Hilbert transform of f as this pointwise limit which is obviously related to the conjugate of f, as defined by a Fourier multiplier operator in Lesson 22. Finally, we present the theorem by Besicovitch and Kolmogorov on the weak type $(1, 1)$ bounds for conjugation, concluding with Marcel Riesz's theorem for conjugation in $L^p(\mathbb{T})$ and corresponding convergence of Fourier series, in norm.

1. Pointwise almost everywhere convergence of Poisson integrals and their conjugates, the Hilbert transform, Besicovitch-Kolmogorov's theorem on the weak type $(1, 1)$ of the conjugation operator and Marcel Riesz's theorem for conjugation in $L^p(\mathbb{T})$, $1 < p < \infty$.

Study material: For this final lesson I followed mostly Katznelson [2, 3] in sections **1 - The The Conjugate Function** and **2 - The Maximal Function of Hardy and Littlewood** from chapter **III - The Conjugate Function and Functions Analytic in the Unit Disc** but a lot of the presentation, including some proofs, are my own adaptation and somewhat different from the book. A lot of this material, on the complex variable methods for proving boundary limits of harmonic functions and conjugation in $L^p(\mathbb{T})$ can be found in many other excellent books focused on advanced complex analysis or the theory of harmonic functions, of which Hoffman's [1] and Rudin's [4] are good examples.

In the last two lessons we introduced and developed a few of the important tools of modern harmonic analysis consisting of weak L^p spaces and the Marcinkiewicz interpolation theorem, for general measure spaces. And the concept of maximal operator of a family, or sequence, of linear operators in L^p, in particular the Hardy-Littlewood maximal function in $\mathbb{R}^n$, as a means to prove pointwise convergence almost everywhere of approximate identities.

Returning now to $\mathbb{T}$, some observations and adaptations should be made, although not crucial. The first one is that $\mathbb{T}$ has finite total Lebesgue measure 2π. So the distribution function[1]

$$\lambda_f(\alpha) = |\{t \in \mathbb{T} : |f(t)| > \alpha\}|,$$

for a Lebesgue measurable function $f : \mathbb{T} \to \mathbb{C}$ is always bounded by 2π, and therefore we do not need to be concerned with its behavior as $\alpha \to 0$. Comparatively, in $\mathbb{R}^n$, as we noted in Lesson 24, the behavior of the distribution function at both ends of the interval $\alpha \in]0, \infty[$ is important - for the rate at which $\lambda_f(\alpha)$ decreases to zero as $\alpha \to \infty$ describes the way f blows up to infinity locally, while the rate at which $\lambda_f(\alpha)$ increases to infinity as $\alpha \to 0$ describes how fast the function is globally decaying to zero as the variable in the unbounded domain moves out to infinity - and both these extremes play a crucial role, and need to be carefully controlled, for the integral

$$\|f\|_{L^p}^p = p \int_0^\infty \alpha^{p-1} \lambda_f(\alpha) d\alpha, \tag{1.1}$$

Date: May 29, 2020.

[1]Recall that we have been using the absolute value of a subset of $\mathbb{T}$ or $\mathbb{R}^n$ to denote its Lebesgue measure.

to be finite. As we saw, a weak L^p function is precisely one for which it can only be guaranteed that $\lambda_f(\alpha) \leq C/\alpha^p$ throughout the whole range $\alpha \in]0, \infty[$ so that it barely fails to be integrable at *both* ends of the interval, diverging logarithmically. To obtain finite L^p norm one needs to ensure a little bit extra, that $\lambda_f(\alpha)$ decreases to zero slightly faster than $1/\alpha^p$ when $\alpha \to \infty$ for the integral (1.1) to converge at infinity, while $\lambda_f(\alpha)$ should grow more slowly to infinity than $1/\alpha^p$ when $\alpha \to 0$ for the integral (1.1) to converge at zero. But on the circle $\mathbb{T}$, or more generally on domains of finite measure, this latter problem related to the infinite measure of an unbounded domain does not exist: $\lambda_f(\alpha)$ is bounded by the finite measure of the whole domain and at $\alpha = 0$ it is finite, at most 2π for $\mathbb{T}$, so that every measurable function will be strongly L^p for (1.1) in the neighborhood of $\alpha = 0$. We therefore only need to be concerned with how fast $\lambda_f(\alpha)$ converges to zero as $\alpha \to \infty$, corresponding to local singularities of the function. So, it is enough to show that

$$\lambda_f(\alpha) \leq \frac{C}{\alpha^p} \quad \text{for} \quad \alpha > R,$$

for some fixed large $R > 0$, to ensure that $f \in L^p_w(\mathbb{T})$[2]. And, for example, the Chebyshev inequality can, in this case, more accurately be stated as

$$\lambda_f(\alpha) \leq \min\left(2\pi, \frac{\|f\|^p_{L^p(\mathbb{T})}}{\alpha^p}\right),$$

for $f \in L^p(\mathbb{T})$. One relevant consequence of this fact, analogous to the nesting of $L^p(\mathbb{T})$ spaces, is that for the circle we have $L^p_w(\mathbb{T}) \subset L^q(\mathbb{T})$, for $0 < q < p \leq \infty$.

