Fourier Analysis

Dr. Felix Schwenninger
University of Hamburg

This is a preliminary (and not-proof-read) version of the lecture notes of the course *Fourier Analysis* that took place at the University of Hamburg in the summer term 2018. Up to a few exception (as for example Ex. 1.3 in the introduction and Section I.7) this manuscript contains most of the material presented in the weekly 2h-lectures. There may and will still be numerous typos and misprints and the file will regularly be updated. As mentioned above this cannot serve as a full replacement to the notes you may have taken in the class, although I tried to incorporate many comments and remarks I made during the lectures. Furthermore, note that the exercises (and their solutions) complement the content presented in the lecture in an essential way. I revised the numbering of the theorems, lemmata, etc in order to match them as good as possible with the numbering I used in the class — as for the contents, this should now match to a large extent, but there are a few exceptions.

Summarizing, this script should be seen as supplementary support and will hopefully serve useful in the preparation for the exam. If you find typos or other mistakes / things that seem to be incorrect, please do not hesitate to let me know. I would be happy to revise these notes.

This lecture notes as well as the exercise sheets can be found at

`https://www.math.uni-hamburg.de/home/schwenninger/Vo_FoAn.html`

Hamburg, July 18, 2018

CHAPTER 0

Introduction

1. What "is" Fourier Analysis?

The point of departure is that we consider a signal, represented by a function $f : \mathbb{R} \to \mathbb{C}$, which we would like to represent and/or approximate.

EXAMPLE 1.1 (Thresholding/Filtering and Fourier series of periodic functions and a reminder on Hilbert spaces). Consider a signal which is 2π-periodic and hence determined by a function $f : [0, 2\pi] \to \mathbb{C}$. Shortly, we we will see that to such a function (under the condition that f is integrable) we can attribute another function, defined by the formal infinite series

$$t \mapsto S(f)(t) := \sum_{n=-\infty}^{\infty} \hat{f}(n)\mathrm{e}^{-int} = \lim_{N\to\infty} \sum_{n=-N}^{N} \hat{f}(n)\mathrm{e}^{-int}$$

Suppose that $f(t) = S(f)(t)$ for all $t \in [0, 2\pi]$ [1] — in fact this is a highly non-trivial property of f and at the heart of the relation between $S(f)$ and f. Representing e^{it} by cos and sin indicates that (the coefficients of) $S(f)$ encode the the contributions of oscillations at fixed frequencies in the signal. When applying filtering or de-noising one may be interested in cancelling certain frequency bands in a signal. Mathematically speaking this may be achieved by "thresholding" the coeffients $\hat{f}(n)$ in $S(f)$: For fixed $\epsilon > 0$, let for all $n \in \mathbb{Z}$,

$$c_n = \begin{cases} \hat{f}(n) & \text{if } |\hat{f}(n)| > \epsilon \\ 0 & \text{else} \end{cases}$$

and define $\tilde{S} = \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{-int}$.
An important question is

How does $\tilde{S}$ relate to $S = f$?

To find a first answer to this, let us additionally assume that f lies in the Hilbert space $L^2(0, 2\pi)$ of square-integrable functions (which clearly is the case if f was continuous). In fact, in the view of basic Hilbert space theory, it is not at all surprising to expect a close relation between f and an element of the form $\sum \hat{f}(n)\mathrm{e}^{-int}$ with special coefficients $\hat{f}(n) \in \mathbb{C}$: Since $(\phi_n) = (\frac{1}{2\pi}\mathrm{e}^{-int})_{n\in\mathbb{Z}}$ defined an orthonormal basis of L^2 (i.e. a complete orthonormal system), we find that $\hat{f}(n) =$ is given by $(f, \phi_n)_{L^2}$, where $(\cdot, \cdot)_{L^2}$ denotes the inner product of $L^2(0, 2\pi)$. By Parseval's identity, we even have

$$\|f\|_{L^2}^2 = \sum_{n\in\mathbb{Z}} |(f, f_n)|^2 = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2$$

and — since $(c_n)_{n\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$ as $(\hat{f}(n))_{n\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$ — also $S(f) - \tilde{S} \in L^2$ with

$$\|S(f) - \tilde{S}\|_{L^2}^2 = \sum_{n\in\mathbb{Z}} |\hat{f}(n) - c_n|^2 = \sum_{n\in\mathbb{Z}, |\hat{f}(n)|>\epsilon} |\hat{f}(n)|^2.$$

[1] in fact, we will see that in general it is only meaningful to assume that f and $S(f)$ coincide almost everywhere with respect to the Lebesgue measure. For the moment let us be a bit vague here.

This identity shows the "deviation" of $f = S(f)$ and $\tilde{S}$ in terms of the "energy-norm" L^2.

EXAMPLE 1.2 (Convergence of Fourier series). Going back to the beginning of the previous example, we reconsider the then-made assumption $f = S(f)$: It turns out that for general functions f, even if we assume that f is continuous, this question is much harder to answer than one may expect after having seen the subsequent procedure in the above example (for more general L^2-functions). In fact, the above mentioned representation by Parseval's identity only says that f equals the L^2-limit $\lim_{N\to\infty} \sum_{n=-N}^{N} \hat{f}(n)\mathrm{e}^{-int}$ in the sense of L^2. This does not imply that $S(f)(t) = f(t)$ for all $t \in [0, 2\pi]$ and not even that this identity holds for almost all t, see Ex. 0.1. Therefore, the following question still remains

Does $S(f)$ equal f pointwise (almost everywhere)? And if not, when does this hold?

Answer: for $f \in L^2$ this is indeed true and subject to a famous result by L. Carleson [2]

EXAMPLE 1.3 (Fourier's investigation on heat flow). see notes

2. The main questions and goals

Q: Does the Fourier series of a function f converge to f pointwise/in (some) norm?

3. Prerequisites and Literature

In this course we will strongly rely on concepts from measure theory (e.g. Lebesgue integration, L^p spaces,..), functional analysis (e.g. bounded linear operators on Banach spaces, uniform boundedness principle, Hahn-Banach, closed graph, basic Hilbert space theory,..) — it should go without saying that fundamentals from analysis (Bachelor courses including Analysis III) will be required too. If any of the mentioned keywords do not "ring a bell" you better consult relevant literature, such as for example Tao's book "An Epsilon of Room, I: Real Analysis" [**4**], available as online version (see [3]) as well as in the library. Alternatively you may consider looking into Analysis III scripts from past courses at the University of Hamburg.

Further reading: Among the vast literature on Fourier analysis, the books by Y. Katznelson ('An Introduction to Harmonic Analysis', Cambridge University Press) [**2**] and L. Grafakos ('Classical Fourier Analysis', Springer Graduate Texts in Mathematics) [**1**] shall be mentioned here. The first part of this course on Fourier series we will loosely follow the corresponding part of Katznelson's book, whereas the introduction to the Fourier transform — the second part — is adapted from [**1**].

[2]moreover, this result was latter shown to generalize to any $f \in L^p(0, 2\pi)$ for $p \in (1, \infty)$.

[3]`https://terrytao.wordpress.com/books/an-epsilon-of-room-pages-from-year-three-of-a-mathematical-blog/`

CHAPTER 1

Fourier Series

This chapter deals with the interplay of 2π-periodic functions $f : \mathbb{R} \to \mathbb{C}$ and associated (formal) objects

$$\sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{int} \tag{0.1}$$

with which we aim to represent the functions f. Sometimes it will be convenient to view such functions as defined on the torus $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ or — upon the isomorphism $t \mapsto \mathrm{e}^{it}$ — on $\{z \in \mathbb{C} \colon |z| = 1\}$. Furthermore, we can identify 2π-periodic functions on $\mathbb{R}$ with functions defined on $[0, 2\pi)$. We will make use of the later identification in the following without specifying. Note that this particularly means that for $f : \mathbb{R} \to \mathbb{C}$ Lebesgue measurable (and integrable on $[0, 2\pi)$,

$$\int_{\mathbb{T}} f(t)\, dt = \int_0^{2\pi} f(x)\, d\lambda(x),$$

where λ denotes the Lebesgue measure.

The presented theory generalizes to any other (positive) period by obvious transformations (in the literature often the 1-periodic functions are discussed).

1. Definitions and basics

Definition 1.2. An object of the form (0.1) is called a **trigonometric series**, and in particular **trigonometric polynomial** if $a_n = 0$ for almost all n.

From the analysis lectures we recall the following fundamental result[1]

> *For $f \in C([a, b]; \mathbb{C})$ there exists a sequence of polynomials $(p_n)_{n\in\mathbb{N}}$ (with complex coefficients) such that $\lim_{n\to\infty} \|f - p_n\|_{\infty,[a,b]} = 0$.*

In particular, this yields

Theorem 1.3 (Weierstrass' approximation theorem). *The trigonometric polynomials lie dense in $C(\mathbb{T})$.*

We remark that there are several ways to prove Weierstass' theorem, among which is a "modern approach" via the more general Stone–Weierstrass theorem which we now have in mind when we take this theorem for granted. However, we will shortly face a constructive proof based on Fourier series.

Having in mind the goal to represent a function f by a trigonometric series, that is, we would like to link a (preferably unique) trigonometric series to f,

$$f \sim \sum_n a_n \mathrm{e}^{int},$$

[1] In these lectures, one may have only seen the 'real version' of this theorem. The 'complex version' however follows from that case by applying it to real- and imaginary part of the considered continuous function.

(this notation is, for the moment, nothing else than formal) and recalling the following fundamental identity for $k \in \mathbb{Z}$

$$\int_0^{2\pi} \mathrm{e}^{ikt}\, dt = \begin{cases} 2\pi & k = 0 \\ 0 & k \neq 0 \end{cases} \tag{1.1}$$

we (formally) derive

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} f(t)\mathrm{e}^{-int}\, dt. \tag{1.2}$$

If f was a trigonometric polynomial (or more generally, if the trigonometric sum converges uniformly), then this formula is indeed justified. From now on and unless stated otherwise, we will assume that functions f are in $L^1(\mathbb{T})$, in which case the integral in (1.2) exists.

Definition 1.4. For a 2π-periodic function $f \in L^1(\mathbb{R})$, the sequence $(a_n)_{n\in\mathbb{Z}}$ defined by (1.2) is called the **Fourier coefficients** of f and we denote it by $\hat{f}(n) = a_n$. The trigonometric series

$$S(f) = \sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{int} = (t \mapsto \sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{int})$$

is called the **Fourier series** of f and we also use the notation $f \sim \sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{int}$. Given a (formal) trigonometric series S of the form (0.1), we call it a Fourier series if there exists an $f \in L^1$ such that S is the Fourier series of f.

Remark 1.5. Trigonometric polynomials take — in some sense — over the role of usual polynomials but on the unit sphere. More generally, trigonometric series take over the role of power series. In that sense, Fourier series can be compared to Taylor series of a function. However, while Taylor series exploit the local behavior of a function, Fourier series encode the *global* behavior, as we will see in this course. The latter fact makes them very powerful on one hand, but far more difficult in study as we shall see.

Theorem 1.6 (Basic facts on Fourier series). *The following holds.*

(1) $f \mapsto \hat{f}$ *is a linear, injective mapping from* $L^1(\mathbb{T})$ *to* $\mathbb{C}^{\mathbb{Z}}$

(2) $f \mapsto \hat{f}$ *bounded from* $L^1(\mathbb{T})$ *to* $\ell^\infty(\mathbb{Z})$*, the space of bounded, complex-valued (double) sequences, i.e.*

$$\|\hat{f}\|_\infty := \sup_{n\in\mathbb{Z}} |\hat{f}(n)| \leq \frac{1}{2\pi}\|f\|_{L^1(\mathbb{T})}.$$

(3) *for any* $s \in \mathbb{R}$, $\widehat{\tau_s f} = (s \mapsto \mathrm{e}^{is}\hat{f}(s))$*, where* $(\tau_s f)(t) = f(t+s)$*,* *(Modulation)*

Proof. The linearity follows directly form (1.2). We will present a proof of the injectivity (and shall see another one in the Exercise classes). Let us assume that $\hat{f} = 0$. By linearity it suffices to show that that $f = 0$. Again by linearity and the definition of $\hat{f}$, we have that

$$\int_0^{2\pi} f(t)g(t)\, dt = 0 \tag{1.3}$$

for all trigonometric polynomials g[2] and hence for all continuous functions g by Weierstrass' theorem. To conclude that $f = 0$ we have several alternatives:

[2]If f is continuous, we can conclude that $f = 0$ by Weierstrass' theorem. For more general $f \in L^1(\mathbb{T})$, we need a more refined argument

(α) If we can show that for any measurable set $E \subset [0, 2\pi]$ it holds that $\int_E f(t)\,dt = 0$, then it is not hard to see that $f = 0$. In fact, if $f \neq 0$, then either $\Re f \neq 0$ or $\Im f \neq 0$ (in the L^1 sense). Let us assume that there exists ε such that $E = \{t \in [0, 2\pi] \colon \Re f(t) > \varepsilon\}$ has positive measure — the other cases follow analogously. Then $\Re \int_E f(t)\,dt > \varepsilon\lambda(E)$ which yields a contradiction. In order to show that $\int_E f(t)\,dt = 0$ for any measurable set E, recall *Lusin's theorem*[3]

Let $h : [a, b] \to \mathbb{C}$ be a measurable function. Then for any $\delta > 0$ there exists a continuous function $h_\delta : [a, b] \to \mathbb{C}$ such $\lambda(\{t \in [a, b] \colon |h(t) - h_n(t)| \neq 0\}) < \delta$ with $\|h_\delta\|_\infty \leq \|h\|_\infty$ [4]

Apply this with $h = \chi_E$ and set $g_n := h_{2^{-n}}$ and conclude that $\lim_{n\to\infty} g_n(t) = g(t)$ for almost every $t \in [a, b]$ [5]. As $\|g_n\|_\infty \leq 1$, dominated convergence together with (1.3) yields

$$0 = \lim_{n\to\infty} \int_0^{2\pi} f(s) g_n(s)\,ds = \int_0^{2\pi} f(s) g(s)\,ds = \int_E f(s)\,ds.$$

(β) For any function g in $L^\infty(\mathbb{T})$ there exists a sequence of continuous functions that converge to g in weak* sense (see e.g. Ex. 2.3), and consequently

$$\int_0^{2\pi} f(t) g(t)\,ds = 0 \quad \forall g \in L^\infty(\mathbb{T}).$$

Using the fact that $L^\infty(\mathbb{T})$ is isometrically isomorph to the dual space of $X = L^1(\mathbb{T})$ (Riesz Representation theorem), this implies that $\langle f, g\rangle_{X,X'} = 0$ for all $g \in X'$. Hence, a consequence of the Hahn-Banach theorem yields $f = 0$.

