MA342A 2017–18

Chapter 2: Convergence results for Fourier series

Contents

2.1 Differentiable functions on the circle

Here we consider Fourier series of functions $F \in C(\mathbb{T})$ but often we will parametrize the circle by $\zeta = e^{2\pi i x}$ and consider the function $f(x) = F(e^{2\pi i x})$ with $x \in \mathbb{R}$ (which is periodic and has 1 as a period).

2.1.1 Definition (Partial sums)**.** If $f \in \mathcal{L}^1[0,1]$ and $N \in \mathbb{N} \cup \{0\}$, then we call the function $S_N f \colon \mathbb{R} \to \mathbb{C}$ given by

$$S_N f(x) = \sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i n x} \qquad (x \in \mathbb{R})$$

the N^{th} (symmetric) *partial sum* of the Fourier series of f.

If $F \in C(\mathbb{T})$ we define $S_N F \colon \mathbb{T} \to \mathbb{C}$ similarly as

$$S_N F(\zeta) = \sum_{n=-N}^{N} \hat{F}(n) \zeta^n \qquad (\zeta \in \mathbb{T}).$$

We would like to show (under suitable hypotheses, if necessary) that the partial sums of the Fourier series converge to the function. One could also argue with the terminology as $S_0 f$ is defined (it is the constant function $\hat{f}(0)$) and so perhaps we should call that the first instead of the zeroth partial sum.

2.1.2 Proposition (Version of Bessel's inequality). *If $F \in C(\mathbb{T})$, then*

$$\sum_{n=-\infty}^{\infty} |\hat{F}(n)|^2 \leq \int_0^1 |F(e^{2\pi i x})|^2 \, dx$$

Proof. For ease of notation, let $f(x) = F(e^{2\pi i x})$. We know $\hat{F}(n) = \hat{f}(n)$. Fix $N \in \mathbb{N}$ and let

$$h(x) = f(x) - \sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i n x}$$

We know that $\int_0^1 |h(x)|^2 \, dx \geq 0$. We expand that using

$$|h(x)|^2 = h(x)\overline{h(x)} = \left(f(x) - \sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i n x} \right) \overline{\left(f(x) - \sum_{m=-N}^{N} \hat{f}(m) e^{2\pi i m x} \right)}.$$

So

$$\begin{aligned} |h(x)|^2 &= f(x)\overline{f(x)} - \sum_{m=-N}^{N} f(x)\overline{\hat{f}(m)} e^{-2\pi i m x} - \sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i n x} \overline{f(x)} \\ &\quad + \sum_{n,m=-N}^{N} \hat{f}(n)\overline{f(m)} e^{2\pi i n x} e^{-2\pi i m x} \end{aligned}$$

Integrate to get

$$\begin{aligned} 0 \leq \int_0^1 |h(x)|^2 \, dx &= \int_0^1 |f(x)|^2 \, dx - \sum_{m=-N}^{N} \hat{f}(m)\overline{\hat{f}(m)} \\ &\quad - \sum_{n=-N}^{N} \hat{f}(n) \overline{\int_0^1 f(x) e^{-2\pi i n x} \, dx} \\ &\quad + \sum_{n,m=-N}^{N} \hat{f}(n)\overline{f(m)} \int_0^1 e^{2\pi i n x} e^{-2\pi i m x} \, dx \\ &= \int_0^1 |f(x)|^2 \, dx - \sum_{m=-N}^{N} |\hat{f}(m)|^2 - \sum_{n=-N}^{N} |\hat{f}(n)|^2 \\ &\quad + \sum_{n=-N}^{N} |\hat{f}(n)|^2 \end{aligned}$$

(using orthonormality of the complex exponentials $e^{2\pi i n x}$). Hence we have

$$\sum_{n=-N}^{N} |\hat{f}(n)|^2 \leq \int_0^1 |f(x)|^2 \, dx.$$

Let $N \to \infty$ to complete the proof. $\square$

2.1.3 Corollary. *If $F \in C(\mathbb{T})$, then $\lim_{|n|\to\infty} \hat{F}(n) = 0$.*

2.1.4 Proposition. *If $f\colon \mathbb{R} \to \mathbb{C}$ has $f(x+1) = f(x)$ and f has a continuous derivative $h = f'$ (on $\mathbb{R}$), then*

$$\hat{f}(n) = \frac{1}{2\pi i n}\hat{h}(n)$$

for $n \in \mathbb{Z} \setminus \{0\}$.

Proof. By integration by parts

$$\hat{f}(n) = \int_0^1 f(x)\, d\left(\frac{e^{-2\pi i n x}}{-2\pi i n}\right) = \left[f(x)\frac{e^{-2\pi i n x}}{-2\pi i n}\right]_0^1 - \int_0^1 \frac{e^{-2\pi i n x}}{-2\pi i n} f'(x)\, dx = 0 + \frac{1}{2\pi i n}\hat{h}(n)$$

□

2.1.5 Corollary. *If $f\colon \mathbb{R} \to \mathbb{C}$ has $f(x+1) = f(x)$ and f has a continuous derivative $h = f'$ (on $\mathbb{R}$), then*

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)| < \infty$$

Proof. Let $h = f'$. Then h is periodic with 1 as a period and there is a corresponding $H \in C(\mathbb{T})$ with $H(e^{2\pi i x}) = h(x)$. Applying Proposition 2.1.2, we conclude that $\sum_{n=-\infty}^{\infty} |\hat{h}(n)|^2 < \infty$.

From Proposition 2.1.4, we have $\hat{f}(n) = \hat{h}(n)/(2\pi i n)$ for $n \neq 0$. It is elementary then (since $2ab \leq a^2 + b^2$ for $a, b \geq 0$) that

$$|\hat{f}(n)| = |\hat{h}(n)|\frac{1}{2\pi n} \leq \frac{1}{2}\left(|\hat{h}(n)|^2 + \frac{1}{4\pi^2 n^2}\right)$$

(for $n \neq 0$) and so $\sum_{n\in\mathbb{Z}} |\hat{f}(n)| < \infty$ (using $\sum_{n=1}^{\infty} 1/n^2 < \infty$). □

While this result seems very close to showing that the Fourier series of continuously differentiable (periodic) functions converge to the function, we need something more to show that result. We will now introduce some terminology and results that will also be useful for more general classes of functions.

2.2 More formalities

2.2.1 Definition (See Definition 1.1.4.5)**.** A *seminorm* on a vector space V over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ is a function from V to the real numbers with the following 3 properties, where we use $\|\cdot\|$ for the function:

(SN1) $\|v\| \geq 0$ for $v \in V$

(SN2) (Triangle inequality) $\|v + w\| \leq \|v\| + \|w\|$ for $v, w \in V$

(SN3) (scaling property) $\|\lambda v\| = |\lambda|\|v\|$ for $\lambda \in \mathbb{K}$, $v \in V$

A seminorm is called a *norm* if it has the additional property that

(N4) $v \in V, \|v\| = 0 \Rightarrow v = 0$

2.2.2 Remark. It is perhaps unusual not to reserve the notation $\|\cdot\|$ for norms (and to use something distinctive like $p(\cdot)$ for seminorms).

2.2.3 Examples. (i) We saw the definitions $\mathcal{L}^1([0,1])$, $\mathcal{L}^1(\mathbb{T})$ and $\mathcal{L}^1(\mathbb{R})$ in Chapter 1. In each case we defined a 'magnitude' $\|f\|_1$ and it is a seminorm in each case. We stated that explicitly for $\mathcal{L}^1([0,1])$ in Lemma 1.1.5.2.

(ii) On $C(\mathbb{T})$ we can define a norm known as the uniform norm by

$$\|F\|_\infty = \sup_{\zeta \in \mathbb{T}} |F(\zeta)| \qquad (F \in C(\mathbb{T}))$$

and on $CP[0,1]$ we have the analogous uniform norm

$$\|f\|_\infty = \sup_{x \in [0,1]} |f(x)|.$$

In fact we should recognize $\|f\|_\infty$ for $f \in C[0,1]$ from MA2223.

2.2.4 Proposition. *If $\|\cdot\|$ is a norm on a vector space V, then we have a (naturally) associated metric on V given by*

$$d(v,w) = \|v - w\|$$

Proof. We omit the details. They are straightforward and should be familiar from MA2223. □

2.2.5 Proposition. *If $\|\cdot\|$ is a seminorm on a vector space V, then $V_0 = \{v \in V : \|v\| = 0\}$ is a vector subspace and we may define a norm on the quotient vector space V/V_0 by*

$$\|v + V_0\| = \|v\|$$

We also then have a (naturally) associated metric on V/V_0 given by

$$d(v + V_0, w + V_0) = \|v - w\|$$

Proof. If $v, w \in V_0$ and $\lambda \in \mathbb{K}$, then

$$\|v + \lambda w\| \leq \|v\| + \|\lambda w\| = \|v\| + |\lambda|\|w\| = 0$$

and so $v + \lambda w \in V_0$. Also $0 \in V_0$ and so V_0 is a vector subspace.

To show that $\|v + V_0\|$ is well defined, suppose $v + V_0 = w + V_0$ for $v, w \in V$. Then $v - w \in V_0$, so $\|v - w\| = 0$ and

$$\|v\| = \|w + (v - w)\| \leq \|w\| + \|v - w\| = \|w\|.$$

Similarly $\|w\| \leq \|v\|$ and so $\|v\| = \|w\|$.

It is straightforward to check that the seminorm properties hold for the function $v + V_0 \mapsto \|v + V_0\|$ and it is a norm since $\|v + V_0\| = 0 \Rightarrow \|v\| = 0 \Rightarrow v \in V_0 \Rightarrow v + V_0 = V_0$ is the zero element of V/V_0. □

2.2.6 Definition. $L^1(\mathbb{T})$ is the quotient vector space $\mathcal{L}^1(\mathbb{T})/\{F \in \mathcal{L}^1(\mathbb{T}) : \|F\|_1 = 0\}$ and we also use $\|\cdot\|_1$ for the norm on $L^1(\mathbb{T})$.

Similarly for $L^1([0,1])$ and $L^1(\mathbb{R})$.

2.2.7 Remark. So the elements of any of these L^1 spaces are not exactly measurable functions, but rather cosets or equivalence classes of measurable functions. Two functions in $\mathcal{L}^1$ represent the same coset (or class) in L^1 when they are equal almost everywhere (because $\|f\|_1 = 0 \iff f$ is almost everywhere zero on the domain in question).

There is some possibility of a misunderstanding because we will tend not to distinguish carefully between a single measurable $f \in \mathcal{L}^1$ and its almost everywhere equivalence class $f \in L^1$, but this is a standard way to proceed. If you ever find yourself tempted to prove something about one value of f, like $f(1/3)$ for $f \in L^1[0,1]$, then it cannot be right because you can change the value of $f(x)$ at $x = 1/3$ without changing the equivalence class. You can only hope to prove things that rely on the values of $f(x)$ for x in a set of positive measure.

What we are aiming for is to prove convergence of Fourier series. That means to prove $\lim_{N\to\infty} S_N f = f$ (and maybe variations of this) for suitable interpretations of what the limit might mean (and we will need some hypotheses on f). We might like it to mean convergence in the uniform norm, that is $\lim_{N\to\infty} \|S_N f - f\|_\infty = 0$ when $f \in C(\mathbb{T})$ or $f \in CP[0,1]$. When $f \in L^1(\mathbb{T})$, we might hope for $\lim_{N\to\infty} \|S_N f - f\|_1 = 0$.