Another point where $\mathbb{T}$ demands a slight modification from what we did before is in the definition of the Hardy-Littlewood maximal function. Again, in $\mathbb{R}^n$ we defined it as the supremum of the averages of $|f|$ over *all* possible balls centered at $x \in \mathbb{R}^n$ and naturally, on $\mathbb{T}$ it only makes sense to consider intervals, i.e. one dimensional balls, whose diameter is at most 2π. So, for $f \in L^1(\mathbb{T}) = L^1_{loc}(\mathbb{T})$ the definition should now be

$$Mf(t_0) = \sup_{0<r<\pi} \frac{1}{|B_r(t_0)|} \int_{B_r(t_0)} |f(t)| dt = \sup_{0<r<\pi} \frac{1}{2r} \int_{t_0-r}^{t_0+r} |f(t)| dt.$$

The proof that this definition of the Hardy-Littlewood maximal function is weak type $(1, 1)$ and strong type (p, p) for any $1 < p \leq \infty$ is exactly like the proof of Theorem 1.5 in the last lesson, for $\mathbb{R}^n$. However, due to the observations made above concerning the finite measure of $\mathbb{T}$ we can now say a bit more. We first need a definition of a new class of functions, though.

Definition 1.1. *Let* $\log^+ : \mathbb{R} \to \mathbb{R}$ *be defined as the positive part of the* log *function, i.e.* $\log^+ x = \log x$ *for* $x \geq 1$ *and* $\log^+ x = 0$ *for* $x < 1$. *We define the Zygmund class* $L \log L(\mathbb{T})$ *as the space of Lebesgue measurable functions* $f : \mathbb{T} \to \mathbb{C}$ *for which* $|f(t)| \log^+ |f(t)| \in L^1(\mathbb{T})$.

It is a simple exercise to show that $L \log L(\mathbb{T}) \subset L^1(T)$. A slight modification of the proof of the Marcinkiewicz interpolation theorem, that generally guarantees the strong (p, p) bounds for the Hardy-Littlewood maximal function, from the interpolation of the weak $(1, 1)$ and strong (∞, ∞) cases, will now yield strong $L^1(\mathbb{T})$ estimates from $f \in L \log L(\mathbb{T})$ by exploiting the finite measure of $\mathbb{T}$. Recall, in comparison, that on $\mathbb{R}^n$, the Hardy-Littlewood maximal function is *never* in $L^1(\mathbb{R}^n)$ unless $f = 0$.

Theorem 1.2. *Let* $f \in L \log L(\mathbb{T})$. *Then* $Mf \in L^1(\mathbb{T})$.

[2]Highlighting this point, Katznelson [2, 3]for example defines the distribution function of f on $\mathbb{T}$, not as the usual $\lambda_f(\alpha) = |\{t \in \mathbb{T} : |f(t)| > \alpha\}|$ but as the measure of the complement $m_f(\alpha) = |\{t \in \mathbb{T} : |f(t)| \leq \alpha\}| = 2\pi - \lambda_f(\alpha)$, for which $L^p_w(\mathbb{T})$ corresponds to the lower bound $m_f(\alpha) \geq 2\pi - \frac{C}{\alpha^p}$ and thus making it very clear that the estimate is only important for large $\alpha \to \infty$, with the rate at which $m_f(\alpha)$ grows to 2π, the full measure of $\mathbb{T}$, because when $\alpha \to 0$ the right hand side becomes negative and the inequality becomes trivial.

Proof. Let $f \in L\log L(\mathbb{T})$. Then, as we did in the proof of Marcinkiewicz, taking $\alpha \geq 1$ and splitting the function into high and low parts we have $f = f_0^\alpha + f_1^\alpha$, with $f_0^\alpha(t) = f\chi_{\{|f(t)|>\alpha\}}$ and $f_1^\alpha(t) = f\chi_{\{|f(t)|\leq\alpha\}}$ where obviously $f_0^\alpha \in L^1(\mathbb{T})$ and $f_1^\alpha \in L^\infty(\mathbb{T})$. Then, being sublinear, the Hardy-Littlewood maximal function satisfies

$$Mf \leq Mf_0^\alpha + Mf_1^\alpha \leq Mf_0^\alpha + \alpha,$$

where we used the (∞,∞) bound, i.e. the fact that the Hardy-Littlewood maximal operator satisfies $\|Mf\|_{L^\infty(\mathbb{T})} \leq \|f\|_{L^\infty(\mathbb{T})}$ and that $\|f_1^\alpha\|_{L^\infty(\mathbb{T})} \leq \alpha$. Therefore $\{t \in \mathbb{T} : Mf(t) > 2\alpha\} \subset \{t \in \mathbb{T} : Mf_0^\alpha(t) > \alpha\}$ and therefore the corresponding distribution functions satisfy

$$\lambda_{Mf}(2\alpha) = |\{t \in \mathbb{T} : Mf(t) > 2\alpha\}| \leq |\{t \in \mathbb{T} : Mf_0^\alpha(t) > \alpha\}| = \lambda_{Mf_0^\alpha}(\alpha),$$

and now using the weak $(1,1)$ bound of the Hardy-Littlewood maximal operator

$$\lambda_{Mf}(2\alpha) \leq \frac{C_1\|f_0^\alpha\|_{L^1(\mathbb{T})}}{\alpha}.$$

Up to this point, the proof has been an exact reproduction of the Marcinkiewicz interpolation theorem's proof for the specific case of the Hardy-Littlewood maximal operator between the end points $(1,1)$ and (∞,∞). Now, however, we exploit the finite measure of $\mathbb{T}$ to estimate the $L^1(\mathbb{T})$ norm of Mf

$$\|Mf\|_{L^1(\mathbb{T})} = \frac{1}{2\pi}\int_{\mathbb{T}} |Mf(t)|dt = \frac{1}{2\pi}\int_0^\infty \lambda_{Mf}(\alpha)d\alpha = \frac{1}{\pi}\int_0^\infty \lambda_{Mf}(2\alpha)d\alpha,$$

and we can disregard the integral in the neighborhood of $\alpha = 0$, so we split it in two

$$\|Mf\|_{L^1(\mathbb{T})} = \frac{1}{\pi}\int_0^\infty \lambda_{Mf}(2\alpha)d\alpha = \frac{1}{\pi}\int_0^1 \lambda_{Mf}(2\alpha)d\alpha + \frac{1}{\pi}\int_1^\infty \lambda_{Mf}(2\alpha)d\alpha \leq 2 + \frac{1}{\pi}\int_1^\infty \lambda_{Mf}(2\alpha)d\alpha,$$