□

THEOREM 1.7 (discrete Riemann–Lebesgue). *$f \mapsto \hat{f}$ maps $L^1(\mathbb{T})$ to $c_0(\mathbb{Z})$, where $c_0(\mathbb{Z})$ denotes the space of sequences that converge to 0 (as $n \to \pm\infty$).*

PROOF. We present a sketch of the proof (the details of each step are left for the exercises). This follows the routine

prove for smooth + functional analysis

First, one shows (sees) that for any trigonometric polynomial p it holds that $\hat{p} \in c_0(\mathbb{Z})$. As a second step, use that $f \mapsto \hat{f}$ is bounded from $L^1(\mathbb{T})$ to $\ell^\infty(\mathbb{Z})$, Theorem 1.6 and conclude the assertion by density of the trigonometric polynomials (Weierstrass' Theorem) and density of continuous functions in $L^1(\mathbb{T})$. □

PROPOSITION 1.8. *The operator $f \mapsto \hat{f}$ is not surjective from $L^1(\mathbb{R})$ to $c_0(\mathbb{Z})$.*

PROOF. See Ex. 1.3 □

PROPOSITION 1.9. *Let $s \in \mathbb{T}$ and define the operator τ_s on $L^1(\mathbb{R})$ by $\tau_s f = f(\cdot + s)$. Then for all $n, m \in \mathbb{Z}$ and $f \in L^1(\mathbb{R})$,*

(1) $\widehat{\tau_s f}(n) = \mathrm{e}^{ins} \hat{f}(n)$,
(2) $\widehat{\mathrm{e}^{im\cdot} f}(n) = \hat{f}(n - m)$.

---

[3] See for instance Tao's book (the condition on the supremum follows easily from the proof) mentioned in the introduction.

[4] where the essential supremum on left hand side of the inequality is allowed to be ∞.

[5] here, one may use that $\lambda(\{t \colon h(t) - h_n(t) \neq 0 \text{ for infinitely many } n \in \mathbb{N}\}) = 0$ by Borel-Cantelli

PROOF. To see (1), note that

$$\widehat{\tau_s f}(n) = \frac{1}{2\pi}\int_{\mathbb{T}} f(t+s)\mathrm{e}^{-int}\,dt = \frac{\mathrm{e}^{ins}}{2\pi}\int_{\mathbb{T}} f(t)\mathrm{e}^{-int}\,dt = \mathrm{e}^{ins}\hat{f}(n).$$

The other assertion follows similarly. □

REMARK 1.10 (a warning). One may be tempted to think that Weierstrass' theorem, Theorem 1.3, already gives an indication on functions that can 'safely' be approximated by Fourier series: the continuous functions on $\mathbb{T}$, as FOURIER himself conjectured. This however turns out to be ('very') wrong as we shall see. In a similar, but different, spirit, we convince ourselves that the approximating sequence in Weierstrass' theorem is in general not given by the partial sums of a Fourier series.

2. Convolutions

In the following let Ω be either $\mathbb{R}$ or $\mathbb{T}$ and consider $L^p = L^p(\Omega) = L^p(\Omega, \lambda)$ where $d\lambda(s) = ds$ refers to the corresponding Lebesgue measure. In the present chapter we shall be interested in the case $\Omega = \mathbb{T}$, but later on the case $\Omega = \mathbb{R}$ will be needed.

DEFINITION 2.1 (Convolution). Let $f, g \in L^1(\Omega)$, then the *convolution* $f * g$ *of* f *and* g is defined as $(f * g)(t) = \int_\Omega f(t-s)g(s)\,ds$ for $t \in \Omega$ such that the integral exists.

Formally one can define the convolution for more general functions f, g only requiring that $f * g$ exists for a.e. $t \in \Omega$. In the case of $L^1(\Omega)$ this is indeed guaranteed as the following result shows.

THEOREM 2.2. *The convolution* $*$ *is a bilinear mapping from* $L^1(\Omega) \times L^1(\Omega) \to L^1(\Omega)$ *that satisfies for all* $f, g, h \in L^1(\Omega)$*:*

$$\begin{aligned} \|f * g\|_{L^1} &\le \|f\|_{L^1}\|g\|_{L^1} \qquad (2.1)\\ f * g &= g * f \qquad \textit{(commutativity)}\\ f * (g * h) &= (f * g) * h \qquad \textit{(associativity).} \end{aligned}$$

PROOF. Bilinearity easily follows from linearity of the integral. All other assertions are based on applying Fubini's theorem. Let us show that $*$ is well-defined from $L^1(\Omega) \times L^1(\Omega)$ to $L^1(\Omega)$ and (2.1). The mapping $I(t,s) = f(t-s)g(s)$ is clearly measurable on $\Omega \times \Omega$ since f and g are measurable. Furthermore, $t \mapsto I(t,s) \in L^1(\Omega)$ for a.e. $s \in \Omega$ because $f \in L^1(\Omega)$ and by

$$\int_\Omega\int_\Omega |I(t,s)|\,dt\,ds = \int_\Omega |g(s)|\int_\Omega |f(t-s)|\,dt\,ds = \|g\|_{L^1(\Omega)}\|f\|_{L^1(\Omega)},$$

where we have used the translation invariance of the Lebesgue measure on our choices for Ω, we get that $s \mapsto \|I(\cdot,s)\|_{L^1(\Omega)}$ is in $L^1(\Omega)$. Thus, Fubini's theorem (or Fubini-Tonelli), yields that $t \mapsto (f * g)(t) = \int_\Omega I(t,s)\,ds$ is integrable and

$$\begin{aligned} \|f * g\|_{L^1(\Omega)} &= \int_\Omega \left|\int_\Omega I(t,s)\,ds\right| dt\\ &\le \int_\Omega\int_\Omega |I(t,s)|\,ds\,dt = \int_\Omega\int_\Omega |I(t,s)|\,dt\,ds = \|g\|_{L^1(\Omega)}\|f\|_{L^1(\Omega)}. \end{aligned}$$

□

Inequality (2.1) is only a special instance of the following result.

THEOREM 2.3 (Young's inequality (for convolutions)). *Let* $\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}$ *for* $p,q,r\in[1,\infty]$ *and* $f\in L^p(\Omega)$, $g\in L^q$. *Then* $f*g\in L^r(\Omega)$ *and*

$$\|f*g\|_{L^r}\le\|f\|_{L^p}\|g\|_{L^q}.$$

PROOF. See Exercise 5.4. $\square$

COROLLARY 2.4 (Minkowski's inequality). *For any* $g\in L^1(\Omega)$, *the mapping* $f\mapsto f*g$ *is bounded on* $L^p(\Omega)$ *with norm less or equal to* 1.

For a separate proof of Minkowski's inequality see Ex. 2.4.

REMARK 2.5. We have seen convolutions on L^p spaces defined on $\mathbb{R}$ and $\mathbb{T}$ with the Lebesgue measure. It is important to realize that the latter are (additive) groups with a translation-invariant measure. In fact, these are the crucial properties needed to define a convolution. We remark that this is also possible for more general *locally compact topological groups* and their, roughly speaking, translation-invariant measures.

THEOREM 2.6 (Convolution and Fourier coefficients). *Let* $f,g\in L^1(\mathbb{T})$, $n\in\mathbb{Z}$. *Then*

$$\widehat{f*g}=2\pi\ \hat{f}\cdot\hat{g},$$

where the multiplication on the right has to be understood pointwise, i.e. $\widehat{f*g}(n)=2\pi\hat{f}(n)\hat{g}(n)$. *Furthermore,*

$$(f*\mathrm{e}^{in\cdot})(t)=2\pi\hat{f}(n)\mathrm{e}^{int}.$$

PROOF. By definition of the convolution and the Fourier coefficients,

$$2\pi\widehat{f*g}(n)=\int_{\mathbb{T}}\int_{\mathbb{T}}f(t-s)g(s)\,ds\,\mathrm{e}^{-int}\,dt=\int_{\mathbb{T}}\int_{\mathbb{T}}f(t-s)\mathrm{e}^{-in(t-s)}\mathrm{e}^{-ins}g(s)\,ds\,dt.$$

Using Fubini (or Fubini-Tonelli)), we can interchange the order of integration, and also using that translation-invariance of the the Lebesgue measure on $\mathbb{T}$, we get

$$\int_{\mathbb{T}}\int_{\mathbb{T}}f(t-s)\mathrm{e}^{-in(t-s)}\mathrm{e}^{-ins}g(s)\,ds\,dt=\int_{\mathbb{T}}\mathrm{e}^{-ins}g(s)\int_{\mathbb{T}}f(t-s)\mathrm{e}^{-in(t-s)}\,dt\,ds=(2\pi)^2\hat{g}(n)\hat{f}(n).$$

The second claim follows either by a similar calculation or by recalling that $\widehat{\mathrm{e}^{in\cdot}}=e_n$, where e_n denotes the n-th canonical basis vector in $c_0(\mathbb{Z})$, and applying the first part of this theorem: Thus $\widehat{f*\mathrm{e}^{in\cdot}}=\hat{f}(n)e_n$ and by injectivity of the linear map $f\mapsto\hat{f}$, Theorem 1.6, we deduce $(f*\mathrm{e}^{in\cdot})(t)=\hat{f}(n)\mathrm{e}^{int}$. $\square$

3. Approximate identities

DEFINITION 3.1 (Approximate identity). A sequence $(k_n)_{n\in\mathbb{N}}\subset C(\mathbb{T})$ is called *approximate identity* if

(1) $\int_{\mathbb{T}}k_n(s)\,ds=1$ for all $n\in\mathbb{N}$,
(2) $\sup_{n\in\mathbb{N}}\|k_n\|_{L^1(\mathbb{T})}<\infty$
(3) $\lim_{n\to\infty}\int_\delta^{2\pi-\delta}|k_n(s)|\,ds=0$ for all $\delta\in(0,\pi)$.

If $k_n\ge 0$ we say that the approximate identity is *positive*.

REMARK 3.2. Note that the notion of an approximate identity can analogously be defined for functions on $\mathbb{R}$ instead of $\mathbb{T}$, see [**1**].

Note that the requirement that the k_n are continuous is made for technical reasons here and actually could be dropped. The name "approximate identity" is justified by the following proposition. The need for such notion also emerges from the fact that there does not exists a function $g \in L^1(\mathbb{T})$ such that $f * g = f$ for all $f \in L^1(\mathbb{T})$ — in other words, there is no neutral element with respect to $*$ in L^1, see Exercise 2.6. For the following recall the Riemann integral for Banach space valued functions[6].

LEMMA 3.3. *Let X be a Banach space, $\phi : \mathbb{T} \to X$ continuous and $(k_n)_n$ be an approximate identity. Then*

$$\lim_{n\to\infty} \int_{\mathbb{T}} k_n(s)\phi(s)\, ds = \phi(0).$$

PROOF. By the properties of an approximating identity we have

$$\int_{\mathbb{T}} k_n(s)\phi(s)\, ds - \phi(0) = \int_{\mathbb{T}} k_n(s)[\phi(s) - \phi(0)]\, ds.$$

To show that the right-hand side equals zero, we split the integral

$$\int_{\mathbb{T}} k_n(s)[\phi(s) - \phi(0)]\, ds = \int_{\delta}^{2\pi-\delta} k_n(s)[\phi(s) - \phi(0)]\, ds + \int_{[0,\delta]\cup[2\pi-\delta,2\pi]} k_n(s)[\phi(s) - \phi(0)]\, ds$$

with $\delta \in (0, \pi)$. Since $\phi(2\pi) = \phi(0)$ and since $\|k_n\|_{L^1}$ is uniformly bounded in n, the second term on the right-hand-side can be made arbitrarily small as δ goes to 0 (independent of n). On the other hand,

$$\left\| \int_{\delta}^{2\pi-\delta} k_n(s)[\phi(s) - \phi(0)]\, ds \right\| \le 2 \sup_{s\in\mathbb{T}} \|\phi(s)\|_X \int_{\delta}^{2\pi-\delta} |k_n(s)|\, ds \longrightarrow 0$$

for any δ as $n \to \infty$ by the definition of an approximate identity. □

PROPOSITION 3.4 (Strong continuity). *Let X be one of the following Banach spaces*

$$C^k(\mathbb{T}) \text{ with norm } \|f\|_{C^k(\mathbb{T})} = \sum_{j=0}^{k} \|f^{(j)}\|_\infty, \quad L^p(\mathbb{T}) \text{ with } p \in [1, \infty).$$

Then for $s \in \mathbb{T}$, the operator

$$\tau_s : X \to X, f \mapsto f_s := f(\cdot - s)$$

is linear and isometric. Furthermore,

$$\forall f \in X : \quad s \mapsto \tau_s f \text{ is continuous.}$$

The latter property is called **strong continuity of** τ.

PROOF. We only give a sketch — see the exercises for details. That τ_s is well-defined, linear and isometric is seen directly (also using the translation invariance of the Lebesgue measure). Hence in particular $\|\tau_s\|_{\mathcal{L}(X)} = 1$ for all $s \in \mathbb{T}$. Let $f \in C^k(\mathbb{T})$. Then $f^{(j)}$ is uniformly continuous on $\mathbb{T}$ and hence $\lim_{t\to s} \|f_s^{(j)}(t) - f_t^{(j)}(s)\|_\infty = 0$ for $j \in \{1, .., k\}$. For $X = L^p(\mathbb{T})$, $p \in [1, \infty)$, recall that (the set) $C(\mathbb{T})$ lies dense in X. Since (the Banach space) $C(\mathbb{T})$ is continuously embedded in X, $s \mapsto \tau_s f$ is continuous for $f \in C(\mathbb{T})$ by what we have already shown. The continuity for general $f \in X$ now follows by approximation with continuous functions and the fact that $s \mapsto \|\tau_s\|$ is uniformly bounded (triangle inequality).

□

[6]which, to large extend, can be introduced in the same way as the usual Riemann integral for scalar-valued functions, but 'replacing the absolute value" by the norm of the Banach space".

REMARK 3.5. (1) The assertion of Prop. 3.4 remains true if $\mathbb{T}$ is replaced by $\mathbb{R}$ and $C(\mathbb{T})$ is replaced by

$$C_0(\mathbb{R}) = \{f \in C(\mathbb{R})\colon \lim_{x\to\pm\infty} f(x) = 0\}.$$

This can be proved similarly, using compactly supported continuous functions in the approximation argument.

(2) By Ex. 1.5, the assertion of Prop. 3.4 does not hold for $p = \infty$ as strong continuity fails in this case.

The previous result motivates the following definition.

DEFINITION 3.6. A Banach space X that is continuously embedded in $L^1(\mathbb{T})$, i.e. $\exists C > 0$ with $\|\cdot\|_X \geq C\|\cdot\|_{L^1(\mathbb{T})}$, is called **homogeneous (Banach) space** if

(1) for $f \in X$ it follows that $f_s = f(\cdot - s) \in X$ with $\|f\|_X = \|f_s\|_X$ and
(2) the shift group τ_s is strongly continuous on X.