Other norms will show up later and in §2.4 we introduce a class of norms (or seminorms) that we can manage.

2.3 Translations and convolutions

2.3.1 Definition (Translations)**.** For $f \colon \mathbb{R} \to \mathbb{C}$ and $y \in \mathbb{R}$ we define the *translate of f by y* to be the function $f_y \colon \mathbb{R} \to \mathbb{C}$ given by $f_y(x) = f(x - y)$.

For $f \in CP[0,1]$ we use the same terminology and the same definition of f_y (for which we need to consider the periodic extension of f to $\mathbb{R}$ — we then restrict f_y back to $[0,1]$ to get $f_y \in CP[0,1]$.

For $F \in C(\mathbb{T})$ and $\eta = e^{2\pi i y} \in \mathbb{T}$ we define $F_\eta \in C(\mathbb{T})$ by $F_\eta(\zeta) = F(\zeta\eta^{-1})$. (Note that for $f(x) = F(e^{2\pi i x})$ we have $f_y(x) = F(e^{2\pi i(x-y)}) = F_\eta(e^{2\pi i x})$, so that this is essentially the same notion of translation again).

For $f \in L^1(\mathbb{R})$ we define f_y by $f_y(x) = f(x - y)$ and then $f_y \in L^1(\mathbb{R})$.

For $f \in L^1[0,1]$ we define f_y by $f_y(x) = f(x - y \pmod 1)$ where we need to reduce $x - y$ to $x - y - n \in [0,1]$ with $n \in \mathbb{Z}$. For $F \in L^1(\mathbb{T})$ we define $F_\eta \in L^1(\mathbb{T})$ by $F_\eta(\zeta) = F(\zeta\eta^{-1})$ as above.

2.3.2 Definition (Convolution)**.** If $f, h \in C_c(\mathbb{R})$ we define their *convolution* $f * h \colon \mathbb{R} \to \mathbb{C}$ by

$$(f * h)(x) = \int_{y\in\mathbb{R}} f(y)h(x-y)\,d\mu(y) = \int_{-\infty}^{\infty} f(y)h(x-y)\,dy = \int_{\mathbb{R}} f(y)h_y(x)\,d\mu(y)$$

(a finite integral because of the compact support of f).

If $f, h \in CP[0,1]$ we define their convolution $f * h \colon [0,1] \to \mathbb{C}$ by

$$(f * h)(x) = \int_{y \in [0,1]} f(y) h_y(x)\, d\mu(y) = \int_0^1 f(y) h(x - y \text{ (mod 1) })\, dy = \int_{[0,1]} f(y) h_y(x)\, d\mu(y)$$

(with the same $f * h$ notation even though it has a different meaning — if f and h are extended periodically to $\mathbb{R}$, they would not have compact support unless they are zero and the previous definition could not be used).

If $F, H \in C(\mathbb{T})$ we define their convolution $F * H \colon \mathbb{T} \to \mathbb{C}$ by

$$(F * H)(\zeta) = (F * H)(e^{2\pi i x}) = \int_0^1 F(e^{2\pi i y}) H(e^{2\pi i (x-y)})\, dy = \int_{\mathbb{T}} F(\eta) H_\eta(\zeta)\, d\lambda(\eta)$$

(where λ is normalized arc length measure on $\mathbb{T}$, $\zeta = e^{2\pi i x} \in \mathbb{T}$).

2.3.3 Remark. The definition of convolution on $CP[0,1]$ is really convolution on $C(\mathbb{T})$, when we regard $f \in CP[0,1]$ as the 'same as' $F \in C(\mathbb{T})$ with $F(e^{2\pi i x}) = f(x)$.

2.3.4 Theorem. *(a) If $F, H \in L^1(\mathbb{T})$, then $\eta \mapsto F(\eta) H_\eta(\zeta)$ is in $L^1(\mathbb{T})$ for almost every $\zeta \in \mathbb{T}$ and defining $F * H \colon \mathbb{T} \to \mathbb{C}$ (almost everywhere) by $(F * H)(\zeta) = \int_{\mathbb{T}} F(\eta) H_\eta(\zeta)\, d\lambda(\eta)$ gives $F * H \in L^1(\mathbb{T})$.*

Moreover $\|F * H\|_1 \leq \|F\|_1 \|H\|_1$.

(b) If $f, h \in L^1[0,1]$, then $y \mapsto f(y) h_y(x)$ is in $L^1[0,1]$ for almost every $x \in [0,1]$ and defining $f * h \colon [0,1] \to \mathbb{C}$ (almost everywhere) by $(f * h)(x) = \int_{[0,1]} f(y) h_y(x)\, d\mu(y)$ gives $f * h \in L^1[0,1]$.

Moreover $\|f * h\|_1 \leq \|f\|_1 \|h\|_1$.

(c) If $f, h \in L^1(\mathbb{R})$, then $y \mapsto f(y) h_y(x)$ is in $L^1(\mathbb{R})$ for almost every $x \in \mathbb{R}$ and defining $f * h \colon \mathbb{R} \to \mathbb{C}$ (almost everywhere) by $(f * h)(x) = \int_{\mathbb{R}} f(y) h_y(x)\, d\mu(y)$ gives $f * h \in L^1(\mathbb{R})$.

Moreover $\|f * h\|_1 \leq \|f\|_1 \|h\|_1$.

Proof. For this we need Fubini's theorem (which was not in MA2224 but is stated at Theorem A.1.13 in Appendix A.1).

We apply it to $\phi(x, y) = f(y) h(x - y)$ where $f(x) = F(e^{2\pi i x})$ and $h(x) = H(e^{2\pi i x})$. One can check that ϕ is measurable on $[0,1] \times [0,1]$.

$$\int_{y \in [0,1]} \int_{x \in [0,1]} |\phi(x, y)|\, d\mu(x)\, d\mu(y) = \int_{y \in [0,1]} |f(y)| \left(\int_{x \in [0,1]} |h(x - y)|\, d\mu(x) \right) d\mu(y)$$

The inner integral is $\|h\|_1 = \|H\|_1$ and then the result is $\|F\|_1 \|G\|_1$. By Fubini, ϕ is integrable on $[0,1] \times [0,1]$ and

$$\|F\|_1 \|G\|_1 = \int_{y \in [0,1]} \int_{x \in [0,1]} |\phi(x, y)|\, d\mu(x)\, d\mu(y) = \int_{x \in [0,1]} \int_{y \in [0,1]} |\phi(x, y)|\, d\mu(y)\, d\mu(x)$$

Hence $\int_{y\in[0,1]} |\phi(x,y)|\,d\mu(y) < \infty$ for almost all x and

$$\begin{aligned}\|F\|_1\|G\|_1 &= \int_{x\in[0,1]}\int_{y\in[0,1]} |\phi(x,y)|\,d\mu(y)\,d\mu(x)\\ &\geq \int_{x\in[0,1]}\left|\int_{y\in[0,1]} \phi(x,y)\,d\mu(y)\right|\,d\mu(x)\\ &= \int_{x\in[0,1]} |(f*h)(x)|\,d\mu(x)\\ &= \|f*h\|_1 = \|F*H\|_1\end{aligned}$$

This shows (a) and (b) is really the same statement as (a).

For (c) the proof is the same, using $\mathbb{R}$ instead of $[0,1]$ and omitting reference to F and H. □

2.3.5 Lemma. *The convolution product in $L^1(\mathbb{T})$ makes $L^1(\mathbb{T})$ an* algebra*, or a 'linear algebra' to distinguish it from an algebra (Boolean algebra) of subsets, that is $L^1(\mathbb{T})$ is a vector space with a product $*$ that satisfies*

(LinAlg1) $F*H \in L^1(\mathbb{T})$ *for* $F,H \in L^1(\mathbb{T})$

(LinAlg2) $F*(H+\lambda K) = F*H + \lambda F*K$ *and*
$(F+\lambda K)*H = F*H + \lambda K*H$
hold for $F,H,K \in L^1(\mathbb{T})$ *and* $\lambda \in \mathbb{C}$

(LinAlg3) (associativity of multiplication)
$(F*H)*K = F*(H*K)$ *for* $F,H,K \in L^1(\mathbb{T})$

It is in fact a commutative *algebra, that is* $F*H = H*F$ *for* $F,H \in L^1(\mathbb{T})$.

The property $\|F*H\|_1 \leq \|F\|_1\|H\|_1$ *(for $H,K \in L^1(\mathbb{T})$) is called submultiplicativity of the norm $\|\cdot\|_1$, and an algebra with a submultiplicative norm is called a* normed algebra*.*

Proof. We have (LinAlg1) from Theorem 2.3.4 and (LinAlg2) follows by linearity of integrals (valid even for integrals of $\mathbb{C}$-valued functions). Showing (LinAlg3) requires Fubini's theorem to exchange the order of integration.

Showing $F*H = H*F$ requires a small change of variables argument. □

2.3.6 Remark. In fact $L^1(\mathbb{T})$ is complete[1] in the distance arising from the norm $\|\cdot\|_1$ (which we show in Appendix A.2) and a complete normed space is called a *Banach space*. A complete normed algebra is called a *Banach algebra* and $L^1(\mathbb{T})$ is then a commutative Banach algebra.

The same is true about $L^1([0,1])$ and $L^1(\mathbb{R})$ with $*$ and $\|\cdot\|_1$ (with the same proofs).

Our aim now is results along the lines of recovering $f \in L^1(\mathbb{T})$ from its Fourier series, but there are results that are more satisfactory for smaller spaces with certain properties as defined now.

[1]Recall that by definition a metric space is complete if every Cauchy sequence in the space has a limit in the space.

2.4 Homogeneity of spaces of functions

2.4.1 Definition. If $(B, \|\cdot\|_B)$ is a Banach space of functions (or of almost everywhere equivalence classes of functions) on $\mathbb{T}$ we call B a *homogeneous Banach space on* $\mathbb{T}$ if it satisfies

(HB1) $B \subseteq L^1(\mathbb{T})$ and $\|F\|_1 \leq \|F\|_B$ for $F \in B$

(HB2) (translation invariance) if $F \in B$ and $\eta \in \mathbb{T}$, then $F_\eta \in B$ and $\|F_\eta\|_B = \|F\|_B$ (where $F_\eta(\zeta) = F(\zeta\eta^{-1})$ as in Definition 2.3.1)

(HB3) (continuity property of translates) if $F \in B$ and $\eta_0 \in \mathbb{T}$, then

$$\lim_{\eta \to \eta_0} \|F_\eta - F_{\eta_0}\| = 0$$

2.4.2 Examples.

(i) $B = C(\mathbb{T})$ with $\|F\|_B = \|F\|_\infty$ is a homogeneous Banach space in this sense.