(this is the reason why we only considered $\alpha > 1$ at the beginning, when separating f into high and low parts) and we can now use the weak $(1,1)$ estimate for the only part of the integral that needs to be controlled, as $\alpha \to \infty$,

$$\begin{aligned}\int_1^\infty \lambda_{Mf}(2\alpha)d\alpha \leq \int_1^\infty \frac{C_1\|f_0^\alpha\|_{L^1(\mathbb{T})}}{\alpha}d\alpha &= \int_1^\infty \frac{C_1}{2\pi\alpha}\left(\int_{\{|f(t)|>\alpha\}} |f(t)|dt\right)d\alpha \\ &= \frac{C_1}{2\pi}\int_{\mathbb{T}} |f(t)| \int_1^{|f(t)|} \frac{1}{\alpha}d\alpha\, dt = \frac{C_1}{2\pi}\int_{\mathbb{T}} |f(t)|\log^+|f(t)|dt.\end{aligned}$$

We therefore conclude that $\|Mf\|_{L^1(\mathbb{T})} \leq 2 + \frac{C_1}{2\pi^2}\int_{\mathbb{T}} |f(t)|\log^+|f(t)|dt < \infty$. □

The fact that the Hardy-Littlewood maximal operator is weak type $(1,1)$ is enough to prove the Lebesgue differentiation theorem on $\mathbb{T}$ as well as the fact that almost every point is a Lebesgue point, like we did on $\mathbb{R}^n$. Actually, we could also have extended any $L^1(\mathbb{T})$ function to its 2π-periodic version on the whole real line $\mathbb{R}$, which would then be $L^1_{loc}(\mathbb{R})$, and apply to it there the theory for $\mathbb{R}^n$ seen in the last lesson, to obtain the same (local) pointwise limits of the averages on balls. Having thus shown that almost every point of $\mathbb{T}$ is a Lebesgue point, we could then apply the Lebesgue and Fatou theorems of Lesson 19, Theorems 1.5 an 1.8 respectively, to conclude that for any $f \in L^1(\mathbb{T})$, both for the Cesàro means, as well as for the Abel means, we have the pointwise limits

$$\lim_{N\to\infty} \sigma_N(f)(t_0) = \lim_{N\to\infty} K_N * f(t_0) = f(t_0),$$

and

$$\lim_{r\to 1^-} P_r * f(t_0) = f(t_0),$$

at every Lebesgue point t_0 of f, which we now know is almost every point. Actually, the condition (1.3) in Lesson 19 is not exactly the same as the limit over centered balls for the Lebesgue points (1.6) of Lesson 25, but the adaptation is obvious. We would then have proved the final sentence of both Theorems 1.5 and 1.8 of Lesson 19.

These two theorems, whose proofs we actually did not present in Lesson 19 (found in Katznelson [2, 3] in section **3 - Pointwise Convergence of $\sigma_n(f)$** from chapter **I - Fourier Series on $\mathbb{T}$**), but that, similarly to the previous proof of the Fejér theorem 1.1 in Lesson 19, consist in carefully estimating the specific Fejér and Poisson kernels in order to convert the limit in the Lebesgue point condition into a pointwise limit of the Cesàro or Abel means, have the advantage of identifying precisely at which points the pointwise limit occurs - the Lebesgue points - but by themselves they do not establish almost everywhere convergence. That, in turn, comes from the Hardy-Littlewood maximal function machinery which guarantees that almost every point in $\mathbb{T}$ is a Lebesgue point. Our goal, however, was to obtain direct results of pointwise convergence almost everywhere of general families of approximate identities, something that we now know, from Theorem 1.2 in Lesson 25, is just a consequence of establishing weak type (p, q) bounds for the corresponding maximal operators. Such a result is what we achieved in Corollary 1.12 at the end of Lesson 25, for approximate identities in $\mathbb{R}^n$ constructed from rescaled families of functions.

However, on $\mathbb{T}$, neither one of our most important summability kernels, Fejér's $K_N(t) = \frac{1}{N+1}\left(\frac{\sin\frac{N+1}{2}t}{\sin\frac{t}{2}}\right)^2$ nor Poisson's $P_r(t) = \frac{1-r^2}{1-2r\cos t+r^2}$, consist of approximate identities obtained by rescaling. We therefore need a slight variant of Theorem 1.10 and Corollary 1.11, from the last lesson, to suit these summability kernels on $\mathbb{T}$ in order to bound their corresponding maximal convolution operators by the Hardy-Littlewood maximal function. Looking carefully into the proof of Theorem 1.10, though, reveals that the rescaling is really not fundamental: its only purpose is to make sure that the L^1 integral of the convolution kernels remains invariant, so that one gets a pointwise uniform bound of all the convolutions by a fixed constant multiple of the Hardy-Littlewood maximal function. That uniform bound is easily achieved by rescaling, as seen in Theorem 1.10 and Corollary 1.11, but not exclusively. In fact, having bounded L^1 norms is one of the conditions for the definition of an approximate identity, and rescaling is just one way to achieve it. For our purposes on $\mathbb{T}$, a similar result for the convolution with a single even (radially symmetric around the origin), decreasing function will suffice, because that is the essential ingredient, as long as we make sure separately that there will be a uniform L^1 bound for the whole family of convolution kernels.