Note that the spaces considered in Prop. 3.4 are homogeneous Banach spaces

THEOREM 3.7 (Approximation in homogenous spaces). *Let $(k_n)_{n\in\mathbb{N}}$ be an approximate identity and X be a homogeneous Banach space. Then for any $f \in X$ it holds that*

$$f = \lim_{n\to\infty} \int_{\mathbb{T}} k_n(s) f_s \, ds = \lim_{n\to\infty} k_n * f,$$

where the limits as well as the integral exist in X.

PROOF. The first identity follows directly from Lem. 3.3. In fact, $s \mapsto \phi(s) = f_s$ is continuous from $\mathbb{T}$ to X since X is a homogenous Banach space and $\phi(0) = f_0 = f$. For the second identity, we will show that for $k \in C(\mathbb{T})$

$$B_1 f := \int_{\mathbb{T}} k(s) f_s \, ds = k * f =: B_2 f$$

for all $f \in L^1(\mathbb{T})$. If this holds, then the second identity follows as X is continuously embedded in $L^1(\mathbb{T})$ and therefore, the integral $B_1 f$ exists as limit of Riemann sums in both L^1 and X and coincides. It is easy to see that B_1, B_2 are bounded linear operators on $L^1(\mathbb{T})$. Hence, it suffices to show that they coincide on a dense subspace. For $t \in \mathbb{T}$ let $C_t : C(\mathbb{T}) \to \mathbb{C}, f \mapsto f(t)$ denote the point evaluation operator at t. Also note that the integral $B_1 f$ exists as limit of Riemann sums in $C(\mathbb{T})$ — as τ_s is strongly continuous on $C(\mathbb{T})$. Since C_t is linear and bounded, we have that

$$(B_1 f)(t) = C_t B_1 f = \int_{\mathbb{T}} C_t \left(k(s) f_s\right) \, ds = (k * f)(t) = (B_2 f)(t).$$

□

This theorem has several consequence concerning approximation of elements in homogeneous Banach spaces and in particular in the context of trigonometric series.

4. The Fejér kernel and Dirichlet kernel

In the following we would like to apply the abstract result derived in the previous section to specific approximate identities. Recalling that

$$(\mathrm{e}^{ik\cdot} * f)(t) = \hat{f}(k)\mathrm{e}^{int} 2\pi,$$

thus

$$(\sum_{k=-n}^{n} \mathrm{e}^{ik\cdot} * f)(t) = 2\pi \sum_{k=-n}^{n} \hat{f}(k)\mathrm{e}^{ikt} \tag{4.1}$$

one is tempted to believe that

$$D_n(t) = \frac{1}{2\pi}\sum_{k=-n}^{n} \mathrm{e}^{ikt} \tag{4.2}$$

is an approximate identity. If this was true, then Theorem 3.7 would imply that the (partial sums of the) Fourier series

$$(D_n * f)(t) = \sum_{k=-n}^{n} \hat{f}(k)\mathrm{e}^{ikt}$$

of $f \in X$ converge to f in X for any homogenous Banach space X.

EXAMPLE 4.1 (The Dirichlet kernel). The sequence k_n defined by (4.2) is called the **Dirichlet kernel** and is *not* an approximate identity. In fact, $\int_{\mathbb{T}} D_n(s)\, ds = 1$, but $\|D_n\|_{L^1(\mathbb{T})}$ is not uniformly bounded in $n \in \mathbb{N}$. More precisely, using the identity

$$D_n(s) = \frac{1}{2\pi}\frac{\sin((n+\frac{1}{2})t)}{\sin\frac{t}{2}},$$

one can show that $\|D_n\|_{L^1(\mathbb{T})} \sim \log n$, see Exercises.

Not only that $(D_n)_n$ fails to be an approximate identity, it can moreover be shown that the assertion of Theorem 3.7 does not hold with $k_n = D_n$ — at least not in the case $X = L^1(\mathbb{T})$ and $X = C(\mathbb{T})$. Motivated by convergence properties for sequences we may instead hope for some better properties of the arithmetic mean of D_n

Motivated by the fact that the convergence of a sequence $(a_n)_{n\in\mathbb{N}_0}$ may "improve" upon considering the sequence of arithmetic means instead $\left(\frac{a_0+a_1+..+a_n}{n+1}\right)_{n\in\mathbb{N}_0}$, see Ex. 2.0, we introduce the following kernel.

DEFINITION 4.2 (Fejér kernel). The sequence $(F_n)_{n\in\mathbb{N}_0} \subset C(\mathbb{T})$ defined by

$$F_n(t) = \frac{1}{n+1}\sum_{k=0}^{n} D_k(t), \qquad t \in \mathbb{T},$$

where $(D_k)_k$ is defined in Ex. 4.1, is called the **Fejér kernel**. For $f \in L^1(\mathbb{T})$, we call

$$F_n * f = \frac{1}{n+1}\sum_{k=0}^{n}(D_k * f)$$

the n-th **Fejér mean** (or Césaro mean) of f. By (4.1), the n-th Fejér mean is nothing else than the arithmetic mean of the first $n+1$ partial sums of the Fourier series of f.

PROPOSITION 4.3. *The Fejér kernel $(F_n)_{n\in\mathbb{N}_0}$ defined in Def. 4.2 is an approximate identity, additionally satisfying for all $n \in \mathbb{N}_0$, $t \in \mathbb{T}$ and $\delta \in (0,\pi)$,*

$$F_n(t) \geq 0, \qquad F_n(t) = F_n(-t), \qquad \lim_{n\to\infty} \|F_n\|_{L^\infty(\delta,2\pi-\delta)} = 0.$$

Furthermore, the following identities hold.

$$F_n(t) = \frac{1}{2\pi}\sum_{k=0}^{n}\left(1 - \frac{k}{n+1}\right)\mathrm{e}^{int} = \frac{1}{2\pi(n+1)}\left(\frac{\sin((n+1)\frac{t}{2})}{\sin\frac{t}{2}}\right)^2. \tag{4.3}$$

PROOF. See Ex. 2.1 (first show (4.3) from which the rest follows rather directly). □

PROPOSITION 4.4. *ATTENTION: there may have been a typo on the blackboard If $(D_n * f)_{n\in\mathbb{N}}$ converges in X where X is a homogeneous Banach space and $f \in X$, then $(F_n * f)_{n\in\mathbb{N}}$ converges in X as well — and to the same limit.*

PROOF. This follows since $F_n * f$ is the arithmetic mean of $D_n * f$ and Ex. 2.0. $\square$

THEOREM 4.5 (Approximation by Fejér kernel). *Let X be a homogeneous Banach space and $f \in X$. Then*

$$(F_n * f) \text{ converges to } f \text{ in } X \text{ as } n \to \infty.$$

PROOF. This follows from Theorem 3.7 and since $(F_n)_n$ is an approximate identity, Proposition 4.3. $\square$

COROLLARY 4.6. *The set of trigonometric polynomials* $\mathrm{Trig}(\mathbb{T})$ *lies dense in any homogeneous Banach space X in the sense that the closure of* $\mathrm{Trig}(\mathbb{T}) \cap X$ *in X equals X. Moreover, for any $f \in X$, the sequence*

$$f_n(t) = \sum_{k=0}^{n} \left(1 - \frac{k}{n+1}\right) \hat{f}(k) \mathrm{e}^{int}$$

converges to f in the X norm. In particular, the trigonometric polynomials lie dense in $C^k(\mathbb{T})$ and $L^p(\mathbb{T})$ for any $p \in [1, \infty)$, $k \in \mathbb{N}_0$.

PROOF. Note that $F_n * f = f_n$ by Proposition 4.3 and Theorem 2.6. The rest follows from Theorem 4.5. $\square$

Corollary 4.6 provides a constructive proof of Weierstrass' theorem. However, we should be aware of the fact that the proof of Theorem 3.7 relied on the fact that $C(\mathbb{T})$ lies dense in $L^1(\mathbb{T})$ — which then should not be argued by Weierstrass' theorem, but more directly (e.g. use that the simple functions are dense in L^1 and apply Lusin's theorem).

THEOREM 4.7. *For $X = L^1(\mathbb{T})$ or $X = C(\mathbb{T})$ there exists a function $f \in X$ such that the partial sums $D_n * f$ of the Fourier series of f do <u>not</u> converge in X. Moreover, there exists $f \in C(\mathbb{T})$ and $t_0 \in \mathbb{T}$ such that $(D_n * f)(t_0)$ does not converge.*

PROOF. This is proved in Ex. 2.5 by an abstract argument employing the uniform boundedness principle. $\square$

REMARK 4.8. Although the proof of Theorem 4.7 was based on an abstract existence argument, it is possible to construct such $f \in C(\mathbb{T})$ explicitly, see e.g. [**2, 1**].

Theorems 4.5 and 4.7 leave us with two main questions/goals.

(I) Do the Fejér means $(F_n * f)_n$ converge pointwise almost everywhere for any $f \in L^1(\mathbb{T})$?
(II) Do the partial sums $(D_n * f)_n$ of a Fourier series converge in $L^p(\mathbb{T})$ for any $f \in L^p(\mathbb{T})$?

We remark that we could have asked the second question for more general homogeneous Banach spaces X other than $C(\mathbb{T})$ or $L^1(\mathbb{T})$, but here will restrict to the important class of L^p spaces only.

Let us first approach (I):

THEOREM 4.9 (Pointwise convergence of Fejér means). *Let $f \in L^1(\mathbb{T})$, $t_0 \in \mathbb{T}$ and assume that there exists $L_{t_0} \in \mathbb{C}$ such that*

$$\lim_{h \to 0^+} \frac{1}{h} \int_0^h \left| \frac{1}{2}(f(t_0 + s) + f(t_0 - s)) - L_{t_0} \right| ds = 0. \tag{4.4}$$

Then $\lim_{n\to\infty}(F_n * f)(t_0) = L_{t_0}$.

PROOF. Let us consider $(F_n * f)(t_0) - L_{t_0}$ for $n \in \mathbb{N}$: Since $\int_{\mathbb{T}} F_n = 1$, we have

$$\begin{aligned}(F_n * f)(t_0) - L_{t_0} &= \int_{\mathbb{T}} F_n(s)(f(t_0 - s) - L_{t_0})\,ds \\ &= \int_{-\pi}^{\pi} F_n(s)(f(t_0 - s) - L_{t_0})\,ds \\ &= \int_0^{\pi} F_n(s)(f(t_0 + s) + f(t_0 - s) - 2L_{t_0})\,ds,\end{aligned}$$

where we used $F_n(-s) = F_n(s)$, see Prop. 4.3, in the last identity. We now split the integral, $\int_0^\pi = \int_0^v + \int_v^\pi$ with $v \in (0, \pi)$ which will be determined later, and estimate by triangle inequality,

$$\left| \int_0^{\pi} F_n(s)(f(t_0 + s) + f(t_0 - s) - 2L_{t_0})\,ds \right| \le \int_0^v |.| + \int_v^\pi |.| =: I_1 + I_2.$$

The goal is to show that $I_1 + I_2 \to 0$ as $n \to \infty$. To estimate I_1 and I_2, let us investigate $|F_n(s)|$, $s \in (0, \pi)$, first, by using Prop. 4.3 again,

$$|F_n(s)| = \frac{1}{2\pi(n+1)} \left| \frac{\sin((n+1)\frac{s}{2})}{\sin\frac{s}{2}} \right|^2 \le \min\left(n+1, \frac{\pi^2}{(n+1)s^2}\right) \tag{4.5}$$

where use the elementary estimates $|\sin(x)| \le |x|$ and $\frac{2}{\pi}x \le \sin(x)$ for $x \in [0, \frac{\pi}{2})$ as follows

$$\left| \frac{\sin((n+1)\frac{s}{2})}{\sin\frac{s}{2}} \right|^2 \le \frac{1}{|\sin\frac{s}{2}|^2} \le \frac{\pi^2}{4}\frac{4}{s^2}, \quad \left| \frac{\sin((n+1)\frac{s}{2})}{\sin\frac{s}{2}} \right|^2 \le \frac{(n+1)^2(\frac{s}{2})^2}{(\frac{2s}{\pi 2})^2} \le (n+1)^2\frac{\pi^2}{4}.$$

The idea is to choose $v = v(n)$ such that I_1 and I_2 converge to 0 as $n \to \infty$. Let $v = n^{-\frac{1}{4}}$. Then we have for I_2,

$$I_2 = \int_{n^{-1/4}}^{\pi} |F_n(s)| \cdot \underbrace{|f(t_0 + s) + f(t_0 - s) - 2L_{t_0}|}_{=:g(s),\ g \in L^1(\mathbb{T})}\,ds \overset{(4.5)}{\le} \frac{\pi^2}{(n+1)n^{-1/2}} \int_0^\pi |g(s)|\,ds \overset{n\to\infty}{\longrightarrow} 0.$$

We still have not used the assumption that the limit in (4.4) is finite. To do so, and to handle I_2 define

$$P(h) = \int_0^h |g(s)|\,ds \quad \text{with } g(s) = f(t_0 + s) + f(t_0 - s) - 2L_{t_0}.$$

By assumption, $\frac{1}{h}P(h) \to 0$ as $h \to 0^+$. By further splitting the integral in I_1 we get We further split $I_1 = \int_0^{n^{-1}} |.| + \int_{n^{-1}}^{n^{-1/4}} |.|$ (having set $v = n^{-1/4}$) and get

$$\begin{aligned}I_1 = \int_0^{n^{-1}} |.| + \int_{n^{-1}}^{n^{-1/4}} |.| &\overset{(4.5)}{\le} \int_0^{n^{-1}} (n+1)|g(s)|\,ds + \int_{n^{-1}}^{n^{-1/4}} \frac{\pi^2}{(n+1)s^2}|g(s)|\,ds \\ &= (n+1)P(\tfrac{1}{n}) + \int_{n^{-1}}^{n^{-1/4}} \frac{\pi^2}{(n+1)s^2}P'(s)\,ds,\end{aligned}$$

where in the last step we used Lebesgue's fundamental theorem of calculus[7]. By assumption, $(n+1)P(\frac{1}{n}) \to 0$ as $n \to \infty$ and it remains to consider the second term on the right-hand side. Integration by parts gives

$$\int_{n^{-1}}^{n^{-1/4}} \frac{\pi^2}{(n+1)s^2}P'(s)\,ds = \frac{\pi^2}{n+1}\frac{P(s)}{s^2}\Big|_{n^{-1}}^{n^{-1/4}} + \frac{2\pi^2}{n+1}\int_{n^{-1}}^{n^{-1/4}} \frac{P(s)}{s^3}\,ds$$

[7] implying that P, as the primitive of an integrable function, is differentiable almost everywhere and the derivative P' coincides with the integrand.