Proof. It is very easy that $\|F\|_1 \leq \|F\|_\infty$ for $F \in C(\mathbb{T})$ because

$$\|F\|_1 = \int_{\mathbb{T}} |F(\zeta)|\, d\lambda \leq \int_{\mathbb{T}} \|F\|_\infty \, d\lambda = \|F\|_\infty$$

It is also clear that $F_\eta \in C(\mathbb{T})$ for $\eta \in \mathbb{T}$ and the last property (HB3) follows from uniform continuity of continuous functions on the compact circle $\mathbb{T}$. $\square$

(ii) We define $C^1(\mathbb{T})$ to be the continuously differentiable functions, where the derivative F' is defined as

$$F'(e^{2\pi i x}) = \frac{1}{2\pi i}\frac{d}{dx}F(e^{2\pi i x}).$$

So $C^1(\mathbb{T}) = \{F \in C(\mathbb{T}) : \exists F' \text{ and } F' \in C(\mathbb{T})\}$. If we norm the space with

$$\|F\|_{\infty,1} = \|F\|_\infty + \|F'\|_\infty$$

then $(C^1(\mathbb{T}), \|\cdot\|_{\infty,1})$ is also a homogeneous Banach space in the above sense.

Proof. The proof is not so much more complicated that the proof for $(C(\mathbb{T}), \|\cdot\|_\infty)$. $\square$

(iii) Our other main examples with be $L^1(\mathbb{T})$ itself and $L^2(\mathbb{T})$ (defined below and in §A.3).

However, to prove that $L^1(\mathbb{T})$ satisfies (HB3) requires knowing that continuous functions are dense in $L^1(\mathbb{T})$, and also knowing that the space is complete. See Theorem A.2.4 and Corollary A.2.7 in the appendices.

For $L^1(\mathbb{T})$, property (HB1) is obvious and (HB2) follows by a change of variables in the integral defining the norm.

If $F \in C(\mathbb{T})$, then $\|F\|_1 \leq \|F\|_\infty$ by the triangle inequality for integrals. It follows that for $\eta, \eta_0 \in \mathbb{T}$

$$\lim_{\eta \to \eta_0} \|F_\eta - F_{\eta_0}\|_1 \leq \limsup_{\eta \to \eta_0} \|F_\eta - F_{\eta_0}\|_\infty = 0$$

(when $F \in C(\mathbb{T})$). For general $F \in L^1(\mathbb{T})$ choose a sequence $F_n \in C(\mathbb{T})$ with $\lim_{n\to\infty} \|F - F_n\|_1 = 0$. Given $\varepsilon > 0$ choose n so large that $\|F - F_n\|_1 < \varepsilon/3$. If η_0 is fixed then, for $|\eta - \eta_0|$ small enough, $\|(F_n)_\eta - (F_n)_{\eta_0}\|_\infty < \varepsilon/3$. So

$$\begin{aligned}
\|F_\eta - F_{\eta_0}\|_1 &\leq \|F_\eta - (F_n)_\eta\|_1 + \|(F_n)_\eta - (F_n)_{\eta_0}\|_1 + \|(F_n)_{\eta_0} - F_{\eta_0}\|_1 \\
&= \|F - F_n\|_1 + \|(F_n)_\eta - (F_n)_{\eta_0}\|_1 + \|F_n - F\|_1 \\
&< \frac{\varepsilon}{3} + \|(F_n)_\eta - (F_n)_{\eta_0}\|_\infty + \frac{\varepsilon}{3} < \varepsilon
\end{aligned}$$

for $|\eta - \eta_0|$ small. This shows (HB3) for $L^1(\mathbb{T})$.

The definition of $L^2(\mathbb{T})$ is as the space of measurable $F \colon \mathbb{T} \to \mathbb{C}$ such that

$$\int_{\mathbb{T}} |F(\zeta)|^2 \, d\lambda(\zeta) < \infty$$

and we define the 'norm' on $L^2(\mathbb{T})$ by

$$\|F\|_2 = \left(\int_{\mathbb{T}} |F(\zeta)|^2 \, d\lambda(\zeta) \right)^{1/2}$$

There are some of the same issues here as with $L^1(\mathbb{T})$. Actually we should say that the definition given is of what should be denoted $\mathcal{L}^2(\mathbb{T})$ (by analogy with $\mathcal{L}^1(\mathbb{T})$) and $L^2(\mathbb{T})$ is the quotient of $\mathcal{L}^2(\mathbb{T})$ by $\{F \in \mathcal{L}^2(\mathbb{T}) : \|F\|_2 = 0\}$, or the almost everywhere equivalence classes of measurable functions $F \in \mathcal{L}^2(\mathbb{T})$.

While it was very easy to check that $\|\cdot\|_1$ is a seminorm, to show that for $\|\cdot\|_2$ requires a little effort (see §A.3).

For $L^2(\mathbb{T})$, to show property (HB1), observe that it follows from the Cauchy-Schwarz inequality that $\|F\|_1 = \langle |F|, 1 \rangle \leq \|F\|_2 \|1\|_2 = \|F\|_2$ holds for $F \in L^2(\mathbb{T})$. (Here 1 means the constant function.) (HB2) is easy to verify by a change of variables in the integral defining the norm.

Property (HB3) for $L^2(\mathbb{T})$ follows from the density of $C(\mathbb{T})$ in and $L^2(\mathbb{T})$ (see Corollary A.3.7) in the same way as for $L^1(\mathbb{T})$.

2.4.3 Definition. If B is a Banach space and $\phi \colon [0, 1] \to B$ is a continuous function (vector-valued or B-valued) then we define

$$\int_0^1 \phi(x) \, dx$$

using the Riemann integral approach (which can be adapted to the vector-valued case).

So
$$\int_0^1 \phi(x)\,dx = \lim_{n\to\infty} \sum_{j=1}^n \phi((j-1)/n)\frac{1}{n}$$
using the partition $0 < 1/n < 2/n < \cdots < (n-1)/n < 1$ of $[0,1]$. We could use other Riemann sums with mesh sizes tending to 0 and get the same result. (The limit exists as a consequence of uniform continuity of ϕ and completeness of B, in the same way as for $\mathbb{R}$-valued Riemann integrals. Proofs of this and simple properties of Banach space valued integrals will generally involve considering more general kinds of Riemann sums than the one above.)

2.4.4 Lemma (Vector-valued triangle inequality for Riemann integrals)**.** *If $\phi\colon [0,1] \to B$ is a continuous function with values in a Banach space B, then*
$$\left\| \int_0^1 \phi(x)\,dx \right\| \le \int_0^1 \|\phi(x)\|\,dx$$

Proof. This follows from
$$\left\| \sum_{j=1}^n \phi((j-1)/n)\frac{1}{n} \right\| \le \sum_{j=1}^n \|\phi((j-1)/n)\|\frac{1}{n}$$
which is true by the ordinary finite triangle inequality. Just let $n\to\infty$. □

2.4.5 Lemma. *If $\phi\colon [0,1] \to B$ is a continuous function with values in a Banach space B, and $T\colon B \to C$ is a continuous linear transformation with values in a Banach space C, then*
$$T\left(\int_0^1 \phi(x)\,dx\right) = \int_0^1 T(\phi(x))\,dx$$

Proof. Since T is linear
$$T\left(\sum_{j=1}^n \phi((j-1)/n)\frac{1}{n}\right) = \sum_{j=1}^n T(\phi((j-1)/n))\frac{1}{n}$$
we can take limits of both sides as $n\to\infty$ and use continuity of T to prove the result. □

2.4.6 Proposition. *Suppose that $\phi, F \in C(\mathbb{T})$. Then the convolution $\phi * F$ may be written as the $C(\mathbb{T})$-valued integral*
$$\phi * F = \int_0^1 \phi(e^{2\pi i x}) F_{e^{2\pi i x}}\,dx \in C(\mathbb{T}).$$

Proof. Because of (HB3) for the example $B = C(\mathbb{T})$ we know that for $F \in C(\mathbb{T})$, $\eta \mapsto F_\eta$ is a $C(\mathbb{T})$-valued continuous function on $\mathbb{T}$. Hence $x \mapsto F_{e^{2\pi i x}}$ is continuous on $[0,1]$ (and periodic), as the composition of continuous functions. So then also $x \mapsto \phi(e^{2\pi i x})F_{e^{2\pi i x}}$ is continuous (as the product of continuous functions, even though one is a vector-valued function). Hence, the $C(\mathbb{T})$-valued integral $\int_0^1 \phi(e^{2\pi i x})F_{e^{2\pi i x}}\,dx$ makes sense in $C(\mathbb{T})$.

If we fix $\zeta \in \mathbb{T}$, then there is an associated point evaluation linear transformation

$$T_\zeta \colon C(\mathbb{T}) \to \mathbb{C}$$

given by $T_\zeta(H) = H(\zeta)$ for $H \in C(\mathbb{T})$. This is continuous because if $\varepsilon > 0$ and $H, K \in C(\mathbb{T})$ have $\|H - K\|_\infty < \varepsilon$, then $|H(\zeta) - K(\zeta)| \leq \|H - K\|_\infty < \varepsilon$. So in the definition of continuity we can take $\delta = \varepsilon$.

If we apply Lemma 2.4.5 to the linear transformation T_ζ, we find

$$\begin{aligned} T_\zeta\left(\int_0^1 \phi(e^{2\pi i x}) F_{e^{2\pi i x}}\,dx\right) &= \int_0^1 T_\zeta(\phi(e^{2\pi i x}) F_{e^{2\pi i x}})\,dx = \int_0^1 \phi(e^{2\pi i x}) F_{e^{2\pi i x}}(\zeta)\,dx \\ &= (\phi * F)(\zeta) \end{aligned}$$

holds for each $\zeta \in \mathbb{T}$. That says

$$\int_0^1 \phi(e^{2\pi i x}) F_{e^{2\pi i x}}\,dx = \phi * F$$

in $C(\mathbb{T})$. □

2.4.7 Proposition. *Suppose that B is a homogeneous Banach space such that $C(\mathbb{T}) \subseteq B$ is a dense subspace of B and $\|F\|_B \leq \|F\|_\infty$ for $F \in C(\mathbb{T})$.*

Then, for $F \in B$ and $\phi \in C(\mathbb{T})$ continuous, $\phi * F$ is equal almost everywhere to

$$\int_0^1 \phi(e^{2\pi i x}) F_{e^{2\pi i x}}\,dx \in B.$$

Proof. Fix $\phi \in C(\mathbb{T})$.

As in the case $B = C(\mathbb{T})$ we can use (HB3) to show that $x \mapsto \phi(e^{2\pi i x}) F_{e^{2\pi i x}}$ is continuous from $[0,1]$ to B and that the B-valued integral

$$\int_0^1 \phi(e^{2\pi i x}) F_{e^{2\pi i x}}\,dx \in B$$

makes sense for $F \in B$.

For $F \in B$ there is a sequence $(F_n)_{n=1}^\infty$ in $C(\mathbb{T})$ such that $\lim_{n\to\infty} F_n = F$, that is $\lim_{n\to\infty} \|F_n - F\|_B = 0$. Then $\lim_{n\to\infty} \|F_n - F\|_1 = 0$ by (HB1).

It follows from Proposition 2.4.6 that

$$\int_0^1 \phi(e^{2\pi i x}) (F_n)_{e^{2\pi i x}}\,dx = \phi * (F_n) \in C(\mathbb{T}) \qquad (n = 1, 2, \ldots).$$

Applying Lemma 2.4.5 to the the continuous (linear) inclusion $C(\mathbb{T}) \subseteq B$ these equations also hold in B. (That is the $C(\mathbb{T})$-valued integral is the same as the B-valued integral for each n.)