Proposition 1.3. *Let $\phi : \mathbb{T} \to \mathbb{R}$ be a nonnegative, even function (with respect to $t_0 = 0$), decreasing on $[0, \pi]$ and such that $\frac{1}{2\pi}\int_{\mathbb{T}} \phi(t)dt = c > 0$. Then, for all $f \in L^1(\mathbb{T})$, we have*

$$|\phi * f(t)| = \left|\frac{1}{2\pi}\int_{\mathbb{T}} \phi(t-s)f(s)ds\right| \leq c\, Mf(t).$$

Proof. The proof is exactly like the one for Theorem 1.10 of Lesson 25, except that there is no rescaling to be considered here because we are now dealing with a single convolution kernel. The function ϕ can be approximated on $\mathbb{T} =]-\pi, \pi]$ by an increasing sequence of positive, even, simple functions $\psi = \sum_{j=1}^{N} a_j \chi_{I_j}(t)$ where $a_j > 0$ and the $I_j = [-t_j, t_j]$ are symmetric intervals centered at the origin.

For these simple functions we have

$$\begin{aligned}\psi * |f|(t) \le \frac{1}{2\pi}\int_{-\pi}^{\pi}\psi(s)|f(t-s)|ds &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{j=1}^{N} a_j\chi_{I_j}(s)|f(t-s)|ds \\ &\le \frac{1}{2\pi}\sum_{j=1}^{N} a_j\int_{-\pi}^{\pi}\chi_{I_j}(s)|f(t-s)|ds = \frac{1}{2\pi}\sum_{j=1}^{N} a_j\int_{I_j}|f(t-s)|ds \\ &= \frac{1}{2\pi}\sum_{j=1}^{N} a_j|I_j|\frac{1}{|I_j|}\int_{I_j}|f(t-s)|ds \le \left(\frac{1}{2\pi}\sum_{j=1}^{N} a_j|I_j|\right)Mf(t) \\ &= \left(\frac{1}{2\pi}\int_{\mathbb{T}}\psi(s)ds\right)Mf(t) \le cMf(t).\end{aligned}$$

The bound for $|\phi * f(t)| \le \phi * |f|(t)$ can then be obtained from these simple functions ψ by the monotone convergence theorem[3]. □

So, if we now look at the Abel means of a function $f \in L^1(\mathbb{T})$, $P_r * f(t)$ then the Poisson kernel for each fixed $0 < r < 1$

$$P_r(t) = \frac{1-r^2}{1-2r\cos t + r^2},$$

is a nonnegative, even, decreasing function of $t \in [0, \pi]$, such that $\frac{1}{2\pi}\int_{\mathbb{T}} P_r(t)dt = 1$. So that, with each such P_r playing the role of ϕ in Proposition 1.3 we obtain the desired uniform pointwise bound for all $0 < r < 1$, $|P_r * f(t)| \le Mf(t)$ and therefore the corresponding maximal operator for the whole approximate identity

$$\sup_{0<r<1} |P_r * f(t)| \le Mf(t).$$

For the Cesàro means, the Fejér kernel

$$K_N = \frac{1}{N+1}\left(\frac{\sin\frac{N+1}{2}t}{\sin\frac{t}{2}}\right)^2,$$

is nonnegative and even, for every fixed N, with $\frac{1}{2\pi}\int_{\mathbb{T}} K_N(t)dt = 1$. But it obviously is not decreasing. However, analogously to Corollary 1.11 of the last lesson, for each N it can be bounded by such even, nonnegative and decreasing kernels with uniformly bounded L^1 norms, which we leave as an exercise to prove. Therefore, its maximal operator also satisfies

$$\sup_N |\sigma_N(f)(t)| = \sup_N |K_N * f(t)| \le cMf(t).$$

Therefore we have proved the important pointwise convergence theorems for our two main summabilty kernels.

Theorem 1.4. *Let $f \in L^1(\mathbb{T})$. Then, both its Cesàro and Abel means converge pointwise almost everywhere to f,*

$$\lim_{N\to\infty}\sigma_N(f)(t) = \lim_{N\to\infty} K_N * f(t) = f(t) \quad \textit{and} \quad \lim_{r\to 1^-} P_r * f(t) = f(t),$$

for almost every $t \in \mathbb{T}$.

[3]A different, very simple, proof can also be found in Katznelson [2, 3] in section **2 - The Maximal Function of Hardy and Littlewood** from chapter **III - The Conjugate Function and Functions Analytic in the Unit Disc**.

We are finally ready to return to our study of harmonic functions on the unit disc and the conjugation problem, that will lead to the proof the convergence in $L^p(\mathbb{T})$ norm of the Fourier series. We start by converting Theorem 1.4 of Lesson 23, on the characterization of harmonic functions in the Hardy spaces $h^p(D)$, in terms of Poisson integrals, into a pointwise limit characterizations as the radius approaches the boundary ∂D.

Theorem 1.5. *Considering the Hardy spaces $h^p(D)$ of harmonic functions on the unit disk $D = \{z \in \mathbb{C} : |z| < 1\}$ we have*

(1) *If $1 < p \leq \infty$ and $u \in h^p(D)$ then $\lim_{r\to 1^-} u(re^{it}) = \lim_{r\to 1^-} u_r(t) = f(t)$ almost everywhere $t \in \mathbb{T}$, where $f \in L^p(\mathbb{T})$ is such that $u = P_r * f$.*
(2) *If $u \in h^1(D)$ then $\lim_{r\to 1^-} u(re^{it}) = \lim_{r\to 1^-} u_r(t) = f(t)$ almost everywhere $t \in \mathbb{T}$, where f is the Radon-Nikodym derivative of the absolutely continuous component of the measure $\mu \in \mathcal{M}(\mathbb{T})$ such that $u = P_r * \mu$. More precisely, if $\mu = \mu_{ac} + \mu_s$ is the Lebesgue decomposition of μ in terms of its absolutely continuous and singular components, with respect to the Lebesgue measure on $\mathbb{T}$, then $d\mu_{ac} = \frac{1}{2\pi} f dt$.*

Proof.