It follows directly from $\lim_{h\to 0^+} h^{-1}P(h) = 0$ that the first term on the right-hand side tends to 0 as $n \to \infty$. For the second term note that by the same reasoning, for every $\epsilon > 0$ it holds that $\frac{P(s)}{s^3} \leq \frac{\epsilon}{s^2}$ for all $s \in (0, s_\epsilon)$ and sufficiently small $s_\epsilon > 0$. Therefore, for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n > N$,

$$\frac{2\pi^2}{n+1}\int_{n^{-1}}^{n^{-1/4}} \frac{P(s)}{s^3}\, ds \leq \frac{2\epsilon\pi^2}{n+1}(n - n^{\frac{1}{4}}) < 2\epsilon\pi^2.$$

Thus, altogether, $I_1 \to 0$ as $n \to \infty$. $\square$

REMARK 4.10. Note that the condition (4.4) is in particular satisfied if

$$\lim_{s\to 0^+} \frac{1}{2}(f(t_0+s) + f(t_0-s)) = L_{t_0} \tag{4.6}$$

and hence especially for any point t_0 if f is in $C(\mathbb{T})$. In this special situation, the version of Theorem 4.9 is referred to as *Fejér's theorem.*

Theorem 4.9 even implies that the Fejér means $F_n * f$ of a function $f \in L^1(\mathbb{T})$ converge pointwise almost everywhere to f. The proof relies on a stronger version of Lebesgue fundamental theorem of calculus saying particularly that for $f \in L^1(\mathbb{T})$, condition (4.4) is satisfied for almost every t_0 in $\mathbb{T}$. In fact, one can use (more involved) methods from Fourier analysis to show this result — we leave out the proof and refer to e.g. [**3**, Lecture 3, Thm. 2.1] and see also [**1**, Thm. 3.4.4].

THEOREM 4.11. *Let $f \in L^1(\mathbb{T})$, then $\lim_{n\to\infty}(F_n * f)(t) = f(t)$ for almost every $t \in \mathbb{T}$.*

Theorem 4.11 provides an affirmative answer to Question (I). For a similar result concerning pointwise convergence for another summability kernel, see Exercise 4.4 on the Poisson kernel.

5. Convergence of Fourier series in homogeneous Banach spaces

We have already seen that for the spaces $X = C(\mathbb{T})$ and $X = L^1(\mathbb{T})$ the partial sums of a Fourier series (i.e. the sequence of Dirichlet means) does not converge in the respective norms, i.e.

$$D_n * f \to f \quad \text{in } X \text{ for } n \to \infty.$$

This failure was seen by relating convergence with uniform boundedness of the operator sequence $(\mathcal{M}_{D_n})_{n\in\mathbb{N}}$ defined by

$$\mathcal{M}_{D_n} : X \to X, f \mapsto D_n * f. \tag{5.1}$$

Let us begin this section with showing that this is actually a manifestation of an argument for general spaces. In the following the phrase "the Fourier series of f converges to f in X always refers to the convergence of the partial sums $D_n * f$.

THEOREM 5.1. *Let X be a homogeneous Banach space. Then the Fourier series of all $f \in X$ converges to f in X if and only if the operators $(\mathcal{M}_{D_n})_{n\in\mathbb{N}}$ defined by* (5.1) *are uniformly bounded.*

PROOF. The sufficiency ("⇒") is a direct consequence of the uniform boundedness principle. In fact, for $f \in X$ the sequence $(\mathcal{M}_{D_n} f)_{n\in\mathbb{N}}$ is bounded in X, i.e. $\sup_{n\in\mathbb{N}} \|\mathcal{M}_{D_n} f\|_X < \infty$ since $\mathcal{M}_{D_n} f$ converges for $n \to \infty$ by assumption. Thus, by the uniform boundedness principle (recalling that $\mathcal{M}_{D_n} : X \to X$ is a bounded operator on X for every $n \in \mathbb{N}$, see Ex. 2.4), we conclude that $(\mathcal{M}_{D_n})_{n\in\mathbb{N}}$ is bounded in $B(X,X)$, the space of bounded linear operators on X. Conversely, assume that $C := \sup_{n\in\mathbb{N}} \|\mathcal{M}_{D_n}\| < \infty$. In order to show that $\mathcal{M}_{D_n} f \to f$ as $n \to \infty$, we first consider the case when f is a trigonometric polynomial and show the general case by density of these polynomials in X and that $C < \infty$.

$\square$

As we have seen in Exercise 2.4, the norm of $S_n = \mathcal{M}_{D_n}$ as operator on X can be bounded by $\|D_n\|_{L^1(\mathbb{T})}$, but this inequality may be strict.

Remark 5.2. As described in the beginning of this section, from Example 4.1 (and see also Ex. 2.5) it can be inferred that $\|\mathcal{M}_{D_n}\|_{L^1(\mathbb{T})\to L^1(\mathbb{T})} \sim \log(n)$ which, by Theorem 5.1, implies that there exists $f \in L^1(\mathbb{T})$ such that the Fourier series of f does not converge to f. On the other hand we can "apply" Theorem 5.1 to the Dirichlet kernel in $X = L^2$. From Exercise 2.4 (d), we know that $\|S_n\|_{L^2(\mathbb{T})\to L^2(\mathbb{T})} \leq 2\pi\|\widehat{D_n}\|_{\ell^\infty(\mathbb{Z})} = 1$ and hence the partial sums of the Fourier series for L^2 functions converge in L^2. However, recall that this is something we knew already from the functional analysis class, as indicated in the introduction.

Remark 5.3. In contrast to the situation in Remark 5.2 (where $X = C(\mathbb{T})$ or $X = L^1(\mathbb{T})$), it is difficult to calculate (or estimate) $\|S_n\|_{X\to X}$ for general homogeneous spaces X and thus to conclude on the convergence of the Fourier series. For that reason, we will reformulate this property once more in the next subsection. Our ultimate goal is to discuss the convergence of $D_n * f$ in L^p (for $f \in L^p$).

5.1. Conjugate trigonometric series.

Definition 5.4 (Conjugate trigonometric series). Given a trigonometric series $\sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{int}$, we call

$$-\sum_{n\in\mathbb{Z}} i\,\mathrm{sgn}(n) a_n \mathrm{e}^{int}$$

the **conjugate (trigonometric) series** [8].

In the following we shall be particularly interested in conjugate series of Fourier series. Let us first study the question whether Fourier series are invariant under conjugation, i.e.

Is the conjugate series of a Fourier series again Fourier series?

Recall that we say that a trigonometric series is a Fourier series if its coefficients are the Fourier coefficients of some function in $L^1(\mathbb{T})$. This question looks harmless, but, as we shall see shortly, has a negative answer. However, more special functions $f \in L^1(\mathbb{T})$ for which an affirmative answer can be given are e.g. $f \in \mathrm{A}(\mathbb{T}) = \{f \in L^1(\mathbb{T})\colon \hat{f} \in \ell^1(\mathbb{Z})\}$ (*think about why!*).

Proposition 5.5.
(1) *Let $(a_n)_{n\in\mathbb{Z}} \in c_0(\mathbb{Z})$ be a given sequence such that $a_n = a_{-n}$, $a_n \geq 0$ and $\frac{1}{2}(a_{n-1} + a_{n+1}) \geq a_n$ for all $n \in \mathbb{N}_0$. Then there exists $f \in L^1(\mathbb{T})$ nonnegative such that $\hat{f}(n) = a_n$ for all $n \in \mathbb{Z}$.*
(2) *Let $f \in L^1(\mathbb{T})$ be such that $\hat{f}(n) = -\hat{f}(-n) \geq 0$ for all $n \in \mathbb{N}_0$. Then $\left(\frac{\hat{f}(n)}{n}\right)_{n\in\mathbb{N}} \in \ell^1(\mathbb{N})$.*

Proof. (1): From the assumption, $(a_n - a_{n+1}) \searrow 0$ and moreover, $n(a_n - a_{n+1}) \to 0$ as $n \to \infty$ (this follows from $\sum_{n\in\mathbb{N}_0}(a_n - a_{n+1}) = a_0$). Define

$$f(t) = 2\pi \sum_{n=1}^{\infty} n(a_{n-1} + a_{n+1} - 2a_n) F_{n-1}(t)$$

for $t \in \mathbb{T}$. To see that the series converges in $L^1(\mathbb{T})$, note that by $\sup_{n\in\mathbb{N}} \|F_n\|_{L^1(\mathbb{T})} < \infty$, it suffices to see that $\sum_{n=1}^{k} n(a_{n-1} + a_{n+1} - 2a_n)$ converges for $k \to \infty$. The latter follows easily from the identity

$$\sum_{n=1}^{k} n(a_{n-1} + a_{n+1} - 2a_n) = a_0 - a_k - k(a_k - a_{k+1}).$$

[8] Recall that without further specifications these series are only formally defined by their coefficients sequence.

This also shows that the limit as $k \to \infty$ equals a_0. It remains to show that $\hat{f}(m) = a_m$ for all $m \in \mathbb{Z}$. By linearity and boundedness of $f \mapsto \hat{f}$ we have

$$\hat{f}(m) = \sum_{n=1}^{\infty} n(a_{n-1} + a_{n+1} - 2a_n) 2\pi \widehat{F_{n-1}}(m).$$

Since $\widehat{F_{n-1}}(m) = 0$ for $|m| \geq n$ and $2\pi\widehat{F_{n-1}}(m) = (1 - \frac{|m|}{n})$ for $|m| < n$, this yields

$$\hat{f}(m) = \sum_{n=|m|+1}^{\infty} n(a_{n-1} + a_{n+1} - 2a_n)(1 - \frac{|m|}{n}).$$

Using the above derived convergence properties of (a_n) it is not hard to see that the last sum equals $a_{|m|}$ (index shift!).

(2): see Ex. 3.1(3). □

Corollary 5.6. *There exists a Fourier series such that its conjugate series is not a Fourier series. More precisely, the following holds: Let $a_n > 0$ for all $n \in \mathbb{Z}$ such that $\sum_{n\in\mathbb{Z}} \frac{a_n}{n} = \infty$. Then $-i \sum_{n\in\mathbb{Z}} a_n \mathrm{sgn}(n) \mathrm{e}^{int}$ is not a Fourier series.*

Proof. Let $a_n = \frac{1}{\log(n)}$ for $|n| \geq 2$ and 0 else. □

Definition 5.8 (Homogeneous spaces invariant under conjugation). We say that a homogeneous Banach space X is closed under conjugation if for any $f \in X$, it follows that there exists $\tilde{f} \in X$ with

$$\hat{\tilde{f}}(m) = -i\mathrm{sgn}(m)\hat{f}(m) \qquad \forall m \in \mathbb{Z}.$$

Note that the property of being invariant under conjugation is non-trivial: In particular, the homogeneous space $L^1(\mathbb{T})$ does not possess it by Corollary 5.6. On the other hand, examples of a space which are invariant under conjugation are given by $X = L^2(\mathbb{T})$ or $X = \mathrm{A}(\mathbb{T})$ as can be seen easily.

Proposition 5.9. *Let X be a homogeneous Banach space. The following assertions are equivalent.*

(1) *X is invariant under conjugation,*
(2) *$f \mapsto \tilde{f}$ is a well-defined, bounded linear operator from X to X.*
(3) *$f \mapsto P_+ f = \frac{1}{2}\hat{f}(0) + \frac{1}{2}(f + i\tilde{f})$ is a well-defined, bounded linear operator from X to X.*

Note that $\widehat{(P_+ f)}(m) = \hat{f}(m)$ for $m \geq 0$, and $\widehat{(P_+ f)}(m) = 0$ for $m < 0$.

Proof. We show (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\Longrightarrow$ (1). The first implication is a consequence of the closed graph theorem, by which it suffices to show that the mapping $f \mapsto \tilde{f}$ is closed. By (1), the mapping is well-defined and obviously linear. Let $f_n, f, g \in X$, $n \in \mathbb{N}$, be such that $f_n \to f$ and $\tilde{f}_n \to g$ in X as $n \to \infty$. To show that $\tilde{f} = g$, consider the Fourier coefficients. Since $\|\cdot\|_X \geq \|\cdot\|_{L^1}$, we have that

$$\lim_{n\to\infty} \widehat{f_n}(m) = \hat{f}(m) \quad \text{and} \quad \lim_{n\to\infty} \widehat{\tilde{f}_n}(m) = \hat{g}(m)$$

for all $m \in \mathbb{Z}$, by Theorem 1.6. On the other hand, we have that $\widehat{\tilde{f}_n}(m) = -i\mathrm{sgn}(m)\widehat{f_n}(m)$ by definition. Therefore, $\hat{\tilde{f}} = \hat{g}$ and thus $\tilde{f} = g$ by uniqueness of the Fourier coefficients, Theorem 1.6. Thus, by the closed graph theorem (X is a Banach space), the mapping $f \mapsto \tilde{f}$ is bounded. (2) $\Longrightarrow$ (3): This is clear since $f \mapsto \hat{f}(0)$ and $f \mapsto f$ are bounded operators on X and therefore P_+ is bounded as sum of bounded operators.

(3) $\Longrightarrow$ (1): This follows by $\tilde{f} = -i(2P_+ f - \hat{f}(0) - f)$ and since the right-hand side is well-defined for $f \in X$. □

In fact, homogeneous spaces which are invariant under conjugation characterize the property that the partial sums of Fourier series converge:

THEOREM 5.10. *Let X be a homogeneous Banach space such that $f \mapsto \mathrm{e}^{in\cdot} f$ is an isometry from X to X* [9] *for any $n \in \mathbb{Z}$. Then the following assertions are equivalent.*

(1) *For all $f \in X$ the Fourier series converges to f in X,*
(2) *X is invariant under conjugation.*

PROOF. By Theorem 5.1 and Proposition 5.9 it suffices to show that $\mathcal{M}_{D_n}$ are uniformly bounded operators on X if and only if P_+ is a bounded operator on X. From the definition of the Dirichlet kernel and Theorem 2.6 we get the following dentity for $f \in X$,

$$\sum_{k=0}^{2n} \hat{f}(k)\mathrm{e}^{ikt} = \mathrm{e}^{int}(D_n * (f\mathrm{e}^{-in\cdot}))(t) = \mathrm{e}^{int}(\mathcal{M}_{D_n}(f\mathrm{e}^{-in\cdot})) \tag{5.2}$$

For $n \in \mathbb{Z}$ let J_n be the operator defined by $(J_n f)(t) = \mathrm{e}^{int} f(t)$. Clearly, $J_{-n} = (J_n)^{-1}$ and, by assumption, $\|J_n f\|_X = \|f\|_X$ and hence particularly $\|J_n\| = 1$. Let f be a trigonometric polynomial. Then for sufficiently large n,

$$P_+ f = \sum_{k=0}^{2n} \hat{f}(k)\mathrm{e}^{ikt} = J_n \mathcal{M}_{D_n} J_{-n} f,$$

where we used (5.2). Thus $\|P_+ f\|_X \le \|\mathcal{M}_{D_n}\| \|f\|_X$. If we assume that $C = \sup_{n\in\mathbb{N}} \|\mathcal{M}_{D_n}\|_{X\to X} < \infty$, then it follows that $\|P_+ f\|_X \le C\|f\|_X$ for any trigonometric polynomial and thus P_+ is bounded on X by density, Corollary 4.6. Conversely, if P_+ is bounded on X, then by

$$(J_{-n}P_+J_n - J_{n+1}P_+J_{-n-1})f = \mathcal{M}_{D_n} f$$

for any $f \in X$, we have that $\|\mathcal{M}_{D_n}\| \le 2\|P_+\|$ for all $n \in \mathbb{N}$. □

THEOREM 5.11 (**Convergence of Fourier series in L^p, Riesz 1927**). *Let $p \in (1, \infty)$. Then for every $f \in X = L^p(\mathbb{T})$ the partial sums $D_n * f$ of the Fourier series of f converge to f in X as $n \to \infty$.*

SKETCH OF THE PROOF. The proof of Theorem 5.11 is split into the following steps

(1) Show Theorem 5.11 for the special case that $p = 2k$ with $k \in \mathbb{N}$, (*Lemma A*)
(2) Show that this implies the assertion for $p \ge 2$ (*interpolation*)
(3) Show the case $p \in (1, 2)$ (*duality*).