Next

$$\begin{aligned}
\left\| \int_0^1 \phi(e^{2\pi ix})(F_n)_{e^{2\pi ix}}\,dx - \int_0^1 \phi(e^{2\pi ix})F_{e^{2\pi ix}}\,dx \right\|_B &= \left\| \int_0^1 \phi(e^{2\pi ix})(F_n - F)_{e^{2\pi ix}}\,dx \right\|_B \\
&\le \int_0^1 \|\phi(e^{2\pi ix})(F_n - F)_{e^{2\pi ix}}\|_B\,dx \\
&\le \int_0^1 \|\phi\|_\infty \|(F_n - F)_{e^{2\pi ix}}\|_B\,dx \\
&= \|\phi\|_\infty \int_0^1 \|F_n - F\|_B\,dx \\
&= \|\phi\|_\infty \|F_n - F\|_B \to 0
\end{aligned}$$

Hence

$$\lim_{n\to\infty} \phi * F_n = \int_0^1 \phi(e^{2\pi ix})F_{e^{2\pi ix}}\,dx$$

holds in B. Applying (HB1), it also holds in $L^1(\mathbb{T})$.

$$\|\phi * (F_n) - \phi * F\|_1 = \|\phi * (F_n - F)\|_1 \le \|\phi\|_1 \|F_n - F\|_1 \le \|\phi\|_1 \|F_n - F\|_B \to 0.$$

This means that in $L^1(\mathbb{T})$ we have two limits, for the sequence $(\phi * F_n)_{n=1}^\infty$, the B-valued integral $\int_0^1 \phi(e^{2\pi ix})F_{e^{2\pi ix}}\,dx$ and the element $\phi * F \in L^1(\mathbb{T})$. They must agree and so we get the result. □

One case of interest to us is the case $B = L^1(\mathbb{T})$.

2.5 Summability kernels

2.5.1 Definition. A sequence $K_n(\zeta) \in C(\mathbb{T})$ $(n = 1, 2, \ldots)$ is called a (positive) *summability kernel* if it satisfies

(SK1) $K_n(\zeta) \ge 0$ $(\forall n \in \mathbb{N}, \zeta \in \mathbb{T})$

(SK2) $\int_{\mathbb{T}} K_n\,d\lambda = 1$ $(\forall n)$

(SK3) for $0 < \delta < 2$

$$\lim_{n\to\infty} \int_{\{\zeta\in\mathbb{T}: |\zeta - 1| \ge \delta\}} K_n(\zeta)\,d\lambda(\zeta) = 0$$

We can also refer to $k_n(x) = K_n(e^{2\pi ix})$ as a summability kernel (on $[0, 1]$).

Notice that (SK3) implies that K_n must be concentrated close to $\zeta = 1$ (to satisfy (SK2)) for n large. For k_n, the bulk of the mass (or the area under the graph) of $k_n(x)$ must be where x is close to the endpoints of $[0, 1]$ (when n is large).

2.5.2 Proposition. *If B is a homogeneous Banach space that contains $C(\mathbb{T})$ as a dense subspace and satisfies $\|F\|_B \leq \|F\|_\infty$ for $F \in C(\mathbb{T})$, and if $(K_n)_{n=1}^\infty$ is a summability kernel, then for $F \in B$,*

$$\lim_{n\to\infty} \|K_n * F - F\|_B = 0$$

Proof. Fix $F \in B$ and $\varepsilon > 0$. Then (by (HB3)) there is $\delta > 0$ such that $\|F - F_\eta\|_B < \varepsilon$ for $|\zeta - 1| < \delta$.

To help our notation, we write $E_\delta = \{x \in [0,1] : |e^{2\pi i x} - 1| < \delta\}$. By Proposition 2.4.7, we can express $K_n * F$ as a B-valued integral

$$\begin{aligned} K_n * F &= \int_0^1 K_n(e^{2\pi i x}) F_{e^{2\pi i x}}\, dx \\ &= \int_{E_\delta} K_n(e^{2\pi i x}) F_{e^{2\pi i x}}\, dx + \int_{[0,1]\setminus E_\delta} K_n(e^{2\pi i x}) F_{e^{2\pi i x}}\, dx \end{aligned} \tag{2.5.1}$$

(Here we split the B-valued integral as a sum where the first integral extends over two intervals at either end of $[0,1]$ and the second over one interval. We have not actually developed B-valued integrals over intervals other than $[0,1]$ or proved that these integrals can be split in this way, but it can be done in the same way as for Riemann integrals of continuous scalar-valued functions.) We will compare that to

$$F = \left(\int_0^1 K_n(e^{2\pi i x})\, dx\right) F = \int_0^1 K_n(e^{2\pi i x}) F\, dx = \int_{E_\delta} K_n(e^{2\pi i x}) F\, dx + \int_{[0,1]\setminus E_\delta} K_n(e^{2\pi i x}) F\, dx$$

Applying Lemma 2.4.4 (or a strictly speaking a variant of it) we can estimate the second integral in (2.5.1)

$$\begin{aligned} \left\| \int_{[0,1]\setminus E_\delta} K_n(e^{2\pi i x}) F_{e^{2\pi i x}}\, dx \right\|_B &\leq \int_{[0,1]\setminus E_\delta} K_n(e^{2\pi i x}) \|F_{e^{2\pi i x}}\|_B\, dx \\ &= \|F\|_B \int_{[0,1]\setminus E_\delta} K_n(e^{2\pi i x})\, dx \end{aligned}$$

and this is small for large n by (SK3). Similarly $\int_{[0,1]\setminus E_\delta} K_n(e^{2\pi i x}) F\, dx$ also has small norm for large n.

On the other hand

$$\begin{aligned} &\left\| \int_{E_\delta} K_n(e^{2\pi i x}) F_{e^{2\pi i x}}\, dx - \int_{E_\delta} K_n(e^{2\pi i x}) F\, dx \right\|_B \\ &= \left\| \int_{E_\delta} K_n(e^{2\pi i x}) (F_{e^{2\pi i x}} - F)\, dx \right\|_B \\ &\leq \int_{E_\delta} K_n(e^{2\pi i x}) \|F_{e^{2\pi i x}} - F\|_B\, dx \\ &\leq \varepsilon \int_{E_\delta} K_n(e^{2\pi i x})\, dx \leq \varepsilon \end{aligned}$$

Hence
$$\left\| \int_{E_\delta} K_n(e^{2\pi ix}) F_{e^{2\pi ix}}\, dx - \left(\int_{E_\delta} K_n(e^{2\pi ix})\, dx \right) F \right\|_B \leq \varepsilon$$
Since the first integral is close to $K_n * F$ and
$$\int_{E_\delta} K_n(e^{2\pi ix})\, dx \approx \int_0^1 K_n(e^{2\pi ix})\, dx = 1$$
for n large, we deduce that $K_n * F$ is close to F for n large. □

2.5.3 Lemma. *Recalling the notation $\chi_n(\zeta) = \zeta^n$ ($n \in \mathbb{Z}$),*
$$\chi_n * F = \hat{F}(n)\chi_n$$
for $F \in L^1(\mathbb{T})$.

Proof.
$$\begin{aligned}(\chi_n * F)(e^{2\pi ix}) &= (F * \chi_n)(e^{2\pi ix}) \\ &= \int_0^1 F(e^{2\pi iy})(e^{2\pi i(x-y)})^n\, dy \\ &= (e^{2\pi ix})^n \int_0^1 F(e^{2\pi iy})(e^{-2\pi iy})^n\, dy \\ &= (e^{2\pi ix})^n \hat{F}(n) = \hat{F}(n)\chi_n(e^{2\pi ix})\end{aligned}$$
□

2.5.4 Definition. The *Dirichlet kernel* is the sequence $D_N(\zeta) = \sum_{n=-N}^{N} \zeta^n$ of functions $D_N \colon \mathbb{T} \to \mathbb{C}$ (for $N = 0, 1, 2, \ldots$).

2.5.5 Proposition. *For $F \in L^1(\mathbb{T})$, the N^{th} partial sum $S_N F$ of the Fourier series for F is given by*
$$S_N F = D_N * F$$

Proof. Use Lemma 2.5.3. □

2.5.6 Lemma.
$$D_N(e^{2\pi ix}) = \frac{\sin((2N+1)\pi x)}{\sin(\pi x)}$$
for $e^{2\pi ix} \neq 1$.

Proof. By the formula $\sum_{j=0}^{2n} ar^j = a(1 - r^{2n+1})/(1-r)$ for a geometric sum (valid for $r \neq 1$)
$$D_N(\zeta) = \zeta^{-N} \frac{1 - \zeta^{2N+1}}{1 - \zeta} = \frac{\zeta^{-N} - \zeta^{N+1}}{1 - \zeta}$$

Multiply above and below by $\zeta^{-1/2}$ to get

$$D_N(\zeta) = \frac{\zeta^{-N-1/2} - \zeta^{N+1/2}}{\zeta^{-1/2} - \zeta^{1/2}}$$

Replacing ζ by $e^{2\pi i x}$, this reduces to $(-2i\sin((2N+1)\pi x))/(-2i\sin(\pi x))$ and so we have the result. $\square$

2.5.7 Remark. The Dirichlet kernel does not provide a summability kernel since it is not positive. Thus we cannot apply Proposition 2.5.2 with D_N. (If we could, we would have a quick proof that Fourier series converge in many different norms $\|\cdot\|_B$. However, not only does this way of attempting to prove it fail but it is not always true in homogeneous Banach spaces, and in particular in $L^1(\mathbb{T})$ and $C(\mathbb{T})$ too.)

2.5.8 Definition. If $(x_n)_{n=1}^\infty$ is a sequence in a normed space E, the *Cesàro averages* are defined as

$$\frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

The sequence of Cesàro averages is called the *Cesàro mean* of the original sequence $(x_n)_{n=1}^\infty$.

If $\sum_{n=1}^\infty x_n$ is a series in a normed space E, the series is said to be *Cesàro summable* to $s \in E$ if

$$\lim_{n\to\infty} \frac{1}{n}(s_1 + s_2 + \cdot + s_n) = s$$

in E, where $s_j = \sum_{k=1}^j x_k = x_1 + x_2 + \cdots + x_j$ is the j^{th} partial sum of the series.

(So this is the limit of the Cesàro averages of the partial sums.)

2.5.9 Proposition. *If $\sum_{n=1}^\infty x_n$ is a convergent series in a normed space E, then $\sum_{n=1}^\infty x_n$ is Cesàro summable to the same sum $s = \lim_{n\to\infty}(x_1 + x_2 + \cdots + x_n)$.*

Proof. Let $s_n = \sum_{j=1}^n x_j$ denote the partial sums, so that $\lim_{n\to\infty} \|s_n - s\| = 0$.