(1) We know, from Theorem 1.4 in Lesson 23 that, for $1 \leq p < \infty$, $u \in h^p(D)$ implies that there exists a unique $f \in L^p(\mathbb{T})$ such that u is its Poisson integral, i.e. $u = P_r * f$. The pointwise almost everywhere existence of the limit as $r \to 1^-$ is then a consequence of Theorem 1.4 above.
(2) Again, from Theorem 1.4 in Lesson 23, we know that $u \in h^1(D)$ corresponds to the Poisson integral of a unique Borel measure $\mu \in \mathcal{M}(\mathbb{T})$, $u = P_r * \mu$. We can then do the Lebesgue-Radon-Nikodym decomposition of μ with respect to the Lebesgue measure on $\mathbb{T}$, so that $\mu = \mu_{ac} + \mu_s$ and the absolutely continuous component satisfies $d\mu_{ac} = \frac{1}{2\pi} f dt$, with $f \in L^1(\mathbb{T})$. We therefore have $P_r * \mu = Pr * \mu_{ac} + P_r * \mu_s = P_r * f + P_r * \mu_s$, and for the first of these integrals Theorem 1.4 above yields $\lim_{r\to 1^-} Pr * \mu_{ac} = \lim_{r\to 1^-} P_r * f = f$ almost everywhere $t \in \mathbb{T}$. We will leave it as an exercise to show that $\lim_{r\to 1^-} Pr * \mu_s = 0$.

□

The problem of pointwise limits of harmonic and analytic functions on D as we approach the boundary ∂D is a classical and important one. Here we answered it for limits at fixed angle, but more general results for limits of harmonic functions in $h^p(D)$ along directions that approach the boundary nontangentially can also be proved (see Rudin [4] or Hoffman [1]).

Although we still do not know whether a function $f \in L^p(\mathbb{T})$ has a conjugate function also in $L^p(\mathbb{T})$, i.e. a boundary function for the harmonic conjugate $v = Q_r * f$ of its Poisson integral $u = P_r * f$, in D, we will now prove that, nevertheless, a pointwise limit for the harmonic conjugate $v = Q_r * f$ of any $f \in L^1(\mathbb{T})$ always exists.

Theorem 1.6. *Let $f \in L^1(\mathbb{T})$ and consider the harmonic conjugate of its Poisson integral $u = P_r * f$ in D, that vanishes at the origin $z = 0$,*

$$v(re^{it}) = v_r(t) = Q_r * f(t) = \frac{1}{2\pi} \int_{\mathbb{T}} \frac{2r\sin(t-s)}{1 - 2r\cos(t-s) + r^2} f(s)ds = -i \sum_{n=-\infty}^{\infty} \operatorname{sgn}(n) r^n \hat{f}(n) e^{int}.$$

Then, the limit $\lim_{r\to 1^-} v_r(t)$ exists for almost all $t \in \mathbb{T}$.

Proof. The map $f \to v = Q_r * f$ is linear, and f can be decomposed $f = f_1 - f_2 + i(f_3 - f_4)$ where the f_i are nonnegative. So, without loss of generality, we can restrict the proof to $f \geq 0$. Now, the function $F(z) = e^{-(u(z)+iv(z))}$, where $u = P_r * f$ and $v = P_r * f$ is holomorphic in D and thus also harmonic. But, as $P_r \geq 0$ and we are also assuming $f \geq 0$, then $u = P_r * f \geq 0$ and $|F(z)| = e^{-u(z)} \leq 1$. So F is a bounded harmonic function on D, i.e. $F \in h^\infty(D)$. But then, from Part (1) of Theorem 1.5 we know that $\lim_{r\to 1^-} F(re^{it})$ exists for almost all $t \in T$, call it $F_1(t) \in L^\infty(\mathbb{T})$, such that $F_r(t) = F(re^{it}) = P_r * F_1(t)$. In particular $\lim_{r\to 1^-} |F(re^{it})| = e^{-f(t)} = |F_1(t)| \neq 0$ almost everywhere, because $f \in L^1(\mathbb{T})$ is finite almost everywhere. So, necessarily $F_1(t)$ must have well defined argument almost everywhere, which means $\lim_{r\to 1^-} v_r(t)$ exists almost everywhere. □

So, somewhat surprisingly we have established the existence of pointwise limits for the harmonic conjugate $v = Q_r * f$ of the Poisson integral $u = P_r * f$ of every $f \in L^1(\mathbb{T})$. This seems paradoxical as we already know that $L^1(\mathbb{T})$ does not admit conjugation because Fourier series do not converge in $L^1(\mathbb{T})$. The issue, though, is that this pointwise limit of the harmonic conjugate might very well not be an $L^1(\mathbb{T})$ function. Because, if given $f \in L^1(\mathbb{T})$ the conjugate distribution $\tilde{f} \in \mathcal{D}'(\mathbb{T})$ is actually a function $\tilde{f} \in L^1(\mathbb{T})$, in the sense of the definition of Lesson 21, i.e. the distribution corresponding to the Fourier multiplier operator $\{-i\,\mathrm{sgn}\,(n)\}_{n\in\mathbb{Z}}$ such that its Fourier coefficients satisfy $\hat{\tilde{f}}(n) = -i\,\mathrm{sgn}\,(n)\hat{f}(n)$, then the harmonic conjugate $v = Q_r * f$ will be the Poisson integral of $\tilde{f}$, $v = P_r * \tilde{f}$ and therefore, from Theorem 1.4 above, $\tilde{f}(t) = \lim_{r\to 1^-} v_r(t)$ almost everywhere. So, if the conjugate of any $f \in L^1(\mathbb{T})$ is a function $\tilde{f} \in L^1(\mathbb{T})$, it will have to coincide almost everywhere with the pointwise limit of $v = Q_r * f$.