□

LEMMA A. The assertion of Theorem 5.11 holds for $p = 2k$ with $k \in \mathbb{N}$.

PROOF. As we have seen in Proposition 5.9, boundedness of the conjugation $f \mapsto \tilde{f}$ on $X = L^p$ characterizes the convergence of the Fourier series. Hence, it suffices to show that there exists some constant $C_{2k} > 0$ such that $\|\tilde{f}\|_{L^{2k}} \le C_{2k}\|f\|_{L^{2k}}$ for all trigonometric polynomials f, since the latter set is dense in X, Corollary 4.6. For a trigonometric polynomial f, $P_+ f = \frac{1}{2}(\hat{f}(0) + f + i\tilde{f}) = \sum_{j=0}^{N} \hat{f}(j)\mathrm{e}^{ik\cdot}$, for sufficiently large N. If we additionally assume that $\hat{f}(0) = 0$, then the latter identity implies using the elementary identity (1.1) that

$$\int_{\mathbb{T}} (f(t) + i\tilde{f}(t))^{2k}\, dt = 0.$$

[9] i.e. $t \mapsto \mathrm{e}^{int} f(t) \in X$ and $\|\mathrm{e}^{in\cdot} f\|_X = \|f\|_X$ for all $f \in X$.

From this identity we will now extract an estimate for $\|\tilde{f}\|_{L^{2k}}$. By expanding the binomial, we derive

$$\sum_{j=0}^{2k} \binom{2k}{j} i^{2k-j} \int_{\mathbb{T}} f(t)^j \tilde{f}(t)^{2k-j}\, dt = 0.$$

Let us for the moment further assume that f and $\tilde{f}$ are real-valued. Thus, taking real parts gives

$$\sum_{j=1}^{k} \binom{2k}{2j} (-1)^{k-j} \int_{\mathbb{T}} f(t)^{2j} \tilde{f}(t)^{2k-2j}\, dt = -(-1)^k \|\tilde{f}\|_{2k}^{2k}$$

and thus by Hölder's inequality

$$\begin{aligned} \|\tilde{f}\|_{L^{2k}}^{2k} &\leq \sum_{j=1}^{k} \binom{2k}{2j} \int_{\mathbb{T}} |f(t)^{2j} \tilde{f}(t)^{2k-2j}|\, dt \\ &\leq \sum_{j=1}^{k} \binom{2k}{2j} \|f\|_{L^{2k}}^{2j} \|\tilde{f}\|_{L^{2k}}^{2k-2j}. \end{aligned}$$

Dividing by $\|f\|_{L^{2k}}^{2k}$ yields that

$$\left(\frac{\|\tilde{f}\|_{L^{2k}}}{\|f\|_{L^{2k}}} \right)^{2k} \leq \sum_{j=1}^{k} \binom{2k}{2j} \left(\frac{\|\tilde{f}\|_{L^{2k}}}{\|f\|_{L^{2k}}} \right)^{2k-2j}$$

for all $f \neq 0$ satisfying the listed assumptions. In other words, $x = \frac{\|\tilde{f}\|_{L^{2k}}}{\|f\|_{L^{2k}}}$ satisfies $p(x) \leq 0$ for the polynomial $p(x) = x^{2k} - \sum_{j=1}^{k} \binom{2k}{2j} x^{2k-2j}$. By the mean value theorem and the fact that $p(0) < 0$, we conclude that p has at least one zero and we can hence denote by C_{2k} the largest zero of p. It follows that

$$\|\tilde{f}\|_{L^{2k}} \leq C_{2k} \|f\|_{L^{2k}}$$

for all real-valued trigonometric polynomials f such that $\tilde{f}$ is also real-valued and $\hat{f}(0) = 0$.
It remains to remove the made conditions on the trigonometric polynomials f. If f is such that $\hat{f}(0) = 0$, then we can decompose f by $f = P + iQ$ where

$$P(t) = \sum_{j=-N}^{N} \underbrace{\frac{1}{2} \left(\hat{f}(j) + \overline{\hat{f}(-j)} \right)}_{c_j} \mathrm{e}^{ijt}, \qquad Q(t) = \sum_{j=-N}^{N} \frac{1}{2i} \left(\hat{f}(j) - \overline{\hat{f}(-j)} \right) \mathrm{e}^{ijt},$$

and where N is sufficiently large. Then $P(t) = \sum_{j=1}^{N} 2\Re c_j \cos(jt)$ and $Q(t) = \sum_{j=1}^{N} 2\Im c_j \sin(jt)$ are real-valued. Therefore, by linearity of $f \mapsto \tilde{f}$ and by what we have already shown,

$$\|\tilde{f}\|_{L^{2k}} = \|\tilde{P} + i\tilde{Q}\|_{L^{2k}} \leq C_{2k}(\|P\|_{L^{2k}} + \|Q\|_{L^{2k}}) \leq 2C_{2k}\|f\|_{L^{2k}}. \tag{5.3}$$

where the last inequality follows by $\max\{|P(t)|, |Q(t)|\} \leq (|P(t)| + |Q(t)|)^{\frac{1}{2}} = |f(t)|$ for $t \in \mathbb{T}$. Finally, we remove the assumption that $\hat{f}(0) = 0$: For a general trigonometric polynomial we have by linearity and the fact that $\tilde{g} = 0$ if g is constant,

$$\|\tilde{f}\|_{L^{2k}} = \|\widetilde{(f - \hat{f}(0))} + \widetilde{\hat{f}(0)}\|_{L^{2k}} = \|\widetilde{(f - \hat{f}(0))}\|_{L^{2k}} \leq 2C_{2k}\|f - \hat{f}(0)\|_{L^{2k}}$$

where we used (5.3). By Theorem 1.6 and Hölder's inequality,

$$\begin{aligned}\|f - \hat{f}(0)\|_{L^{2k}} &\le \|f\|_{L^{2k}} + |\hat{f}(0)|(2\pi)^{\frac{1}{2k}} \\ &\le \|f\|_{L^{2k}} + \|f\|_{L^1}(2\pi)^{\frac{1}{2k}-1} \\ &\le \|f\|_{L^{2k}} + \|f\|_{L^{2k}}(2\pi)^{\frac{1}{k}-1} \le 2\|f\|_{L^{2k}}.\end{aligned}$$

Thus, altogether, $\|f\|_{L^{2k}} \le 4C_{2k}\|f\|_{L^{2k}}$ for trigonometric polynomials f and hence for all $f \in L^{2k}$. □

Recall that a function $f : \mathbb{C} \to \mathbb{C}$ is called entire if there exists a sequence $(a_n)_{n\in\mathbb{N}_0}$ of complex numbers such that $f(z) = \sum_{n=0}^{\infty} a_n z^n$ for all $z \in \mathbb{C}$. For entire functions one has the following *maximum modulus principle*:

$$\Omega \subset \mathbb{C} \text{ open, bounded } \implies \sup_{z\in\Omega} |f(z)| = \max_{z\in\partial\Omega} |f(z)|, \tag{5.4}$$

where $\partial\Omega$ denotes the boundary of Ω, see Exercise 5.2 for a proof. We even need version of the maximum modulus principe for unbounded domains, which is covered by the following result.

LEMMA 5.12 (**Hadamard's Three Lines lemma** — for entire functions). *Let H be entire and assume that H grows at most exponentially on the vertical strip $S = \{z \in \mathbb{C} \colon \Re z \in (0,1)\}$, i.e. there exists $c_1, c_2 > 0$ such that $|H(z)| \le c_1 e^{c_2 \Im(z)}$ for all $z \in S$. Further suppose that $C_j := \sup_{\Re z = j} |H(z)| < \infty$ for $j \in \{0,1\}$. Then*[10]

$$|H(z)| \le C_0^{1-\theta} C_1^{\theta} \qquad \textit{if } \Re z = \theta.$$

PROOF. Define G by $G(z) = H(z)(C_0^{1-z}C_1^{z})^{-1}$. Since $|C_0^{1-z}C_1^{z}| = C_0^{1-\theta}C_1^{\theta}$ for $\Re z = \theta$, it thus remains to show that $|G(z)| \le 1$ for all $z \in S$. As we cannot apply the maximum modulus principle to the unbounded set S directly, we first assume that $|G(x+iy)| \to 0$ as $|y| \to \infty$ uniformly in $x \in (0,1)$. Hence there exists $y_0 > 0$ such that $|G(z)| < 1$ for $z \in S$ with $\Im(z) > y_0$. Applying the maximum modulus principle, (5.4), to G and $\Omega = \{z \in S \colon |\Im z| < y_0\}$, yields that $|G(z)| \le 1$ for $z \in \Omega$ since it follows from the assumption on H that $|G(z)| \le 1$ when either $\Re z = 0$ or $\Re z = 1$. Together this shows that $|G(z)| \le 1$ for all $z \in S$. In the general case when H is only assumed to grow at most exponentially, we consider G_n defined by

$$G_n(z) = G(z) e^{\frac{z^2-1}{n}}$$

which converges to $G(z)$ for any $z \in S$ as $n \to \infty$. Since $|e^{\frac{z^2-1}{n}}| = e^{\frac{1}{n}(-(\Im z)^2 + (\Re z)^2 - 1)}$, it follows that for every fixed n, $|G_n(z)|$ goes to 0 uniformly in S as $\Im(z) \to \infty$ and thus satisfies the assumption made in the first proof step. Therefore, $|G_n(z)| \le 1$ for all $n \in \mathbb{N}$ and thus $|G(z)| \le 1$ for all $z \in S$. □

Recall that for a measure space (Ω, μ) a function $f : \Omega \to \mathbb{C}$ is called simple if it is of the form

$$f = \sum_{j=1}^{n} f_j \chi_{A_j}$$

with $n \in \mathbb{N}$, $f_j \in \mathbb{C}$ and measurable, disjoint sets $A_j \subset \Omega$. We say that a simple function f has finite measure support if $\mu(A_j) < \infty$ for $j \in \{1, \dots, n\}$. Note that for $p \in [1, \infty)$, the set of simple functions with finite measure support $M_{simple,finite}(\Omega, \mu)$ lies dense in any $L^p(\Omega, \mu)$ if the measure space is σ-finite.

[10] Note that this statement actually holds more generally for functions f which are analytic on S and continuous on the closure $\overline{S}$.

THEOREM 5.13 (**Riesz–Thorin interpolation theorem**). *Let* (Ω_i, μ_i), $i = 1, 2$ *be* σ*-finite measure spaces and let* $p_i, q_i \in [1, \infty]$. *Let* T *be an operator mapping the space of simple functions* $\Omega_1 \to \mathbb{C}$ *with finite measure support* $M_{simple,finite}(\Omega, \mu)$ *to* $L^{q_1}(\Omega_2, \mu_2) \cap L^{q_2}(\Omega_2, \mu_2)$ *such that*

$$\|Tf\|_{L^{q_i}(\Omega_2)} \leq C_i \|f\|_{L^{p_i}(\Omega_1)} \qquad \forall f \in M_{simple,finite}(\Omega_1, \mu). \tag{5.5}$$

Then there exists a constant $C_{p,q} > 0$ *such that*

$$\|Tf\|_{L^q(\Omega_2)} \leq C_{p,q} \|f\|_{L^p(\Omega_1)} \qquad \forall f \in M_{simple,finite}(\Omega_1, \mu),$$

where

$$\frac{1}{p} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}, \quad \frac{1}{q} = \frac{1-\theta}{q_1} + \frac{\theta}{q_2}, \theta \in (0,1), \tag{5.6}$$

and $C_{p,q} \leq C_1^{1-\theta} C_2^{\theta}$.
In particular, if $p < \infty$, *then* T *extends to a bounded operatore from* $L^p(\Omega_1, \mu_1)$ *to* $L^q(\Omega_2, \mu_2)$.