If $\varepsilon > 0$ is fixed, by assumption there is n_0 such that $\|s_n - s\| < \varepsilon/2$ all $n \geq n_0$. It follows that for $n \geq n_0$

$$\left\| \frac{s_{n_0} + s_{n_0+1} + \cdots + s_n}{n - n_0 + 1} - s \right\| = \left\| \sum_{j=n_0}^n \frac{s_j - s}{n - n_0 + 1} \right\| = \sum_{j=n_0}^n \frac{\|s_j - s\|}{n - n_0 + 1} < \frac{\varepsilon}{2}$$

Hence

$$\begin{aligned}
\left\| \frac{\sum_{j=1}^n s_j}{n} - s \right\| &= \left\| \frac{\sum_{j=1}^{n_0-1} s_j - s}{n} + \frac{\sum_{j=n_0}^n s_j - s}{n} \right\| \\
&\leq \frac{1}{n} \left\| \sum_{j=1}^{n_0-1} s_j - s \right\| + \frac{n - n_0 + 1}{n} \left\| \frac{\sum_{j=n_0}^n s_j - s}{n - n_0 + 1} \right\| \\
&\leq \frac{1}{n} \left\| \sum_{j=1}^{n_0-1} s_j - s \right\| + \frac{n - n_0 + 1}{n} \frac{\varepsilon}{2} < \varepsilon
\end{aligned}$$

for n large enough. $\square$

2.5.10 Example. Consider the series $\sum_{n=1}^{\infty} x_n$ in $\mathbb{R}$ where $x_n = (-1)^{n+1}$. The partial sums s_n are 1 for n odd and 0 for n even. It follows that

$$\frac{1}{2n}(s_1 + s_2 + \cdots + s_{2n}) = \frac{1}{2}$$

and

$$\frac{1}{2n-1}(s_1 + s_2 + \cdots + s_{2n-1}) = \frac{2n-2}{2n-1} \to \frac{1}{2}$$

Thus the series is Cesàro summable to $1/2$ (but not summable in the usual sense).

2.5.11 Definition. The *Fejér kernel* is the sequence $K_N(\zeta) = (1/N)\sum_{n=0}^{N-1} D_n(\zeta)$ of functions $K_N\colon \mathbb{T} \to \mathbb{C}$ (for $N = 1, 2, \ldots$).

2.5.12 Proposition. . *For $F \in L^1(\mathbb{T})$, the convolution $K_N * F$ is the N^{th} Cesàro average*

$$K_N * F = \frac{1}{N}(S_0F + S_1F + \cdots + S_{N-1}F)$$

of the partial sums of the Fourier series for F.

Proof. This follows immediately from the definition of K_N and Proposition 2.5.5. $\square$

2.5.13 Proposition.

$$K_N(e^{2\pi i x}) = \frac{1}{N}\left(\frac{\sin(\pi N x)}{\sin(\pi x)}\right)^2 = \frac{1}{N}\frac{1-\cos(2\pi N x)}{1-\cos(2\pi x)}$$

for $0 < x < 1$.

Proof. We have

$$\begin{aligned}
K_N(\zeta) &= \frac{1}{N}\sum_{n=0}^{N-1} D_n(\zeta) \\
&= \frac{1}{N}\sum_{n=0}^{N-1} \frac{\sin((2n+1)\pi x)}{\sin(\pi x)} \\
&= \frac{1}{N\sin(\pi x)} \operatorname{Im} \sum_{n=0}^{N-1} \operatorname{Im} e^{(2n+1)\pi i x} \\
&= \frac{1}{N\sin(\pi x)} \operatorname{Im} e^{\pi i x}\frac{e^{2N\pi i x}-1}{e^{2\pi i x}-1} \\
&= \frac{1}{N\sin(\pi x)} \operatorname{Im} \frac{e^{2N\pi i x}-1}{e^{\pi i x}-e^{-\pi i x}} \\
&= \frac{1}{N\sin(\pi x)} \operatorname{Im} \frac{e^{2N\pi i x}-1}{2i\sin(\pi x)} \\
&= \frac{1}{N\sin^2(\pi x)} \operatorname{Re} \frac{1-e^{2N\pi i x}}{2} = \frac{1-\cos(2N\pi x)}{2N\sin^2(\pi x)}
\end{aligned}$$

The two formulae for K_N follow from the trigonometric identity $\cos(2\theta) = 1 - 2\sin^2\theta$. $\square$

2.5.14 Proposition. *The* Fejér kernel *sequence* $(K_N)_{N=1}^{\infty}$ *is a summability kernel.*

Proof. It is easy to see that $\int_0^1 D_n(e^{2\pi i x})\,dx = \sum_{j=0}^n \int_0^1 e^{2\pi i j x}\,dx = 1$ and then

$$\int_0^1 F_N(e^{2\pi i x})\,dx = \frac{1}{N}\sum_{n=0}^{N-1}\int_0^1 D_n(e^{2\pi i x})\,dx = 1$$

From Proposition 2.5.13, $K_N(e^{2\pi i x}) > 0$ for $0 < x < 1$ (and $K_N(1) = N > 0$).

If $|e^{2\pi i x} - 1| > \delta$, then dividing by $1 = |e^{\pi i x}|$ we get $|e^{\pi i x} - e^{-\pi i x}| > \delta$ or $2|\sin(\pi x)| > \delta$. Hence

$$K_N(e^{2\pi i x}) = \frac{1}{N}\left(\frac{\sin(\pi N x)}{\sin(\pi x)}\right)^2 \leq \frac{1}{N}\frac{\sin^2(\pi N x)}{\delta^2/4} \leq \frac{4}{N\delta^2}$$

for $|e^{2\pi i x} - 1| > \delta$ and

$$\int_{\{x\in[0,1]:|e^{2\pi i x}-1|>\delta} K_N(e^{2\pi i x})\,dx \leq \frac{4}{N\delta^2} \to 0$$

Thus (SK3) holds and the proof is complete. □

2.6 Fejér's theorem

2.6.1 Corollary (Fejér's theorem)**.** *If B is a homogeneous Banach space that contains $C(\mathbb{T})$ as a dense subspace, and satisfies $\|F\|_B \leq \|F\|_\infty$ for $F \in C(\mathbb{T})$, then the Fourier series of each $F \in B$ converges to F in the Cesàro sense*

$$\lim_{N\to\infty}\left\|\frac{S_0F + S_1F + \cdots + S_{N-1}F}{N} - F\right\|_B = 0$$

Proof. This is immediate from Propositions 2.5.2 and 2.5.14. □

2.6.2 Theorem (Continuously differentiable functions have convergent Fourier series)**.** *If $f\colon \mathbb{R} \to \mathbb{C}$ has $f(x+1) = f(x)$ and f has a continuous derivative f', then*

$$\lim_{N\to\infty}\sup_{x\in[0,1]}\left|\left(\sum_{n=-N}^{N}\hat{f}(n)e^{2\pi i n x}\right) - f(x)\right| = 0$$

In other words, the partial sums $S_Nf(x)$ converge to $f(x)$ (as $N \to \infty$) uniformly in x.

Proof. From Corollary 2.1.5 we know that $\sum_{n=-\infty}^{\infty}|\hat{f}(n)| < \infty$. Since absolutely convergent series are convergent, we may now define $f_0\colon \mathbb{R} \to \mathbb{C}$ as

$$f_0(x) = \sum_{n=-\infty}^{\infty}\hat{f}(n)e^{2\pi i n x} = \lim_{N\to\infty} S_Nf(x)$$

and the limit is uniform in x because

$$|f_0(x) - S_N f(x)| \leq \sum_{|n|>N} |\hat{f}(n)| \qquad (x \in [0,1])$$

gives

$$\|f_0 - S_N f\|_\infty = \sup_{x\in[0,1]} |f_0(x) - S_N f(x)| \leq \sum_{|n|>N} |\hat{f}(n)| \to 0 \text{ (as } N \to \infty).$$

Then Proposition 2.5.9 gives

$$\lim_{N\to\infty} \left\| \frac{S_0 f + S_1 f + \cdots + S_{N-1} f}{N} - f_0 \right\|_\infty = 0$$

But Fejér's theorem (Corollary 2.6.1) (applied to $F(e^{2\pi i x}) = f(x)$) implies the same with f_0 replaced by f. Hence $f_0 = f$ and $\lim_{N\to\infty} \|f - S_N f\|_\infty = 0$. $\square$

Fejér's theorem has several other consequences.

2.6.3 Notation. For $F \in L^1(\mathbb{T})$ we use $\sigma_N F$ to denote the Cesàro average

$$\sigma_N F = \frac{S_0 F + S_1 F + \cdots + S_{N-1} F}{N}$$

Then the conclusion of Fejér's theorem can be stated more concisely as $\lim_{N\to\infty} \|\sigma_N F - F\|_B = 0$.

Note that $\sigma_N F(\zeta) = (1/N)\sum_{n=0}^{N-1} S_n F(\zeta) = (1/N)\sum_{n=0}^{N-1}\sum_{j=-n}^{n} \hat{F}(j)\zeta^j$ and this can be rearranged as

$$\sum_{n=-N+1}^{N-1} (1 - |n|/N)\hat{F}(n)\zeta^n$$

2.6.4 Corollary (Wierstrass theorem)**.** *The trigonometric polynomials are dense in* $C(\mathbb{T})$*.*

Proof. By Fejér's theorem (Corollary 2.6.1) applied to $B = C(\mathbb{T})$, for $F \in C(\mathbb{T})$ we have $\lim_{N\to\infty} \|\sigma_N F - F\|_\infty = 0$.

But $\sigma_N F(\zeta)$ is a trigonometric polynomial for each N. $\square$

2.6.5 Corollary (Fourier series determine $L^1(\mathbb{T})$ functions)**.** *An element $F \in L^1(\mathbb{T})$ is uniquely determined by its Fourier series.*

(In other words, if $F, H \in L^1(\mathbb{T})$ have $\hat{F}(n) = \hat{H}(n)$ for each $n \in \mathbb{Z}$, then $F = H$.)

Proof. If $F, H \in L^1(\mathbb{T})$ have $\hat{F}(n) = \hat{H}(n)$ for each $n \in \mathbb{Z}$, the $\sigma_N F = \sigma_N H$ for each N. By Fejér's theorem (Corollary 2.6.1) applied to $B = L^1(\mathbb{T})$, we have

$$F = \lim_{N\to\infty} \sigma_N F = \lim_{N\to\infty} \sigma_N H = H$$

(limits in $(L^1(\mathbb{T}), \|\cdot\|_1)$). $\square$

2.6.6 Corollary (L^2 convergence of Fourier series)**.** *If $F \in L^2(\mathbb{T})$, then*

$$\lim_{N\to\infty} \|S_N F - F\|_2 = 0$$

Proof. In $L^2(\mathbb{T})$ we can write Fourier coefficients as inner products $\hat{F}(n) = \langle F, \chi_n \rangle$ where $\chi_n(\zeta) = \zeta^n$. So

$$S_N F = \sum_{n=-N}^{N} \langle F, \chi_n \rangle \chi_n$$

for $F \in L^2(\mathbb{T})$.

We can also write $\sigma_N F$ as a linear combination $\sigma_N F = \sum_{n=-N+1}^{N-1} (1 - |n|/N)\hat{F}(n)\chi_n$ of the χ_n with $|n| < N$. (We will not need the exact values $(1 - |n|/N)\hat{F}(n)$ of the coefficients in our proof.)

As in the proof of Proposition 2.1.2, it is easy to verify $\langle \chi_n, \chi_m \rangle = \delta_{n,m}$ (that is 0 if $n \neq m$ and 1 if $n = m$). It follows that $\langle F - S_N F, \chi_n \rangle = 0$ for $|n| \leq N$ and then that $\langle F - S_N F, p \rangle = 0$ for any trigonometric polynomial p of degree ast most N.