Definition 1.7. *Let $f \in L^1(\mathbb{T})$. We define the Hilbert transform of f, and denote it by $\mathcal{H}(f)$ the function defined almost everywhere on $\mathbb{T}$ by the pointwise limit, as $r \to 1^-$ of the harmonic conjugate $v = Q_r * f$ of the Poisson integral $u = P_r * f$*

$$\mathcal{H}(f)(t) = \lim_{r\to 1^-} Q_r * f(t).$$

With this definition, the previous observations can then be stated as.

Proposition 1.8. *Let $f \in L^1(\mathbb{T})$. Then if the conjugate of f, defined as the distribution $\tilde{f} \in \mathcal{D}(\mathbb{T})$ defined by the Fourier multiplier operator $\hat{\tilde{f}}(n) = -i\,\mathrm{sgn}\,(n)\hat{f}(n)$, satisfies $\tilde{f} \in L^1(\mathbb{T})$ then*

$$\mathcal{H}f(t) = \tilde{f}(t),$$

almost everywhere $t \in \mathbb{T}$.

It is tempting to use (1.6), with the fact that $\lim_{r\to 1^-} \frac{2r\sin(t)}{1-2r\cos(t)+r^2} = \frac{1}{\tan t/2}$, to say that

$$\mathcal{H}f(t) = \frac{1}{2\pi}\int_{\mathbb{T}} \frac{f(t-s)}{\tan\frac{s}{2}}ds.$$

The problem, of course, is that $\frac{1}{\tan t/2}$ is not locally integrable in a neighborhood of the origin, and this convolution does not make sense as an $L^1(\mathbb{T})$ integral with such a singular kernel. However, $\frac{1}{\tan t/2}$ is an odd function, and a cancellation effect of its signs around the origin can be exploited in what is called the Cauchy Principal Value. A careful study of the pointwise limit $\lim_{r\to 1^-} Q_r * f(t)$, that we will not do here (see, for example, Katznelson [2, 3]), shows that indeed this is the appropriate way to define the singular kernel operator for the Hilbert transform

$$\mathcal{H}f(t) = \lim_{r\to 1^-} Q_r * f(t) = \lim_{\varepsilon\to 0} \frac{1}{2\pi}\int_{|s|>\varepsilon} \frac{f(t-s)}{\tan\frac{s}{2}}ds = \frac{1}{2\pi}\text{p.v.}\int_{\mathbb{T}} \frac{f(t-s)}{\tan\frac{s}{2}}ds,$$

where "p.v." stands for Principal Value of Cauchy and is defined precisely as the limit of the integrals as a removed symmetric interval around the origin shrinks to zero.

So, although we know now that, when the conjugate exists as an $L^1(\mathbb{T})$ function it coincides with the Hilbert transform, the converse is not obvious. In other words, if for $f \in L^1(\mathbb{T})$ we have $\mathcal{H}f \in L^1(\mathbb{T})$ this only means that the pointwise limit $\lim_{r\to 1^-} v_r(t) = \lim_{r\to 1^-} Q_r * f(t)$ of v at the boundary ∂D is in $L^1(\mathbb{T})$ but from this pointwise limit we *cannot*, at the moment, conclude that v is given by the Poisson integral $v = P_r * \mathcal{H}f$ in order to conclude that $\tilde{f} = \mathcal{H}f \in L^1(\mathbb{T})$. We do know that if such a limit exists in the $L^p(\mathbb{T})$ norm, then it implies that v is the Poisson integral with that limit, but with what we have so far, an analogous conclusion does not follow from pointwise limits only. In fact, one just needs to think about the Poisson kernel itself $P_r(t)$ as an example of a harmonic function on the disk D whose limit $\lim_{r\to 1^-} P_r = 0$ almost everywhere, but of course $P_r \neq P_r * 0$. Or, more generally, Part (2) of Theorem 1.5 to obtain many examples of harmonic functions in $h^1(D)$ whose boundary pointwise limits "only see" the $L^1(\mathbb{T})$ Radon-Nikodym derivative of the absolutely continuous component of the Lebesgue decomposition of the boundary measure, but are "blind" to the singular measure component (which is precisely what happens with the Poisson kernel, whose boundary value is the Dirac-δ at the origin). So it could very well happen that, for $f \in L^1(\mathbb{T})$ the conjugate $\tilde{f}$ is only a distribution, for example a measure with nonzero singular component, while the Hilbert transform is in $L^1(\mathbb{T})$ because it only extracts the Radon-Nikodym derivative of the absolutely continuous component of the measure. In particular, we cannot deduce that the spaces $L^p(\mathbb{T})$ admit conjugation - and from it conclude the convergence of Fourier series in $L^p(\mathbb{T})$ - by just looking exclusively at whether the Hilbert transform is in $L^p(\mathbb{T})$ for $f \in L^p(\mathbb{T})$.

So we will really look at the conjugation problem from the interior of the disk D and not only by looking at the pointwise boundary values. The following is one of the central results on which the conjugation in $L^p(\mathbb{T})$ hinges.