PROOF. In the following we will abbreviate $L^p = L^p(\Omega_1)$ and $L^q = L^q(\Omega_2)$ as the space Ω_i is clear from the context. Recall that q' denotes the Hölder conjugate of q, i.e. $\frac{1}{q} + \frac{1}{q'} = 1$. By a Hahn-Banach argument and duality of L^p spaces, it suffices to show that

$$|\int_{\Omega_2} (Tf) g \, d\mu_2| \leq C_{p,q}$$

for simple functions $f : \Omega_1 \to \mathbb{C}$ and $g : \Omega_2 \to \mathbb{C}$ with finite measure support and with $\|f\|_{L^p}, \|g\|_{L^{q'}} = 1$ (*why?!*, check this in particular for limiting cases $p, q' = \infty$). Our goal is to apply Hadamard's three lines lemma. Let us factorize f and g as follows

$$f = \mathrm{e}^{i \arg(f)} F_1^{1-\theta} F_2^{\theta}, \quad g = \mathrm{e}^{i \arg(g)} G_1^{1-\theta} G_2^{\theta},$$

with[11] F_i, G_i non-negative and such that $F_i \in L^{p_i}$, $G_i \in L^{q_i'}$ with norms $\|F_i\|_{L^{p_i}} = \|G_i\|_{L^{q_i'}} = 1$ — this can be achieved by setting $F_i = |f|^{\frac{p}{p_i}}$ and $G_i = |g|^{\frac{q'}{q_i'}}$. Now define $h : \mathbb{C} \to \mathbb{C}$ by

$$h(z) = \int_{\Omega_2} T(\mathrm{e}^{i \arg(f)} F_1^{1-z} F_2^{z}) \mathrm{e}^{i \arg(g)} G_1^{1-z} G_2^{z} \, d\mu_2,$$

which is well-defined and entire since F_i and G_i are simple functions with finite measure support: More precisely, this by the fact that there exists pairwise disjoint measurable sets $A_j \subset \Omega_1$, and $B_k \subset \Omega_2$, and $a_j, b_j, c_k, d_k > 0$, for $j = 1, .., m$, $k = 1, .., l$ such that

$$h(z) = \sum_{j,k} a_j^{1-z} b_j^{z} c_k^{1-z} d_k^{z} \int_{\Omega_2} T(\mathrm{e}^{i \arg(f)} \chi_{A_j}) \mathrm{e}^{i \arg(g)} \chi_{B_k} \, d\mu_2.$$

To apply Lemma 5.12, it remains to check that h is bounded in modulus along the lines $i\mathbb{R}$ and $1 + i\mathbb{R}$. Let us first check that $F_1^{1-z} F_2^{is} \in L^{p_1}$ and $G_1^{1-is} G_2^{is} \in L^{q_1'}$. This holds with norms equal to 1 since $|F_1^{1-is} F_2^{is}| = F_1$ and $|G_1^{1-is} G_2^{is}| = G_1$ (since F_i, G_i are nonnegative). Hence, for $s \in \mathbb{R}$,

$$\begin{aligned}
|h(is)| &= \left| \int_{\Omega_2} T(\mathrm{e}^{i \arg(f)} F_1^{1-is} F_2^{is}) \mathrm{e}^{i \arg(g)} G_1^{1-is} G_2^{is} \, d\mu_2 \right| \\
&\leq \|T(\mathrm{e}^{i \arg(f)} F_1^{1-is} F_2^{is})\|_{L^{q_1}} \, \|G_1^{1-is} G_2^{is}\|_{L^{q_1'}} \\
&\leq C_1 \|F_1^{1-is} F_2^{is})\|_{L^{p_1}} \, \|G_1^{1-is} G_2^{is}\|_{L^{q_1'}} \\
&= C_1
\end{aligned}$$

[11]here, arg denotes a principal branch of the argument and we "set" $\mathrm{e}^{i \arg(x)} = 1$ if $x = 0$.

where the first estimate follows from Hölder's inequality and the second is clear by the assumed boundedness of T. Analogously, we get $|h(1+is)| \le C_2$ for all $s \in \mathbb{R}$. Thus by Lemma 5.12, $|h(z)| \le C_1^{1-\theta} C_2^{\theta}$ for $\Re z = \theta$. Setting $z = \theta$, the assertion follows since

$$h(\theta) = \int_{\Omega_2} (Tf) g \, d\mu_2.$$

□

Proof of Theorem 5.11. Let $k \in \mathbb{N}$. We apply the Riesz–Thorin Theorem to the operators $Tf = \tilde{f}$ defined from $L^{2k}(\mathbb{T})$ to $L^{2k+2}(\mathbb{T})$ and from $L^{2k+2}(\mathbb{T})$ to $L^{2k+2}(\mathbb{T})$. By Lemma A, T is bounded on both spaces and hence Theorem 5.13 yields that $T : L^p(\mathbb{T}) \to L^p(\mathbb{T})$ is bounded as well for all $p \in (2k, 2k+2)$. Since this holds for arbitrary $k \in \mathbb{N}$, we conclude that L^p is invariant under conjugation for all $p \ge 2$ and thus the assertion of the Theorem holds for such p by Proposition 5.9.
The final step — that the statement also holds for $p \in (1,2)$ now follows by duality: For that consider the dual operator $T' : (L^q(\mathbb{T}))' \to (L^q(\mathbb{T}))'$ of the conjugation operator on $L^q(\mathbb{T})$ for $q > 2$. By duality of L^q-spaces and the definition of the dual operator, this means that

$$\langle f, Tg \rangle_{L^{q'}, L^q} = \langle T'f, g \rangle_{L^{q'}, L^q} \qquad \forall f \in L^{q'}(\mathbb{T}), g \in L^q(\mathbb{T}),$$

where q is the Hölder conjugate of q, i.e. $\frac{1}{q} + \frac{1}{q'} = 1$. Let us consider $f, g \in \mathrm{Trig}(\mathbb{T})$. Then by $Tf = \tilde{f}$,

$$2\pi i \sum_{m=-N}^{N} \hat{f}(m)\overline{\hat{g}(m)} \mathrm{sgn}(m) = 2\pi \sum_{m=-N}^{N} \widehat{T'f}(m)\overline{\hat{g}(m)},$$

and therefore, $\widehat{T'f}(m) = i\mathrm{sgn}(m)\hat{f}(m) = -\hat{\tilde{f}}(m)$ for all $m \in \{-N, .., N\}$ where N is chosen sufficiently large. By density of the trigonometric polynomials it follows that $T'f = -\tilde{f}$. Since with T also T' is a bounded operator, we hence conclude that $f \mapsto \tilde{f} = -T'f$ is bounded on $L^{q'}$. Since this holds for all $q \in (2, \infty)$, the assertion follows.

□

Remark 5.14. We emphasize that Theorem 5.11 does not show that the partial sums $(D_n * f)_{n\in\mathbb{N}}$ of the Fourier series of $f \in L^p(\mathbb{T})$ converge pointwise a.e. to $f(t)$ (compare with Ex. 0.1). However, this statement is in fact true, but relies on a deep result by Carleson–Hunt from the 1960ies.

We give another application of the Riesz–Thorin interpolation theorem, which comes natural knowing that $f \mapsto \hat{f}$ is a bounded operator from both $L^1(\mathbb{T})$ to $\ell^\infty(\mathbb{Z})$ and $L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$.

Theorem 5.15 (**Hausdorff–Young**). *For $p \in [1,2]$ and $q \in [2,\infty]$ defined by $1 = \frac{1}{p} + \frac{1}{q}$, the operator $f \mapsto \hat{f}$ is bounded from $L^p(\mathbb{T})$ to $\ell^q(\mathbb{Z})$ with operator norm equal to $(2\pi)^{-\frac{2}{q}}(2\pi)^{-\frac{1}{q}}$.*

Proof. See Exercise 6.1. The claim holds for $p = 1$ and $p = 2$ by Theorem 1.6 and Hilbert space theory respectively. Thus we can apply Theorem 5.13 for the σ-finite measure spaces $(\mathbb{T}, \lambda)$ and $(\mathbb{Z}, |\cdot|)$ — $|\cdot|$ refers to the counting measure — and get that $f \mapsto \hat{f}$ is bounded from $L^p(\mathbb{T})$ to $\ell^q(\mathbb{Z})$ for all pairs (p,q) of the form

$$\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2}, \qquad \frac{1}{q} = \frac{1-\theta}{\infty} + \frac{\theta}{2}, \qquad \theta \in (0,1).$$

Hence, $\frac{1}{p} = 1 - \frac{\theta}{2}$ and $q = \frac{2}{\theta}$, which gives the assertion.

□

6. Coming back to the heat equation

see notes.

CHAPTER 2

The Fourier tranform

Note: In the following L^p will — unless state otherwise — refer to $L^p(\mathbb{R}) = L^p(\mathbb{R}, \lambda)$ where λ refers to the Lebesgue measure (on the Lebesgue measurable sets).

Motivated by the question

How to represent a non-periodic function on $\mathbb{R}$ in an analogous way as with Fourier series?

we are lead to a "continuous" transformation of a function f — in contrast to the Fourier series which is "discrete".

1. Basics on the Fourier transform and the Schwartz class

Definition 1.1 (Fourier transform on L^1). For $f \in L^1(\mathbb{R})$ the function $\mathcal{F}(f) : \mathbb{R} \to \mathbb{C}$ defined by

$$\mathcal{F}(f)(s) = \int_{\mathbb{R}} f(s)\mathrm{e}^{-ist}\, ds, \qquad s \in \mathbb{R}$$

is called **Fourier transform** of f.

In the following we use the operators, $\omega \in \mathbb{R} \setminus \{0\}$, $g \in L^\infty(\mathbb{R})$,

$$\tau_w f = f(\cdot + \omega), \qquad m_g f = fg, \qquad Rf = f(-\cdot), \qquad D_\omega f = f(w\cdot)$$

defined for $f \in L^p(\mathbb{R})$ for any $p \in [1, \infty]$.

Theorem 1.2 (Basics of the Fourier transform). *Let $\omega \in \mathbb{R} \setminus \{0\}$.*

(1) *$\mathcal{F} : L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$ is linear and bounded*

(2) *The range* ran $\mathcal{F}$ *of $\mathcal{F}$ lies in $C_0(\mathbb{R}) = \{f \in C(\mathbb{R})\colon \lim_{x\to\pm\infty} f(x) = 0\}$, that is $\mathcal{F} \in C_0(\mathbb{R})$ for all $f \in L^1(\mathbb{R})$ (Riemann–Lebesgue)*

(3) *For $\mathrm{e}_{i\omega}(s) := \mathrm{e}^{i\omega s}$ we have*

$$\begin{aligned}
\mathcal{F} m_{\mathrm{e}_{-i\omega}} &= \tau_\omega \mathcal{F} && \textit{(Modulation)}\\
\mathcal{F}\tau_\omega &= m_{\mathrm{e}_{i\omega}} \mathcal{F} && \textit{(Translation)}\\
\mathcal{F}R &= R\mathcal{F} && \textit{(Reflection)}\\
\mathcal{F}\,\bar{\cdot} &= \overline{R\mathcal{F}} && \textit{(Conjugation)}\\
\mathcal{F}D_\omega &= m_{1/|\omega|} D_{\frac{1}{\omega}} \mathcal{F} && \textit{(Dilation)}
\end{aligned}$$

considered on $L^1(\mathbb{R})$.

(4) *$\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g)$ (Convolution Theorem)*

(5) *for any $f \in C^k(\mathbb{R})$ such that $f^{(\ell)} \in L^1(\mathbb{R})$ for all $\ell = 0, .., k$,*

$$\mathcal{F}(f^{(k)}) = m_{(i\mathbf{s})^k} \mathcal{F}(f),$$

where $i\mathbf{s}$ refers to the function $s \mapsto is$. Conversely, if $x \mapsto x^\ell f(x) \in L^1(\mathbb{R})$ for all $\ell = 0, .., k$,

$$[\mathcal{F}(f)]^{(k)} = \mathcal{F}(m_{(-i\mathbf{s})^k} f)$$

PROOF. See Exercise 6.1. □

REMARK 1.3. Compared to the definition of the Fourier series for a given function in $L^1(\mathbb{T})$, it is not clear how to define the Fourier transform of a general function in $L^p(\mathbb{R})$ for $p > 1$ since the integral defining $\mathcal{F}$ may not converge absolutely then: Recall that $L^p(\mathbb{R}) \not\subset L^1(\mathbb{R})$ since $\lambda(\mathbb{R}) = \infty$! This shows already a crucial technical difficulty when dealing with the Fourier transform of a function. However, as we shall see shortly, it is possible to 'define' the Fourier transform on $L^2(\mathbb{R})$ in a consistent way. In fact, one may naively think that this is possible if the integral $\int_{\mathbb{R}} \mathrm{e}^{ist}$ is "allowed" to converge conditionally, which is not meaningful in terms of the Lebesgue integral however.

In the following we will use the notation

$$\partial^\alpha f = \frac{\partial^\alpha f}{\partial x^\alpha} = f^{(\alpha)}$$

for $\alpha \in \mathbb{N}_0$ and a α-times differentiable function $f : \mathbb{R} \to \mathbb{C}$.

DEFINITION 1.4 (The Schwartz class). A function $f \in C^\infty(\mathbb{R})$ is defined to be in the **Schwartz class** $S(\mathbb{R})$ if

$$\forall \alpha, \beta \in \mathbb{N}_0 : \qquad \rho_{\alpha,\beta}(f) := \sup_{x \in \mathbb{R}} |x^\alpha \partial^\beta f(x)| < \infty.$$

Note that $(x \mapsto \mathrm{e}^{-x^2})$ and any compactly supported $C^\infty(\mathbb{R})$-function is in $S(\mathbb{R})$. Further note that it is easily seen that a function $f \in C^\infty(\mathbb{R})$ lies in $S(\mathbb{R})$ if and only if

$$\forall N, \beta \in \mathbb{N}_0 \exists C_{N,\beta} > 0 \, \forall x \in \mathbb{R} : \qquad |f^{(\beta)}(x)| \leq C(1 + |x|)^{-N}.$$

If $f \in S(\mathbb{R})$, the property follows by the binomial formula. Conversely, we have $|x^\alpha f^{(\beta)}(x)| \leq C \frac{|x|^\alpha}{(1+|x|)^N}$ which is bounded on $\mathbb{R}$ for $N \geq \alpha$. Let us introduce the following type of topology on the Schwartz space.

DEFINITION 1.5 (Convergence in the Schwartz space). Let $(f_n)_{n \in \mathbb{N}} \subset S(\mathbb{R})$ and $f \in S(\mathbb{R})$. We say that the sequence (f_n) converges to f in $S(\mathbb{R})$, $f_n \overset{S}{\to} f$, if

$$\forall \alpha, \beta \in \mathbb{N} : \qquad \rho_{\alpha,\beta}(f - f_n) = \sup_{x \in \mathbb{R}} |x^\alpha (f - f_n)^{(\beta)}(x)| \to 0, \qquad \text{as } n \to \infty.$$

REMARK 1.6. One may wonder how the sequential convergence defined in Def. 1.5 relates to a topology on the space $S(\mathbb{R})$. In fact, this convergence can be shown to be equivalent to convergence of a sequence in the following metric defined on $S(\mathbb{R})$

$$d(f, g) = \sum_{j \in \mathbb{N}} 2^{-j} \frac{\rho_{\alpha_j,\beta_j}(f - g)}{1 + \rho_{\alpha_j,\beta_j}(f - g)},$$

where $j \mapsto (\alpha_j, \beta_j)$ is a bijection from $\mathbb{N} \to \mathbb{N}_0 \times \mathbb{N}_0$. See Ex. 6.2. In other words, the local convex topological vector space defined by defined by the seminorms $\rho_{\alpha,\beta}$, $\alpha, \beta \in \mathbb{N}_0$ — as the initial topology for which $\rho_{\alpha,\beta}$ is continuous — is metrizable. Moreover, it can be shown that the above metric even yields a complete space.

PROPOSITION 1.7 (Properties of $S(\mathbb{R})$).

(1) *The inclusions $C_c^\infty(\mathbb{R}) \subset S(\mathbb{R}) \subset L^p(\mathbb{R})$ are dense in the $\|\cdot\|_{L^p}$-norm for $p \in [1, \infty)$.*
(2) *$S(\mathbb{R})$ is continuously embedded in $L^p(\mathbb{R})$ for all $p \in [1, \infty]$* [1].

[1] Here "continuously embedded" means that the identity is sequentially continuous, i.e. $f_n \overset{S}{\to} f$ implies that $f_n \overset{L^p}{\to} f$. However, by Remark 1.6, sequential continuity coincides with continuity between the topological spaces.