Observe that

$$F - \sigma_N F = (F - S_N F) + (S_N F - \sigma_N F)$$

is a sum of $F - S_N F$ and a trigonometric polynomial $S_N F - \sigma_N F$ of degree at most N. It follows that $F - S_N F$ is orthogonal to $S_N F - \sigma_N F$.

For orthogonal vectors v, w in an inner product space we have

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle w, w \rangle = \|v\|^2 + \|w\|^2$$

and so

$$\begin{aligned} \|F - \sigma_N F\|_2^2 &= \|(F - S_N F) + (S_N F - \sigma_N F)\|_2^2 \\ &= \|F - S_N F\|_2^2 + \|S_N F - \sigma_N F\|_2^2 \\ &\geq \|F - S_N F\|_2^2 \end{aligned}$$

But Fejér's theorem (Corollary 2.6.1) applied to $B = L^2(\mathbb{T})$, we have $\lim_{N\to\infty} \|F - \sigma_N F\|_2 = 0$ and so the preceding inequality now implies $\lim_{N\to\infty} \|F - S_N F\|_2 = 0$. $\square$

2.6.7 Corollary (Parseval's identity for Fourier series)**.** *If $F \in L^2(\mathbb{T})$, then*

$$\|F\|_2 = \left(\sum_{n=-\infty}^{\infty} |\hat{F}(n)|^2 \right)^{1/2}$$

Proof. Using orthonormality of the caracters χ_n,

$$\|S_N F\|_2^2 = \langle S_N F, S_N F \rangle = \sum_{n=-N}^{N} |\hat{F}(n)|^2$$

Since $\lim_{N\to\infty} \|F - S_N F\|_2 = 0$, $\lim_{N\to\infty} (\|F\|_2 - \|S_N F\|_2) = 0$ (because the triangle inequality for norms implies $|\|v\| - \|w\|| \leq \|v - w\|$ for vectors v and w in any normed space). Hence $\|F\|_2 = \lim_{N\to\infty} \|S_N F\|_2$ gives the result. $\square$

2.6.8 Corollary (Fourier characterization of $L^2(\mathbb{T})$)**.** *If $F \in L^1(\mathbb{T})$, then $F \in L^2(\mathbb{T})$ holds if and only if*

$$\sum_{n=-\infty}^{\infty} |\hat{F}(n)|^2 < \infty$$

In fact, if $(a_n)_{n\in\mathbb{Z}}$ is a sequence of scalars with $\sum_{n=-\infty}^{\infty} |a_n|^2 < \infty$, then there is $F \in L^2(\mathbb{T})$ with $\hat{F}(n) = a_n$ for all $n \in \mathbb{Z}$.

Proof. We already know (from Parseval's identity, Corollary 2.6.7) that if $F \in L^2(\mathbb{T})$, then the Fourier coefficients are square summable.

For the converse, if $\sum_{n=-\infty}^{\infty} |\hat{F}(n)|^2 < \infty$, let $s_N \in L^2(\mathbb{T})$ be the trigonometric polynomial $s_N(\zeta) = S_N F(\zeta)$.

We claim that (s_N) is a Cauchy seqeunce in L^2 and that the Fourier coefficients of the limit $s = \lim_{N\to\infty} s_N \in L^2(\mathbb{T})$ agree with those of F.

If $N < M$, we have $s_M - s_N = \sum_{n=N+1}^{M} \hat{F}(n)\chi_n + \sum_{n=-M}^{-N-1} \hat{F}(n)\chi_n$ (where $\chi_n(\zeta) = \zeta^n$). Using orthonormaility of the χ_n we get

$$\|s_M - s_N\|_2^2 = \sum_{N<|n|\leq M} |\hat{F}(n)|^2 = \|s_M\|_2^2 - \|s_N\|^2$$

By assumption

$$\sum_{n=-\infty}^{\infty} |\hat{F}(n)|^2 = \lim_{N\to\infty} \sum_{n=-N}^{N} |\hat{F}(n)|^2 = \lim_{N\to\infty} \|s_N\|_2^2 < \infty$$

and so $\|s_M\|_2^2 - \|s_N\|^2$ is small if $M > N$ with N large enough. Thus $\|s_M - s_N\|_2$ is small for M, N large, which verifies the Cauchy condition of $(s_N)_{N=1}^{\infty}$ in $L^2(\mathbb{T})$.

Let $s = \lim_{N\to\infty} s_N \in L^2(\mathbb{T})$. Then $\hat{s}(n) = \langle s, \chi_n\rangle = \lim_{N\to\infty}\langle s_N, \chi_n\rangle$ (using continuity of the inner product, something that follows quickly from the Cauchy-Schwarz inequality). But $\langle s_N, \chi_n\rangle = 0$ if $N < |n|$ and $\langle s_N, \chi_n\rangle = \hat{F}(n)$ for $|n| \leq N$. So $\lim_{N\to\infty}\langle s_N, \chi_n\rangle = \hat{F}(n)$.

Thus $\hat{s}(n) = \hat{F}(n)$ for all $n \in \mathbb{Z}$. We have $F \in L^1(\mathbb{T})$ and $s \in L^2(\mathbb{T}) \subset L^1(\mathbb{T})$. Thus $F = s$ by Corollary 2.6.5. Hence $F \in L^2(\mathbb{T})$.

If we do not assume that we have $F \in L^1(\mathbb{T})$ to begin with, but just assume that we have scalars a_n with $\sum_{n=-\infty}^{\infty} |a_n|^2 < \infty$, then we can define $s_N(\zeta) = \sum_{n=-N}^{N} a_n\zeta^n$. We get in the same way as above that there is a limit $F = \lim_{N\to\infty} s_N$ in $L^2(\mathbb{T})$. To show that $\hat{F}(n) = a_n$ for each $n \in \mathbb{Z}$, use $\hat{F}(n) = \langle F, \chi_n\rangle = \lim_{N\to\infty}\langle S_N, \chi_n\rangle$ where $\chi_n(\zeta) = \zeta^n$. (This is by continuity of the inner product on $L^2(\mathbb{T})$.) But $\langle S_N, \chi_n\rangle = a_n$ once $N > |n|$ by orthonormality of the χ_n and hence the limit is a_n. That is $\hat{F}(n) = a_n$. □

2.6.9 Corollary (Riemann Lebesgue lemma)**.** *If $F \in L^1(\mathbb{T})$, then $\lim_{|n|\to\infty} \hat{F}(n) = 0$.*

Proof. We have $F = \lim_{N\to\infty} \sigma_N F$ in $L^1(\mathbb{T})$, or in other terms $\lim_{N\to\infty} \|\sigma_N F - F\|_1 = 0$ by Fejér's theorem (Corollary 2.6.1) applied to $B = L^1(\mathbb{T})$. So if $\varepsilon > 0$ is given, then we can find N so that $\|\sigma_N F - F\|_1 < \varepsilon$.

Let $H = F - \sigma_N F$. Then $\hat{H}(n) = \hat{F}(n)$ for $n > N$ since $\sigma_N F$ is a trigonometric polynomial of degree at most N (and so $\widehat{(\sigma_N F)}(n) = 0$ for $n > N$). But $|\hat{H}(n)| \leq \|H\|_1 < \varepsilon$ for all n.

Thus $|\hat{F}(n)| = |\hat{H}(n)| < \varepsilon$ for $n > N$. $\square$

2.6.10 Remark. We cannot have uniform convergence of $\sigma_N F$ to F unless $F \in C(\mathbb{T})$ because the $\sigma_N F$ are trigonometric polynomials, so that $\sigma_N F \in C(\mathbb{T})$, and uniform limits of continuous functions are continuous.

But we can have pointwise converegence in some cases, that is $\lim_{N\to\infty} \sigma_N F(\zeta) = F(\zeta)$ for all $\zeta \in \mathbb{T}$ (or for almost all ζ). For instance we say that $F(\zeta)$ has a jump discontinuity at $\zeta = e^{2\pi i x_0}$ if both one-sided limits

$$f(x_0^{\pm}) = \lim_{h\to 0+} F(e^{2\pi i (x_0 \pm h)})$$

exist. In this case, taking $\check{f}(x_0) = (f(x_0^+) + f(x_0^-))/2$ we have that $f(x) = F(e^{2\pi i x})$ satisfies

$$\lim_{h\to 0^+} \int_0^h \left| \frac{f(x_0+\tau) + f(x_0-\tau)}{2} - \check{f}(x_0) \right| d\tau = 0 \tag{2.6.1}$$

2.6.11 Theorem. *If $F \in L^1(\mathbb{T})$, $f(x) = F(e^{2\pi i x})$ and $x_0 \in \mathbb{R}$ is a point where there exists a value $\check{f}(x_0) \in \mathbb{C}$ satisfying (2.6.1), then*

$$\lim_{N\to\infty} \sigma_N F(e^{2\pi i x_0}) = \check{f}(x_0)$$

Proof. We will not give this. See Y. Katznelson, *An introduction to Harmonic Analysis* (Dover edition, 1976), p. 20, where it is attributed to Lebesgue. $\square$

A Appendix

A.1 General measures and Fubini's theorem

A.1.1 Definition (MA2224 Definition 3.1.2)**.** If X is a set, then a collection Σ of subsets of X is called a *σ-algebra of subsets* of X if it satisfies

(σAlg-1) $\emptyset \in \Sigma$

(σAlg-2) $E \in \Sigma \Rightarrow X \setminus E \in \Sigma$

(σAlg-3) $E_1, E_2, \ldots \in \Sigma$ implies $\bigcup_{n=1}^{\infty} E_n \in \Sigma$

The pair (X, Σ) is then called a *measurable space*.

A.1.2 Definition (similar to MA2224 Definition 3.2.1)**.** If (X, Σ) is a measurable space, then a function $f \colon X \to \mathbb{R}$ is called *Σ-measurable* if for each $a \in \mathbb{R}$

$$\{x \in X : f(x) \leq a\} = f^{-1}((-\infty, a]) \in \Sigma.$$

A.1.3 Definition (MA2224 Definition 3.1.3)**.** If (X, Σ) is a measurable space and $\mu \colon \Sigma \to [0, \infty]$ is a function, then we call μ a *measure* on Σ if it satisfies

(Meas1) $\mu(\emptyset) = 0$

(Meas2) μ is countably additive, that is whenever $E_1, E_2, \ldots \in \Sigma$ are disjoint, then $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$.

The triple (X, Σ, μ) is then called a *measure space*.

A.1.4 Definition. A function $f \colon X \to \mathbb{R}$ is called a *simple function* if the range $f(X)$ is a finite set.

If (X, Σ_X) is a measurable space, then we have also *measurable simple functions* $f \colon X \to \mathbb{R}$, that is simple functions that are measurable.