Theorem 1.9. (Besicovitch, Kolmogorov) *Let $f \in L^1(\mathbb{T})$. Then, the map $f \to v_r = Q_r * f$ is weak type $(1,1)$ uniformly in r, i.e. there exists a constant C that does not depend on r (it can be made $C = 128$) for which*

$$\lambda_{v_r}(\alpha) = |\{t \in \mathbb{T} : |v_r(t)| > \alpha\}| \leq C\frac{\|f\|_{L^1(\mathbb{T})}}{\alpha},$$

for $\alpha > 0$. In particular, these uniform bounds imply $\mathcal{H}f \in L^1_w(\mathbb{T})$ and that the Hilbert transform is weak type $(1,1)$.

Proof. We start by assuming $f \geq 0$ and with normalized $L^1(\mathbb{T})$ norm $\|f\|_{L^1(\mathbb{T})} = 1$.

The proof relies on a smart use of a particular harmonic function. If, for $\alpha > 0$, we consider the function

$$H_\alpha(z) = 1 + \frac{1}{\pi}\arg\frac{z - i\alpha}{z + i\alpha} = 1 + \frac{1}{\pi}\mathrm{Im}\left(\log\frac{z - i\alpha}{z + i\alpha}\right) = 1 - \frac{1}{\pi}\arctan\frac{2\alpha x}{x^2 + y^2 - \alpha^2},$$

then it is harmonic and positive on the right half plane of $\mathbb{C}$, of complex numbers with positive real part $x = \mathrm{Re}(z) > 0$. The level line $H_\alpha(z) = \frac{1}{2} \Leftrightarrow \arg\frac{z-i\alpha}{z+i\alpha} = -\frac{\pi}{2} \Leftrightarrow \arctan\frac{2\alpha x}{x^2+y^2-\alpha^2} = \frac{\pi}{2} \Leftrightarrow x^2 + y^2 - \alpha^2 = 0$ is the half circle $z = \alpha e^{i\theta}$ with $-\pi/2 < \theta < \pi/2$, so that $H_\alpha(z) > \frac{1}{2}$ when z is outside the half circle, $|z| > \alpha$. Also, for $z = 1$, $H_\alpha(1) = 1 - (2/\pi)\arctan\alpha < \frac{2}{\pi\alpha}$.

So, considering $u = P_r * f$ and its harmonic conjugate $v = Q_r * f$ and composing the harmonic function $H_\alpha(z)$ with the holomorphic $u(z) + iv(z)$ on D we obtain the harmonic function on D, $H_\alpha(u(z) + iv(z))$ which is nonnegative. Notice that, as $f \geq 0$ we have $u = P_r * f \geq 0$ and that, actually, u cannot be zero in the interior of D, so that for $z \in D$ the complex number $u(z) + iv(z)$ will be in the right half plane where H is positive. We can thus apply the mean value theorem for harmonic functions to obtain, from the estimates of the previous paragraph for H_α,

$$\frac{1}{2\pi}\int_{\mathbb{T}} H_\alpha(u(re^{it}) + iv(re^{it}))dt = H_\alpha(f(0)) = H_\alpha(1) < \frac{2}{\pi\alpha}, \tag{1.2}$$

for all $0 < r < 1$. We will now use this estimate to obtain the weak L^1 bound for v_r. In fact, if $|v_r(t)| = |v(re^{it})| > \alpha$ then $|u(re^{it}) + iv(re^{it})| > \alpha$ and we know that outside of the half circle of radius

α, $H_\alpha(u(re^{it})+iv(re^{it}))>\frac{1}{2}$. So that, for (1.2) to hold

$$\frac{1}{4\pi}|\{t\in\mathbb{T}:|v_r(t)|>\alpha\}|\le\frac{1}{2\pi}\int_{\{t\in\mathbb{T}:|v(re^{it})|>\alpha\}}H_\alpha(u(re^{it})+iv(re^{it}))dt\le\frac{2}{\pi\alpha},$$

and we obtain the estimate for the distribution function of v_r

$$\lambda_{v_r}(\alpha)=|\{t\in\mathbb{T}:|v_r(t)|>\alpha\}|\le\frac{8}{\alpha}.$$

The end of the proof now follows easily. If we consider nonnormalized functions $f\ge 0$ then this weak L^1 estimate becomes a weak type $(1,1)$ bound

$$\lambda_{v_r}(\alpha)=|\{t\in\mathbb{T}:|v_r(t)|>\alpha\}|\le\frac{8\|f\|_{L^1(\mathbb{T})}}{\alpha}.$$

And finally, for any complex $f\in L^1(\mathbb{T})$ we decompose, as in the proof of Theorem 1.6, $f=f_1-f_2+i(f_3-f_4)$ so that $\lambda_{v_r}(\alpha)\le\sum_{j=1}^4\lambda_{v(j)_r}(\alpha/4)$ to obtain

$$\lambda_{v_r}(\alpha)\le\frac{128\|f\|_{L^1(\mathbb{T})}}{\alpha}.$$

□

If we recall the example of the Fourier series $\sum_{n\ge2}\frac{\cos nt}{\log n}$ whose conjugate trigonometric series $\sum_{n\ge2}\frac{\sin nt}{\log n}$, although convergent at every $t\in\mathbb{T}$ is not a Fourier series, it all becomes clear now. The conjugate function is not in $L^1(\mathbb{T})$, but it is in $L^1_w(\mathbb{T})$.