(3) *For $f, g \in S(\mathbb{R})$ we have that fg and $f * g$ are elements of $S(\mathbb{R})$ as well as $\partial^\alpha f \in S(\mathbb{R})$ with $\partial^\alpha(f * g) = \partial^\alpha f * g = f * \partial^\alpha g$.*

PROOF. See Ex. 6.2. for (1), (2) and (3) without the part on $\partial^\alpha f$. The fact that for $f \in S(\mathbb{R})$ also $\partial^\alpha f \in S(\mathbb{R})$ simply follows by definition of $S(\mathbb{R})$. To see that ∂^α commutes with the convolution consider $\alpha = 1$: Since $s \mapsto |\partial_t f(t-s)g(s)|$ is dominated by an integrable function of the form $s \mapsto C(1+|t-s|)^{-2}\|g\|_\infty$, we conclude that $\partial(f * g) = \partial f * g$, which, by symmetry also equals $f * \partial g$ by (a consequence of) dominated convergence. The rest follows by induction. $\square$

DEFINITION 1.8 (The inverse Fourier transform). For $f \in L^1(\mathbb{R})$ the function $\mathcal{F}^*(f)$ defined by

$$\mathcal{F}^\star(f)(s) = \frac{1}{2\pi}\mathcal{F}(f)(-s) = \frac{1}{2\pi}\int_{\mathbb{R}} f(s)\mathrm{e}^{ist}\, ds$$

is called "inverse Fourier transfrom" of f.

Let us show that the latter notion is meaningful.

LEMMA 1.9. *Let $f, g, h \in S(\mathbb{R})$. Then*

(1) $\int_{\mathbb{R}} f\mathcal{F}(g)\, d\lambda = \int_{\mathbb{R}} (\mathcal{F}f)g\, d\lambda$.
(2) $\mathcal{F}(\mathcal{F}^\star(f)) = f = \mathcal{F}^\star(\mathcal{F}(f))$
(3) $\langle f, h\rangle = \frac{1}{2\pi}\langle \mathcal{F}(f), \mathcal{F}(h)\rangle$
(4) $\langle \mathcal{F}f, \mathcal{F}^\star \overline{h}\rangle = \langle f, h\rangle$
(5) $\|f\|_{L^2(\mathbb{R})} = \sqrt{\frac{1}{2\pi}}\|\mathcal{F}(f)\|_{L^2(\mathbb{R})} = \sqrt{2\pi}\|\mathcal{F}^\star f\|_{L^2(\mathbb{R})}$ *(Parseval's identity)*

where $\langle f, h\rangle = \int_{\mathbb{R}} f(s)\overline{g(s)}\, ds$.

PROOF. Assertion (1) follows by Fubini's theorem. The idea to prove (2) is as follows: Choose a function $h \in L^1(\mathbb{R})$ with the following properties

- $h \in S(\mathbb{R})$,
- $\frac{1}{2\pi}\mathcal{F}(h(\frac{\cdot}{2\pi}) = h$,
- $h(0) = 1$.

The existence of such a function will be shown in the exercises (hint: $h(x) = a\mathrm{e}^{bx^2}$ for suitable $a, b \in \mathbb{R}$). The idea is to represent f by an approximate identity, $f = \lim_{\lambda\to\infty}\int_{\mathbb{R}} f(s)h_\lambda(t-s)\, ds$ and then use (1) in order to find to prove the equality in (2). By exercise 6.2 and the assumptions on h, $h_\lambda = \lambda h(\lambda\cdot)$, $\lambda > 0$, defines an approximate identity. Hence, by a version of Theorem 3.7, we have that $f(t) = \lim_{\lambda\to\infty}(f * h_\lambda)(t)$. On the other hand — motivated by idea to choose g such that $\mathcal{F}(g) = h_\lambda(t-\cdot) = \lambda R\tau_t D_\lambda h$ — we find $g(t) = \frac{1}{2\pi}Rm_{\mathrm{e}_{-it}}D_{\frac{1}{2\pi\lambda}}h$, where we used the basic properties of the Fourier transform and that $\mathcal{F}(h) = 2\pi h$. By (1), we now have

$$\begin{aligned}(f * h_\lambda)(t) = \int_{\mathbb{R}} f(s)\mathcal{F}(g)(s)\, ds &= \int_{\mathbb{R}} (\mathcal{F}(f)(s))g(s)\, ds \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} (\mathcal{F}(f)(s))(Rm_{\mathrm{e}_{-it}}D_{\frac{1}{\lambda}}h)(s)\, ds \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} (\mathcal{F}(f)(s))\mathrm{e}^{its}h(\frac{-s}{2\pi\lambda})\, ds\end{aligned}$$

Since $h(0) = 1$, the last integral converges to $\frac{1}{2\pi}\int_{\mathbb{R}}(\mathcal{F}(f)(s))\mathrm{e}^{its}\, ds$ for $\lambda \to \infty$, by dominated convergence. Altogether, we thus have $f(t) = (\mathcal{F}^\star\mathcal{F}f)(t)$ for all $t \in \mathbb{R}$. The other identity follows by applying what we have shown to $Rf = f(-\cdot)$.
Assertion (3) and (4) follows from (1) and (2) by setting $g = \mathcal{F}^\star(\overline{h})$ and noting that $\mathcal{F}^\star(\overline{h}) = \frac{1}{2\pi}\overline{\mathcal{F}(h)}$. Assertion (5) follows direcltly from (3). $\square$

THEOREM 1.10 (The Fourier transform on $S(\mathbb{R})$). *The Fourier transform is a well-defined linear mapping from $S(\mathbb{R})$ to $S(\mathbb{R})$. Moreover,*

(1) *$\mathcal{F}$ is continuous and bijective on $S(\mathbb{R})$ (with in the natural topology on $S(\mathbb{R})$)*
(2) *$(2\pi)^{-\frac{1}{2}}\mathcal{F}$ is isometric on $S(\mathbb{R})$ with respect to the L^2-norm, i.e.*

$$\forall f \in S(\mathbb{R}): \qquad \|\mathcal{F}(f)\|_{L^2(\mathbb{R})} = (2\pi)^{\frac{1}{2}}\|f\|_{L^2(\mathbb{R})}$$

Thus $\mathcal{F}$ is also continuous with continuous inverse on the $S(\mathbb{R})$ equipped with the L^2-norm.

The inverse $\mathcal{F}^{-1}$ on $S(\mathbb{R})$ is given by $\mathcal{F}^\star$.

PROOF. By Lemma 1.9(2), $\mathcal{F}$ is bijective on $S(\mathbb{R})$ with inverse $\mathcal{F}^{-1} = \mathcal{F}^\star$. For the continuity in (1) see Ex. 6.2 and use the analogous argument for the continuity of $\mathcal{F}^\star$. Assertion (2) follows from Lemma 1.9(5). $\square$

2. The Fourier transform on L^2

THEOREM 2.1 (The Fourier transform on L^2). *The definition of the Fourier transform $\mathcal{F}$ can be extended to functions in $L^2(\mathbb{R})$ such that $(2\pi)^{\frac{1}{2}}\mathcal{F}$ is an isometric isomorphism from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$* [2]. *For functions f in $L^1(\mathbb{R})\cap L^2(\mathbb{R})$, this abstract extension coincides pointwise almost everywhere with the action of $\mathcal{F}$.*

PROOF. Let $\tilde{\mathcal{F}}: L^2 \to L^2$ denote the unique bounded operator which extends the operator $\mathcal{F}: S(\mathbb{R}) \to S(\mathbb{R})$ such that $\tilde{\mathcal{F}}|_{L^2} = \mathcal{F}$ — note that we use the density of $S(\mathbb{R})$ in L^2 here. By Theorem 1.10, we even have that

$$\forall f \in L^2(\mathbb{R}): \qquad \|\tilde{\mathcal{F}}(f)\|_{L^2(\mathbb{R})} = (2\pi)^{\frac{1}{2}}\|f\|_{L^2(\mathbb{R})}$$

Thus $(2\pi)^{-\frac{1}{2}}\tilde{\mathcal{F}}$ is an isometric isomorphism. Let $f \in L^1 \cap L^2$. It remains to show that $\tilde{\mathcal{F}}(f) = \mathcal{F}f$ pointwise almost everywhere. Note that this is not clear a-priori as $\tilde{\mathcal{F}}$ is only defined as operator on $L^2(\mathbb{R})$ and $\tilde{\mathcal{F}}f$ is defined as the L^2-limit of the Cauchy sequence $\mathcal{F}f_n$ where $(f_n)_{n\in\mathbb{N}}$ is a sequence in $S(\mathbb{R})$ which converges to $f \in L^2(\mathbb{R})$ in the L^2-norm, i.e. $\|\mathcal{F}f_n - \tilde{\mathcal{F}}f\|_{L^2} = 0$ [3]. Let $g_n = f\chi_{\{|x|\le n\}}$ and let $f_n \in C^\infty(\mathbb{R})$ with support in $[0,n]$ such that $\|g_n - f_n\|_{L^2(\mathbb{R})} = \|g_n - f_n\|_{L^2([0,n])} < \frac{1}{n}$. Then $\|f - f_n\|_{L^2(\mathbb{R})} \to 0$ as $n \to \infty$ and also for fixed $s \in \mathbb{R}$,

$$|(\mathcal{F}f - \mathcal{F}f_n)(s)| \le \|f - f_n\|_{L^1(\mathbb{R})} \le \|f - g_n\|_{L^1(\mathbb{R})} + \|g_n - f_n\|_{L^1(0,n)}$$

By definition of g_n, it is easy to see that the term $\|f - g_n\|_{L^1(\mathbb{R})}$ goes to 0 as $n \to \infty$ since $f \in L^1$. The second term $\|g_n - f_n\|_{L^1(0,n)}$, can be estimated using Cauchy-Schwarz so that we get

$$\|g_n - f_n\|_{L^1(0,n)} \le \sqrt{n}\|g_n - f_n\|_{L^2(0,n)} \le \frac{1}{\sqrt{n}} \to 0, \quad (n \to \infty).$$

Thus $\mathcal{F}f_n(s) \to \mathcal{F}f(s)$ for any $s \in \mathbb{R}$. On the other hand, we know that L^2-convergence always implies the existence of a subsequence which converges pointwise almost everywhere (see Ex. 1.0) — therefore, $\mathcal{F}f_{n_k} \to \tilde{\mathcal{F}}f$ pointwise almost everywhere. Thus, $\mathcal{F}f(s) = \tilde{\mathcal{F}}f(s)$ for a.e. $s \in \mathbb{R}$. $\square$

REMARK 2.2. Note that isometric isomorphisms on Hilbert spaces are nothing else than unitary operators — the latter being defined as operators $T: X \to X$ such that $T^*T = T^*T = I$.

[2] Note that we identify the operator $\mathcal{F}$ with its "extension" defined on L^2.

[3] As we know from Ex.1.0, L^2-convergence need not imply pointwise convergence.

Theorem 2.3 (Hausdorff–Young inequality for the Fourier transform). *Let $p \in (1,2)$ and let q denote the Hölder conjugate of p. The Fourier transform $\mathcal{F}$ can be uniquely extended to a bounded operator from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ with*

$$\|\mathcal{F}f\|_{L^q} \leq C_p \|f\|_{L^p} \qquad \forall f \in L^p(\mathbb{R}),$$

where $C_p = (2\pi)^{1-\frac{1}{p}}$.

Proof. This is an application of Riesz–Thorin's theorem since $\mathcal{F}$ act as (extension of a) bounded operator on the spaces

$$\begin{aligned} L^1 \to L^\infty, \quad & \|\mathcal{F}f\|_{L^\infty} \leq \|f\|_{L^1} \\ L^2 \to L^2, \quad & \|\mathcal{F}f\|_{L^2} \leq (2\pi)^{\frac{1}{2}} \|f\|_{L^2}, \end{aligned}$$

see the basics of the Fourier transform and Theorem 2.1. This thus implies that $\|Tf\|_{L^p} \leq C_p\|f\|_{L^q}$ for simple functions f with finite measure support and p, q, θ linked through (5.6). The constant C_p can be computed from $C_p = 1^{1-\theta}(2\pi)^{\frac{\theta}{2}}$ where θ is given through $\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2}$. Thus $C = (2\pi)^{\frac{\theta}{2}} = (2\pi)^{1-\frac{1}{p}}$. By density of the simple functions with finite measure support in L^p, $p < \infty$, the assertion follows. $\square$

3. Tempered distributions

In the previous section we have seen how to define the Fourier transform of a function $f \in L^p$ for $p \in [1,2]$. This approach, however, can be seen to fail if $p > 2$.

Definition 3.1. Any linear functional $u : S(\mathbb{R}) \to \mathbb{C}$ which is continuous with respect to sequential convergence on $S(\mathbb{R})$, i.e.

$$f_n \overset{S}{\to} f \qquad \Longrightarrow \qquad u(f_n - f) \to 0,$$

as $n \to \infty$ is called *tempered distribution*. The set of all tempered distributions is denoted by $S'(\mathbb{R})$.

Proposition 3.2. *A linear functional*[4] *$u : S(\mathbb{R}) \to \mathbb{C}$ is a tempered distribution if and only if there exists $k, m \in \mathbb{N}_0$ and $C > 0$ such that*

$$|u(f)| \leq C \sum_{\beta \leq m, \alpha \leq k} \rho_{\alpha,\beta}(f)$$

Proof. If a linear functional u satisfies such an estimate, we trivially have that u is continuous in the sense of Definition 3.1. Conversely, let u be a tempered distribution. Let us first prove the statement

$$\exists m, k \in \mathbb{N}_0, \delta > 0\, \forall g \in S(\mathbb{R}) : \quad (\rho_{\alpha,\beta}(g) \leq \delta \quad \forall \alpha \leq m, \beta \leq k \implies |u(g)| < 1) \tag{3.1}$$

To do so, assume conversely that the statement was false:

$$\forall m, k \in \mathbb{N}_0, \delta > 0\, \exists g \in S(\mathbb{R}) : \quad \rho_{\alpha,\beta}(g) \leq \delta \quad \forall \alpha \leq m, \beta \leq k \wedge |u(g)| > 1$$

But this implies that (setting $\delta = \frac{1}{n}$) there exists a sequence g_n (g_n) in $S(\mathbb{R})$ such that

$$\forall m, k \in \mathbb{N}_0 : \quad \rho_{m,k}(g_n) \leq \frac{1}{n} \wedge |u(g_n)| > 1$$

[4]Here "linear functional" simply refers to a *linear* mapping that maps to the field of complex numbers.