A.1.5 Definition. If (X, Σ, μ) is a measurable space and $f \colon X \to [0, \infty)$ is a (nonnegative) measurable simple function with range $f(X) = \{y_1, y_2, \ldots, y_n\}$, then the Lebesgue integral $\int_X f \, d\mu$ is defined as

$$\sum_{j=1}^{n} y_j \mu(f^{-1}(\{y_j\})$$

A.1.6 Definition. If (X, Σ, μ) is a measurable space and $f \colon X \to [0, \infty)$ is a (nonnegative) measurable function then the Lebesgue integral $\int_X f \, d\mu$ is defined as

$$\int_X f \, d\mu = \sup \left\{ \int_X s \, d\mu : s \text{ simple measurable and } 0 \leq s(x) \leq f(x) \forall x \in X \right\}$$

A.1.7 Definition. If (X, Σ, μ) is a measurable space and $f \colon X \to \mathbb{R}$ is a measurable function f is called *integrable* if $\int_X |f| \, d\mu < \infty$. If f is integrable then $\int_X f \, d\mu$ is defined as

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu$$

where $f^+, f^- \colon X \to \mathbb{R}$ are defined by

$$f^+(x) = \max(f(x), 0) \text{ and } f^-(x) = \max(-f(x), 0).$$

A.1.8 Definition. If (X, Σ) is a measurable space and $f \colon X \to \mathbb{C}$ is a function, then f is called measurable if $\operatorname{Re} f$ and $\operatorname{Im} f$ are both measurable ($\mathbb{R}$-valued functions on X).

If (X, Σ, μ) is a measure space and $f \colon X \to \mathbb{C}$ is a function, then f is called *integrable* if $\operatorname{Re} f$ and $\operatorname{Im} f$ are both integrable (equivalently if f is measurable and $\int_X |f| \, d\mu < \infty$) and then we define

$$\int_X f \, d\mu = \int_X \operatorname{Re} f \, d\mu + i \int_X \operatorname{Im} f \, d\mu$$

A.1.9 Example. If (X, d) is a metric space and Σ is a σ-algebra of subsets of X that contains all the open sets, then every continuous $f \colon X \to \mathbb{C}$ is measurable.

Proof. Then $u = \operatorname{Re} f \colon X \to \mathbb{R}$ is continuous and for each $a \in \mathbb{R}$, $u^{-1}((-\infty, a])$ is closed (as the inverse image of a closed set under a continuous function). The complement $X \setminus u^{-1}((-\infty, a])$ is open, hence in Σ by hypothesis and so $u^{-1}((-\infty, a]) \in \Sigma$ by the definition of a σ-algebra.

Similarly for $v = \operatorname{Im} f$. □

A.1.10 Definition. If (X, Σ, μ) is a measurable space and $P(x)$ is a statement about $x \in X$ that may be true or false for each $x \in X$, we say that $P(x)$ holds for *almost every* $x \in X$ (strictly for μ-almost every $x \in X$) if there is a set $E \in \Sigma$ with $\mu(E) = 0$ such that

$$\{x \in X : P(x) \text{ is false}\} \subseteq E$$

A.1.11 Definition. If (X, Σ_X) and (Y, Σ_Y) are measurable spaces, then a subset of $X \times Y$ of the form

$$A \times B \qquad (A \in \Sigma_X, B \in \Sigma_Y)$$

is called a *measurable rectangle* in $X \times Y$.

The smallest σ-algebra of subsets of $X \times Y$ that contains all measurable rectangles is denoted $\Sigma_X \times \Sigma_Y$.

A.1.12 Theorem. *If (X, Σ_X, μ_X) and (Y, Σ_Y, μ_Y) are measure spaces, then there is a measure λ on $(X \times Y, \Sigma_X \times \Sigma_Y)$ called the* product measure *with the properties*

(PM1) if $A \times B$ is a measurable rectangle in $X \times Y$, then $\lambda(A \times B) = \mu_X(A)\mu_Y(B)$

(PM2) *if* $Q \in \Sigma_X \times \Sigma_Y$ *and*

$$Q_x = \{y \in Y : (x, y) \in Q\}, Q^y = \{x \in X : (x, y) \in Q\}$$

then $Q_x \in \Sigma_Y \forall x \in X$ *and* $Q^y \in \Sigma_X \forall y \in Y$.

(PM3) *if* $Q \in \Sigma_X \times \Sigma_Y$ *and*

$$\phi(x) = \mu_Y(Q_x), \psi(y) = \mu_X(Q^y)$$

then ϕ *is measurable on* (X, Σ_X), ψ *is measurable on* (Y, Σ_Y) *and*

$$\lambda(Q) = \int_X \phi d\mu_X = \int_Y \psi\, d\mu_Y$$

A.1.13 Theorem (Fubini's theorem)**.** *Let* (X, Σ_X, μ_X) *and* (Y, Σ_Y, μ_Y) *be measure spaces,* λ *the product measure on* $X \times Y$ *and* $f \colon X \times Y \to \mathbb{C}$ *a function which is measurable with respect to* $\Sigma_X \times \Sigma_Y$.

Then $x \mapsto f(x, y)$ *is measurable (with respect to* (X, Σ_X)*) for each* $y \in Y$ *and* $y \mapsto f(x, y)$ *is measurable (with respect to* (Y, Σ_Y)*) for each* $x \in X$.

If

$$\phi^*(x) = \int_{y \in Y} |f(x, y)|\, d\mu_Y(y) < \infty$$

for almost every $x \in X$*, and if* $\int_X \phi^*\, d\mu_X < \infty$*, then* f *is integrable.*

If f *is integrable, then* $\phi^*(x) < \infty$ *for almost every* $x \in X$, $\phi(x) = \int_{y \in Y} f(x, y)\, d\mu_Y(y)$ *defines an integrable function on* X *(where we take* $\phi(x) = 0$ *when* $\phi^*(x) = \infty$*) and*

$$\int_{X \times Y} f\, d\lambda = \int_X \phi\, d\mu_X = \int_{x \in X} \left(\int_{y \in Y} f(x, y)\, d\mu_Y(y) \right) d\mu_X(x)$$

Similar statements hold with the rôles of X *and* Y *reversed and in particular, if* f *is integrable, then*

$$\int_{X \times Y} f\, d\lambda = \int_{y \in Y} \left(\int_{x \in X} f(x, y)\, d\mu_X(x) \right) d\mu_Y(y)$$

For proofs, see W. Rudin, *Real and complex analysis*, McGraw-Hill (1974) [Chapter 7]

A.2 Facts about L^1

We need to know that $L^1(\mathbb{T})$ is complete in the norm $\|\cdot\|_1$ and that this is also true about $L^1[0,1]$ (which is the same space up to a change of variables in our definition) and $L^1(\mathbb{R})$.

We will also need to know that $C(\mathbb{T}) \subseteq L^1(\mathbb{T})$ is dense and that $C_c(\mathbb{R}) \subseteq L^1(\mathbb{R})$ is dense. (Recall that a subset of a metric space is called *dense* if its closure is the whole space.)

Here we outline how these are proved.

A.2.1 Definition. If $(E, \|\cdot\|)$ is a normed space then a *series* in E is just a sequence $(x_n)_{n=1}^\infty$ of terms $x_n \in E$.

We say that the *series converges* in E if the sequence of partial sums has a limit — $\lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{j=1}^n x_j$ exists in E, or there exists $s \in E$ so that

$$\lim_{n\to\infty} \left\| \left(\sum_{j=1}^n x_j \right) - s \right\| = 0$$

We say that a series $\sum_{n=1}^\infty x_n$ is *absolutely convergent* if $\sum_{n=1}^\infty \|x_n\| < \infty$.

A.2.2 Lemma. *Let (X,d) be a metric space in which each Cauchy sequence has a convergent subsequence. Then (X,d) is complete.*

Proof. Omitted. □

A.2.3 Proposition. *Let $(E, \|\cdot\|)$ be a normed space. Then E is a Banach space (that is complete) if and only if each absolutely convergent series $\sum_{n=1}^\infty x_n$ of terms $x_n \in E$ is convergent in E.*

Proof. Assume E is complete and $\sum_{n=1}^\infty \|x_n\| < \infty$. Then the partial sums of this series of positive terms

$$S_n = \sum_{j=1}^n \|x_j\|$$

must satisfy the Cauchy criterion. That is for $\varepsilon > 0$ given there is N so that $|S_n - S_m| < \varepsilon$ holds for all $n, m \geq N$. If we take $n > m \geq N$, then

$$|S_n - S_m| = \left| \sum_{j=1}^n \|x_j\| - \sum_{j=1}^m \|x_j\| \right| = \sum_{j=m+1}^n \|x_j\| < \varepsilon.$$

Then if we consider the partial sums $s_n = \sum_{j=1}^n x_j$ of the series $\sum_{n=1}^\infty x_n$ we see that for $n > m \geq N$ (same N)

$$\|s_n - s_m\| = \left\| \sum_{j=1}^n x_j - \sum_{j=1}^m x_j \right\| = \left\| \sum_{j=m+1}^n x_j \right\| \leq \sum_{j=m+1}^n \|x_j\| < \varepsilon.$$

It follows from this that the sequence $(s_n)_{n=1}^\infty$ is Cauchy in E. As E is complete, $\lim_{n\to\infty} s_n$ exists in E and so $\sum_{n=1}^\infty x_n$ converges.

For the converse, assume that all absolutely convergent series in E are convergent. Let $(u_n)_{n=1}^\infty$ be a Cauchy sequence in E. Using the Cauchy condition with $\varepsilon = 1/2$ we can find $n_1 > 0$ so that

$$n, m \geq n_1 \Rightarrow \|u_n - u_m\| < \frac{1}{2}.$$

Next we can (using the Cauchy condition with $\varepsilon = 1/2^2$) find $n_2 > 1$ so that

$$n, m \geq n_2 \Rightarrow \|u_n - u_m\| < \frac{1}{2^2}.$$

We can further assume (by increasing n_2 if necessary) that $n_2 > n_1$. Continuing in this way we can find $n_1 < n_2 < n_3 < \cdots$ so that

$$n, m \geq n_j \Rightarrow \|u_n - u_m\| < \frac{1}{2^j}.$$

Consider now the series $\sum_{j=1}^\infty x_j = \sum_{j=1}^\infty (u_{n_{j+1}} - u_{n_j})$. It is absolutely convergent because

$$\sum_{j=1}^\infty \|x_j\| = \sum_{j=1}^\infty \|u_{n_{j+1}} - u_{n_j}\| \leq \sum_{j=1}^\infty \frac{1}{2^j} = 1 < \infty.$$

By our assumption, it is convergent. Thus its sequence of partial sums

$$s_J = \sum_{j=1}^J (u_{n_{j+1}} - u_{n_j}) = u_{n_{J+1}} - u_{n_1}$$

has a limit in E (as $J \to \infty$). It follows that

$$\lim_{J\to\infty} u_{n_{J+1}} = u_{n_1} + \lim_{J\to\infty} (u_{n_{J+1}} - u_{n_1})$$

exists in E. So the Cauchy sequence $(u_n)_{n=1}^\infty$ has a convergent subsequence. By Lemma A.2.2 E is complete. □

A.2.4 Theorem. $(L^1[0,1], \|\cdot\|_1)$ *is a Banach space.*

Proof. We have discussed the fact that it is a normed space earlier.