We can finally conclude with Marcel Riesz's theorem on the conjugation problem in $L^p(\mathbb{T})$. Once we have a weak type $(1,1)$ estimate and the strong $(2,2)$ coming from the $L^2(\mathbb{T})$ theory, we just need to apply the Marcinkiewicz interpolation theorem to obtain the strong type (p,p) conclusion for the conjugation operator for intermediate $1<p<2$. The remaining exponents are obtained by duality.

Theorem 1.10. (M. Riesz) *Let $1<p<\infty$. Let $f\in L^p(\mathbb{T})$, then its conjugate $\tilde f$ exists and the map $f\mapsto\tilde f$ is a bounded linear operator in $L^p(\mathbb{T})$. Also, in this case $\mathcal{H}f=\tilde f$ and $\|\mathcal{H}f\|_{L^p(\mathbb{T})}\le C_p\|f\|_{L^p(\mathbb{T})}$.*

Proof. We have already seen in Lesson 22 that conjugation holds (strongly) in $L^2(\mathbb{T})$ because the conjugation Fourier multiplier $\{-i\,\mathrm{sgn}\,n\}$ is bounded, i.e. in l^∞ , and therefore $\|\tilde f\|_{L^2(\mathbb{T})}\le\|f\|_{L^2(\mathbb{T})}$. Or, equivalently, in terms of the harmonic conjugate of the Poisson integral of f,

$$\|v_r\|_{L^2(\mathbb{T})}=\|P_r*\tilde f\|_{L^2(\mathbb{T})}=\|Q_r*f\|_{L^2(\mathbb{T})}\le\|f\|_{L^2(\mathbb{T})}.$$

We can now interpolate between this strong type $(2,2)$ estimate and the weak type $(1,1)$ estimate coming from Kolmogorov's theorem, using Marcinkiewicz, to obtain the strong bound at every $1<p<2$

$$\|v_r\|_{L^p(\mathbb{T})}=\|Q_r*f\|_{L^p(\mathbb{T})}\le C_p\|f\|_{L^p(\mathbb{T})},$$

for some constant C_p and uniformly in $0<r<1$. We thus conclude that $v\in h^p(D)$ and therefore, from the characterization Theorem 1.4 in Lesson 23, that there exists $\tilde f\in L^p(\mathbb{T})$ for which $v=P_r*\tilde f$, with $\|\tilde f\|_{L^p(\mathbb{T})}\le C_p\|f\|_{L^p(\mathbb{T})}$.

For $2\le p<\infty$ we use duality. Let $f,g\in L^2(\mathbb{T})$. Then, from Parseval's identity, Theorem 1.11, Part (3), in Lesson 16, we have

$$\langle f,\tilde g\rangle=\frac{1}{2\pi}\int_{\mathbb{T}}f(t)\overline{\tilde g(t)}dt=\sum_{n=-\infty}^{\infty}\hat f(n)\overline{-i\,\mathrm{sgn}\,(n)\hat g(n)}=\sum_{n=-\infty}^{\infty}i\,\mathrm{sgn}\,(n)\hat f(n)\overline{\hat g(n)}=\langle-\tilde f,g\rangle.$$

Let now $f \in L^p(\mathbb{T}) \subset L^2(\mathbb{T})$ for $p > 2$ and $g \in L^2(\mathbb{T}) \subset L^{p'}(\mathbb{T})$, with p' the conjugate exponent of p, $1/p + 1/p' = 1$. Then, the conjugate of f surely exists and it is, at least, an $L^2(\mathbb{T})$ function. What we just need to show is that it is actually in $L^p(\mathbb{T})$. But, from the previous computations we have

$$\left| \frac{1}{2\pi} \int_{\mathbb{T}} \tilde{f}(t)\overline{g(t)}dt \right| = |\langle \tilde{f}, g\rangle| = |\langle f, \tilde{g}\rangle| \leq \|f\|_{L^p(\mathbb{T})}\|\tilde{g}\|_{L^{p'}(\mathbb{T})} \leq C_p\|f\|_{L^p(\mathbb{T})}\|g\|_{L^{p'}(\mathbb{T})}.$$

And because this holds for every $g \in L^2(\mathbb{T}) \subset L^{p'}(\mathbb{T})$, which includes simple functions, it implies, from our results about duality and dual norms at the beginning of the course, that

$$\|\tilde{f}\|_{L^p(\mathbb{T})} \leq C_p\|f\|_{L^p(\mathbb{T})}.$$

□

We could now also do the exact same type of interpolation as we did in Theorem 1.2 to obtain the following.

Theorem 1.11. *Let $f \in L\log L(\mathbb{T})$. Then $\mathcal{H}f \in L^1(\mathbb{T})$.*

And we finish with the convergence in $L^p(\mathbb{T})$ of Fourier series, which is a direct consequence of Marcel Riesz's theorem for the conjugation operator in $L^p(\mathbb{T})$, $1 < p < \infty$.

Corollary 1.12. *Let $1 < p < \infty$ and $f \in L^p(\mathbb{T})$. Then, the partial sums of the Fourier series of f converge to f in the $L^p(\mathbb{T})$ norm, i.e.*

$$\left\| \sum_{n=-N}^{N} \hat{f}(n)e^{int} - f \right\|_{L^p(\mathbb{T})} \to 0 \quad \textit{as} \quad N \to \infty.$$

References

[1] Kenneth Hoffman *Banach Spaces of Analytic Functions,* Dover Publications, 1988.
[2] Yitzhak Katznelson *An Introduction to Harmonic Analysis,* 2nd Edition, Dover Publications, 1976.
[3] Yitzhak Katznelson *An Introduction to Harmonic Analysis,* 3rd Edition, Cambridge University Press, 2004.
[4] Walter Rudin *Real and Complex Analysis,*