This, however, contradicts the continuity of u. Hence, (3.1) holds. For general $f \in S(\mathbb{R})$ let $g = f \frac{\delta}{\sum_{\alpha \le m, \beta \le k} \rho_{\alpha,\beta}(f)}$. Clearly, $\rho_{\alpha,\beta}(g) \le \delta$ for $\alpha \le m, \beta \le k$ and hence $|u(g)| < 1$. By linearity of u, this rewrites to

$$|u(f)| < \frac{1}{\delta} \sum_{\alpha \le m, \beta \le k} \rho_{\alpha,\beta}(f),$$

which shows the assertion for $C = \frac{1}{\delta}$. □

Clearly, $S'(\mathbb{R})$ is a vector space. We can equip it with a topology in a natural way:

Definition 3.3. A sequence of tempered distributions $u_n \in S'(\mathbb{R})$ converges to $u \in S'(\mathbb{R})$ if

$$u_n(f) = \langle u_n, f \rangle \to \langle u, f \rangle = u(f) \quad (n \to \infty) \qquad \forall f \in S(\mathbb{R}).$$

Example 3.4 ((see Exercise sheet 7).
(1) Dirac δ_0 is in $S'(\mathbb{R})$.
(2) $L^p(\mathbb{R})$ functions, $p \in [1, \infty]$ are in $S'(\mathbb{R})$ via the identification $f \mapsto \int_{\mathbb{R}} u(t) f(t)\, dt$.
(3) Finite Borel measures μ on $\mathbb{R}$ via $f \mapsto \int_{\mathbb{R}} f(x) d\mu(x)$.
(4) $u(x) = \log|x|$.
(5) $\langle u, \phi \rangle = \lim_{\epsilon \to 0^+} \int_{|x| \ge \epsilon} \frac{\phi(x)}{x}\, dx$

Remark 3.5 (Motivation for Fourier transform of more general functions). We have seen that the definition of the Fourier transform — a-priori only made on $L^1(\mathbb{R})$ — can be lifted to L^p, $p \in [1, 2]$ with the help of special properties of the transform (boundedness in the L^2-norm $\rightsquigarrow$ Parseval's identity) and basic functional analysis (extension of a bounded operator as well as interpolation). It can be shown (by means of examples) that this procedure can not be extended to L^p for $p > 2$, and hence particularly not for functions in L^∞. However, a ('version' of the) Fourier transform in the latter cases is desirable: for instance to obtain an analogy to the theory we have encountered on Fourier series and on the other hand since signals, which we would like to regard in the "frequency domain" (that is, as Fourier transforms of other functions/objects) in application. Let us elaborate on this in more detail: Coming back to the "thresholding" we have already touched in the introduction of this course, consider an input–output behavior of a signal u given by

$$y = \mathcal{F}^{-1}(h \cdot \mathcal{F}u)$$

where h is a fixed function. For example, h may be an indicator function that 'selects' only specific frequency values $\omega \in \mathbb{R}$, $h(\omega) = 1$, and sets the others to zero, $h(\omega) = 0$. Formally, we would expect from the convolution theorem that y can be rewritten as

$$y = \mathcal{F}^{-1}(h) * u,$$

which, however, may not be well-defined as $\mathcal{F}^{-1} = \mathcal{F}^\star$ is only defined for $h \in L^p(\mathbb{R})$, $p \in [1, 2]$. But do we get if e.g. $h = \chi_{[0,\infty)}$? Shortly, we will see how to abstractly circumvent this problem, but before let us approach the answer intuitively. For that let $h \equiv 1$ and we would like to find a suitable definition for $\mathcal{F}h$.

Definition 3.6 (**The Fourier transform for tempered distributions**). For a tempered distribution $u \in S'(\mathbb{R})$, we define the Fourier transform $\mathcal{F}(u)$ as the tempered distribution given by

$$\langle \mathcal{F}u, \phi \rangle = \langle u, \mathcal{F}\phi \rangle \quad \forall \phi \in S(\mathbb{R}).$$

Note that the $\mathcal{F}u$ is indeed well-defined as a distribution since $\mathcal{F}$ maps $S(\mathbb{R})$ to $S(\mathbb{R})$, Theorem 1.10.

PROPOSITION 3.7. *The distributional definition of the Fourier transform is consistent with the one for $L^p(\mathbb{R})$, $p \in [1,2]$, where we identify functions f with tempered distributions u_f by means of*

$$f \mapsto \langle u_f, \psi \rangle = \int_{\mathbb{R}} f(s)\psi(s)\,ds.$$

PROOF. In Exercise 7.1 we showed that $u_f \in S'(\mathbb{R})$ for $f \in L^p(\mathbb{R})$, $p \in [1,\infty]$. To show the assertion consider □

DEFINITION 3.8 (Modulation, translation, reflection, dilation, convolution on $S'(\mathbb{R})$). Let $u \in S'(\mathbb{R})$, $f \in S(\mathbb{R})$, $w \in \mathbb{R} \setminus \{0\}$, $\alpha \in \mathbb{N}$. We extend the definitions of the following operators to $S'(\mathbb{R})$:

$$\tau_w, \qquad m_f, \qquad R, \qquad D_\omega$$

by

(1) $\langle \tau_w u, \phi \rangle = \langle u, \tau_{-w}\phi \rangle$
(2) $\langle m_f u, \phi \rangle = \langle u, m_f \phi \rangle$
(3) $\langle Ru, \phi \rangle = \langle u, R\phi \rangle$
(4) $\langle D_w u, \phi \rangle = \langle u, m_{\frac{1}{|\omega|}} D_{\frac{1}{\omega}} \phi \rangle$
(5) $\langle u * f, \phi \rangle = \langle u, Rf * \phi \rangle$
(6) $\langle \partial^\alpha u, \phi \rangle = \langle u, (-1)^\alpha \partial \phi \rangle$

for all $\phi \in S(\mathbb{R})$.

It is easy to see that Definition 3.8 indeed extends the definition of the operators on $L^1(\mathbb{R})$. Note that the convolution $u * f$ of a tempered distribution and a function f is in general not well-defined if $f \notin S(\mathbb{R})$ as $Rf * \phi$ may not be in $S(\mathbb{R})$ for all $\phi \in S(\mathbb{R})$ then.

PROPOSITION 3.9 (Convolution of tempered distribution with Schwartz function). *Let $\phi \in S(\mathbb{R})$ and $u \in S'(R)$. Then $u * \phi \in C^\infty(\mathbb{R})$ and*

$$(u * \phi)(t) = \langle u, \tau_{-t} R\phi \rangle,$$

where we identify a function with its corresponding tempered distribution. (Note that the $u * \phi$ was defined by $\langle u * \phi, \psi \rangle = \langle u, R\phi * \psi \rangle$.)

PROOF. We only prove the identity (which is clear in the special case where $u \in S(\mathbb{R})$). Let $u \in S'(\mathbb{R})$ and $\psi \in S(\mathbb{R})$.

$$\begin{aligned}\langle u * \phi, \psi \rangle = \langle u, R\phi * \psi \rangle &= \langle u, \int_{\mathbb{R}} (R\phi)(\cdot - s)\psi(s)\,ds \rangle \\ &= \langle u, \int_{\mathbb{R}} \tau_{-s} R\phi(\cdot)\psi(s)\,ds \\ &\stackrel{!}{=} \int_{\mathbb{R}} \langle u, \tau_{-s} R\phi(\cdot) \rangle \psi(s)\,ds\end{aligned}$$

where the last step follows since it can be shown that the Riemann sums of the $S(\mathbb{R})$-valued integral converge in $S(\mathbb{R})$, for a detailed argument see e.g. [**1**, Thm. 2.3.20]. This shows the identity. To show that $(u * \phi) \in C^\infty$, one explicitly shows that $t \mapsto \langle u, \tau_{-t} R\phi \rangle$ is C^∞ which follows since $t \mapsto \tau_{-t} R\phi$ is continuous from $\mathbb{R}$ to $S(\mathbb{R})$ (which in turn holds as the translation operator $t \mapsto \tau_{-t}$ is strongly continuous). □

THEOREM 3.10 (Fourier transform on tempered distributions). (1) *$\mathcal{F} : S'(\mathbb{R}) \to S'(R)$ is continuous.*

(2) *We have that*

$$\begin{aligned}
\mathcal{F} m_{\mathrm{e}_{-i\omega}} &= \tau_\omega \mathcal{F} && \textit{(Modulation)}\\
\mathcal{F} \tau_\omega &= m_{\mathrm{e}_{i\omega}} \mathcal{F} && \textit{(Translation)}\\
\mathcal{F} R &= R \mathcal{F} && \textit{(Reflection)}\\
\mathcal{F}\, \bar{\cdot} &= \overline{R\mathcal{F}} && \textit{(Conjugation)}\\
\mathcal{F} D_\omega &= m_{1/|\omega|} D_{\frac{1}{\omega}} \mathcal{F} && \textit{(Dilation)}
\end{aligned}$$

where the corresponding operators are defined in Def. 3.8.

(3) *Let* $u \in S'(\mathbb{R})$ *and* $\phi \in S(\mathbb{R})$*, then*

$$\mathcal{F}(u * \phi) = \mathcal{F}(u) \cdot \mathcal{F}(\phi) \quad \textit{and} \quad \mathcal{F}(u\phi) = \mathcal{F}(u)\mathcal{F}(\phi)$$

(Convolution Theorem)

(4) *for any* $u \in S'(\mathbb{R})$*,* $\alpha \in \mathbb{N}$ *we have that*

$$\mathcal{F}(\partial^\alpha u) = m_{(i\mathbf{s})^\alpha} \mathcal{F}(u), \quad \textit{and} \quad \partial^\alpha \mathcal{F}(u) = \mathcal{F}(m_{(-i\mathbf{s})^\alpha} u)$$

PROOF. The assertions follow directly by the definition of the considered operators via duality and the basics of the Fourier transform (for L^1- and hence $S(\mathbb{R})$-functions). □

REMARK 3.11. It can be shown that Schwartz functions lie dense in tempered distributions.

4. Operators that commute with translations

At several occasions we have already encountered linear operators defined by convolution with fixed functions h, i.e.

$$T_h = h * \cdot$$

acting from L^p to L^q for some p, q. Thinking of the respective results on (partial sums of) Fourier series, the following fact does not come as a surprise.

PROPOSITION 4.1. *Let* $p, q, r \in [1, \infty]$ *with* $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$ *and* $h \in L^q(\mathbb{R})$*. Let* $T_h : L^p \to L^r$ *be defined by* $T_h f = h * f$*. Then* $T_h \in \mathcal{B}(L^p, L^q)$ *and*

$$T_h \tau_s = \tau_s T_h,$$

where $\tau_s f = f(\cdot + s)$*.*

This motivates the following definition.

DEFINITION 4.2. We say that an operator $T \in \mathcal{B}(L^p(\mathbb{R}), L^q(\mathbb{R}))$ *commutes with translations* if

$$T\tau_s = \tau_s T \qquad s \in \mathbb{R}.$$

Such operators appear quite naturally when studying transformation of signals and the goal of this section is to characterize when an operator is translation-invariant. Note that although we will be concerned with the "continuous" version here, that is, with functions defined on the whole real line, there exist corresponding results for the torus $\mathbb{T}$.

THEOREM 4.3 (Characterization of operators which commute with translations). *Let* T *be an operator from* L^p *to* L^q *which commutes with translations. Then there exists* $h \in S'(\mathbb{R})$ *such that*

$$T\phi = h * \phi \quad \textit{a.e.} \qquad \forall \phi \in S(\mathbb{R}).$$

This h *is unique.*

PROOF. (Sketch) The proof relies on two facts which we will take for granted here[5].

[5] a detailed proof can be found in [**1**, Theorem ??]

(1) The assumptions made imply that T commutes with (distributional) derivatives, i.e. $T\partial^\alpha f = \partial^\alpha T f$ for $f \in S(\mathbb{R})$ where it is implicit that $\partial^\alpha T f \in L^q$ for all α.
(2) (Sobolev embedding): If $g \in L^q(\mathbb{R})$ is such that the distributional derivatives $\partial^\alpha h$ are in L^q for all $\alpha \in \mathbb{N}_0$, then g is continuous[6] and

$$|g(0)| \leq C_q \leq \sum_{|\alpha|\leq n+1} \|\partial^\alpha g\|_{L^q}$$

Approach: set $g_\phi := T(\phi)$ for $\phi \in S(\mathbb{R})$ where g can be chosen to be continuous by (2), and define the functional u by

$$\langle u, \phi\rangle = g_\phi(0)$$

It is easy to see that u is well-defined. Next we show that $u \in S'(\mathbb{R})$. This follows by the inequality in (2), (1), the boundedness of T and the fact that $S(\mathbb{R})$ is continuously embedded in L^p. Now let $h = Ru$. To show that $T = h * \cdot$ on $S(\mathbb{R})$, we first show that for all $s \in \mathbb{R}$,

$$\langle u, \tau_s\phi\rangle = g_\phi(s).$$

Let $g[s] = T(\tau_s\phi)$ be continuous (via (2)). Then

$$g[s](t) = T(\tau_s\phi)(t) = \tau_s T(\phi)(t) = T(\phi)(t+s) = g(t+s) = \tau_s g(t)$$

holds for a.e. $t \in \mathbb{R}$, but since the left and the right-hand side are continuous in t, even for all $t \in \mathbb{R}$. Hence, $g[s](0) = \tau_s g(0)$ which shows $\langle u, \tau_s\phi\rangle = g_\phi(s)$. We conclude the proof by

$$(h * \phi)(t) = \langle Ru, \tau_{-t}R\phi\rangle = \langle u, \tau_t\phi\rangle = g_\phi(t) = T(\phi)(t)$$

which holds for a.e. $t \in \mathbb{R}$. □

For special choices of the parameters p, q, we can expect more specific information on the tempered distribution given by Theorem 4.3. We state these results without proof and refer the interested reader to [**1**].

THEOREM 4.4. *Let $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ commute with translations. Then there exists $h \in L^\infty(\mathbb{R})$ such that $T_h = T$ (more precisely, T_h is the unique bounded extension of the operator $\phi \mapsto h * \phi$ defined on $S(\mathbb{R})$).*

THEOREM 4.5. *Let $T : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ commute with translations. Then there exists $h \in$ M$(\mathbb{R})$ such that $T_h = T$ (more precisely, T_h is the unique bounded extension of the operator $\phi \mapsto h * \phi$ defined on $S(\mathbb{R})$).*

5. The Hilbert transform

DEFINITION 5.1. The convolution operator $T_h = h * \cdot$ with h given by

$$\langle h, \phi\rangle = \lim_{\epsilon\to 0^+} \int_{|x|\geq\epsilon} \frac{\phi(x)}{x}\, dx$$

is called *Hilbert transform*, see also Exercise 7.1 for a proof that $h \in S'(\mathbb{R})$.

THEOREM 5.2. *The Hilbert transform is bounded from $L^p(\mathbb{R})$ to $L^p(\mathbb{R})$ for $p \in (1, \infty)$.*

PROOF. The proof is omitted, but can be done in a similar spirit as the proof of Theorem 5.11. □

[6] up to identification in L^q

Bibliography

[1] L. Grafakos, *Classical Fourier Analysis*, volume 243 of *Graduate Texts in Mathematics*, 3rd edition, Springer, 2014.

[2] Y. Katznelson *An Introduction to Harmonic Analysis*, volume 243 of *Graduate Texts in Mathematics*, 3rd edition, Cambridge University Press, 1968.

[3] T. Tao, *Math 247A: Fourier Analysis*, lecture notes available at `http://www.math.ucla.edu/~tao/247a.1.06f/` (last access July 2018)

[4] T. Tao, *An epsilon of room, I: real analysis*, volume 117 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, 2010.