To show completeness, we use Proposition A.2.3. If $\sum_{n=1}^\infty \|f_n\|_1 < \infty$, let

$$h(x) = \lim_{N\to\infty} \sum_{n=1}^N |f_n(x)| \quad (x \in [0,1])$$

with the understanding that $h(x) \in [0, +\infty]$. By the monotone convergence theorem

$$\begin{aligned}
\int_{[0,1]} h(x)\, d\mu(x) &= \lim_{N\to\infty} \int_{[0,1]} \sum_{n=1}^N |f_n(x)|\, d\mu(x) \\
&= \lim_{N\to\infty} \sum_{n=1}^N \int_{[0,1]} |f_n(x)|\, d\mu(x) = \lim_{N\to\infty} \sum_{n=1}^N \|f_n\|_1 \\
&= \sum_{n=1}^\infty \|f_n\|_1 < \infty
\end{aligned}$$

It follows that $h(x) < \infty$ almost everywhere on $[0,1]$ and so we can define $f(x)$ by taking $f(x) = \sum_{n=1}^\infty f_n(x)$ when $h(x) < \infty$ (and say $f(x) = 0$ if $h(x) = \infty$).

This $f(x)$ will be measurable and integrable because $|f(x)| \leq h(x)$ (and so $f \in L^1[0,1]$). We also have

$$\left| f(x) - \sum_{n=1}^N f_n(x) \right| = \left| \sum_{n=N+1}^\infty f_n(x) \right| \leq \sum_{n=N+1}^\infty |f_n(x)| \leq h(x)$$

for each x. The dominated convergence theorem then implies that

$$\begin{aligned} \lim_{N\to\infty} \left\| f - \sum_{n=1}^N f_n \right\|_1 &= \lim_{N\to\infty} \int_{[0,1]} \left| f(x) - \sum_{n=1}^N f_n(x) \right| \, d\mu(x) \\ &= \int_{[0,1]} \lim_{N\to\infty} \left| f(x) - \sum_{n=1}^N f_n(x) \right| \, d\mu(x) = 0 \end{aligned}$$

Thus each absolutely convergent series in $L^1[0,1]$ converges and so the space is complete (by Proposition A.2.3). □

A.2.5 Remark. The same proof works for $L^1(\mathbb{R})$.

A.2.6 Theorem (Lusin's theorem)**.** *Suppose that $f\colon [0,1] \to \mathbb{C}$ is measurable and $M = \sup_{x\in[0,1]} |f(x)| < \infty$. If $\varepsilon > 0$ is arbitrary, then there is a continuous $h\colon [0,1] \to \mathbb{C}$ such that*

(a) $\mu(\{x \in [0,1] : f(x) \neq h(x)\}) < \varepsilon$*, and*

(b) $\sup_{x\in[0,1]} |h(x)| \leq M$.

A proof can be found in W. Rudin, *Real and complex analysis*, McGraw-Hill (1974) [Theorem 2.23].

A.2.7 Corollary. $C(\mathbb{T}) \subseteq L^1(\mathbb{T})$ *is dense (or, more formally, the equivalence classes in* $L^1(\mathbb{T})$ *which have a continuous element are dense).*

Proof. We will show instead that $CP[0,1] = \{f \in C[0,1] : f(0) = f(1)\}$ is dense in $L^1[0,1]$ (which is equivalent in view of our definitions 1.1.18, 1.2.23 and 2.2.6).

Although there is a distinction between $CP[0,1]$ and $C[0,1]$, this distinction turns out to be not important in this proof, because $C[0,1]$ is contained in the closure of $CP[0,1]$ in the $\|\cdot\|_1$ norm. If $f \in C[0,1]$ we can define $f_n \in CP[0,1]$ by saying that $f_n(x) = f(x)$ for $0 \leq x \leq 1 - (1/n)$ and $f_n(x)$ is linear on the remainder, matching $f(0)$ at $x = 1$, that is

$$f_n(x) = \left(1 - \frac{x - 1 + \frac{1}{n}}{\frac{1}{n}}\right) f(1 - \frac{1}{n}) + \left(\frac{x - 1 + \frac{1}{n}}{\frac{1}{n}}\right) f(0) \qquad (1 - \frac{1}{n} < x \leq 1).$$

Then $\|f_n - f\|_1 = \int_{1-1/n}^1 |f_n(x) - f(x)| \, dx \to 0$ as $n \to \infty$.

Thus if we establish that $C[0,1]$ is dense in $L^1[0,1]$, it will follow that $CP[0,1]$ is dense.

If we consider first $f \in L^1[0,1]$ (or a representative $f \in \mathcal{L}^1[0,1]$ of its almost everywhere equivalence class) such that $M = \sup_{x\in[0,1]} |f(x)| = M < \infty$, then we can apply Lusin's theorem A.2.6 with $\varepsilon = 1/n$ to find $f_n \in C[0,1]$ with $\|f_n\|_\infty \leq M$ and $\mu(\{x : f(x) \neq f_n(x)\}) < 1/n$. Then

$$\begin{aligned} \|f_n - f\|_1 &= \int_{[0,1]} |f_n(x) - f(x)|\, d\mu(x) \\ &= \int_{\{x:f(x)\neq f_n(x)\}} |f_n(x) - f(x)|\, d\mu(x) \\ &\leq \mu(\{x : f(x) \neq f_n(x)\})2M < \frac{2M}{n} \end{aligned}$$

(using $|f_n(x) - f(x)| \leq |f_n(x)| + |f(x)| \leq 2M$).

For general $f \in L^1[0,1]$, let $h_n(x) = f(x)$ for $|f_n(x)| \leq n$ and $h_n(x) = 0$ for $|f_n(x)| \geq n$. Then $h_n \in L^1[0,1]$ and h_n is in the closure of $C[0,1]$ by the previous paragraph. We have

$$\begin{aligned} \|h_n - f\|_1 &= \int_{[0,1]} |h_n(x) - f(x)|\, d\mu(x) \\ &= \int_{\{x:|f(x)|\geq n\}} |h_n(x) - f(x)|\, d\mu(x) \\ &= \int_{\{x:|f(x)|\geq n\}} |f(x)|\, d\mu(x) \\ &= \int_{[0,1]} \chi_{\{y:|f(y)|\geq n\}}(x)|f(x)|\, d\mu(x) \\ &\to 0 \text{ as } n \to \infty \end{aligned}$$

by the dominated convergence theorem. (Here the notation χ_E denotes the characteristic function of a subset E.) It follows that f is in the closure of $C[0,1]$, hence that $C[0,1]$ is dense. $\square$

A.2.8 Definition. If (X, d) is a metric space, the a *completion* of (X, d) is a complete metric space (Y, ρ) together with a distance-preserving map $\alpha\colon X \to Y$ (*i.e.* one that satisfies $d(x_1, x_2) = \rho(\alpha(x_1), \alpha(x_2))$ for $x_1, x_2 \in X$) such that $\alpha(X)$ is dense in Y.

A.2.9 Remark. For each $y \in Y$ there is a sequence $(x_n)_{n=1}^\infty$ such that $\lim_{n\to\infty} \alpha(x_n) = y$. The sequence $(x_n)_{n=1}^\infty$ must be a Cauchy sequence and if $(x'_n)_{n=1}^\infty$ is another choice of such a sequence then $\lim_{n\to\infty} d(x_n, x'_n) = 0$. This latter relation is an equivalence relation on Cauchy sequences in X and the points of Y must correspond bijectively to equivalence classes of Cauchy sequences in X. This allows one to construct a completion and to show that all completions of X are 'essentially' the same.

One often treats the completion Y as 'containing' the original metric space, so that the map α becomes inclusion.

A.2.10 Corollary. *$L^1(\mathbb{T})$ with the distance arising from the norm $\|\cdot\|_1$ is the completion of $C(\mathbb{T})$ with the same distance.*

$L^1[0,1]$ is the completion of $CP[0,1]$ and of $C[0,1]$ (in the distance arising from the norm $\|\cdot\|_1$).

A.2.11 Theorem. *$(L^1(\mathbb{R}), \|\cdot\|_1)$ is a Banach space.*

$C_c(\mathbb{R})$ is dense in $L^1(\mathbb{R})$ and $L^1(\mathbb{R})$ is the completion of $C_c(\mathbb{R})$ (in the distance arising from the norm $\|\cdot\|_1$).

Proof. The proof of completion is basically identical to that of Theorem A.2.4.

The proof that $C_c(\mathbb{R})$ is dense relies on Lusin's theorem more or less as in Corollary A.2.7 but one has to reduce to $L^1[-N,N]$ and continuous functions supported in $[-N,N]$. This is not hard because if $f \in L^1(\mathbb{R})$, then

$$\lim_{N\to\infty} \|f\chi_{[-N,N]} - f\|_1 = 0$$

by the dominated convergence theorem. □

A.3 Facts about L^2

We defined $L^2(\mathbb{T})$ in Examples 2.4.2. Here we repeat the definition a little more formally. Recall that $\mathbb{K}$ denotes one of $\mathbb{R}$ or $\mathbb{C}$. We usually deal with the case $\mathbb{K} = \mathbb{C}$ but the definitions and results in this section are also valid over $\mathbb{R}$, with the same proofs.

A.3.1 Definition. The space $\mathcal{L}^2[0,1]$ is the space of measurable $f\colon [0,1] \to \mathbb{K}$ such that

$$\int_{[0,1]} |f(x)|^2 \, d\mu(x) < \infty$$

We define a 'magnitude' $\|f\|_2$ for $\mathcal{L}^2[0,1]$ as the square root of the above integral.

The space $\mathcal{L}^2(\mathbb{T})$ is the space of measurable $F\colon \mathbb{T} \to \mathbb{K}$ such that $f(x) = F(e^{2\pi i x})$ is in $\mathcal{L}^2[0,1]$. We define $\|F\|_2 = \|f\|_2$.

The space $\mathcal{L}^2(\mathbb{R})$ is the space of measurable $f\colon \mathbb{R} \to \mathbb{K}$ such that

$$\int_{\mathbb{R}} |f(x)|^2 \, d\mu(x) < \infty$$

We define $\|f\|_2$ for $f \in \mathcal{L}^2(\mathbb{R})$ as $\|f\|_2 = \left(\int_{\mathbb{R}} |f|^2 \, d\mu\right)^{1/2}$.

A.3.2 Lemma. *If $(V, \langle \cdot, \cdot \rangle)$ is an inner product space, then $\|v\| = \sqrt{\langle v, v \rangle}$ defines a norm on V.*

Proof. We take this as familiar. (See Lemma 1.1.4.6.) The triangle inequality follows from the Cauchy-Schwarz inequality $|\langle v, w \rangle| \leq \sqrt{\langle v, v \rangle}\sqrt{\langle w, w \rangle}$, which holds for $v, w \in V$. The inequality follows from $\langle v + \lambda w, v + \lambda w \rangle \geq 0$ by choosing a suitable $\lambda \in \mathbb{K}$. $\square$

A.3.3 Proposition. *$\mathcal{L}^2([0,1])$ is vector space over $\mathbb{K}$ and $\|\cdot\|_2$ is a seminorm on the space.*

The associated normed space $L^2([0,1]) = \mathcal{L}^2([0,1])/\{f : \|f\|_2 = 0\}$ is an inner product space with inner product given by

$$\langle f, h \rangle = \int_{[0,1]} f(x)\overline{h(x)} \, d\mu(x)$$

Proof. It is quite straightforward that $\lambda f \in \mathcal{L}^2([0,1])$ if $f \in \mathcal{L}^2([0,1])$ and $\lambda \in \mathbb{K}$, also that $\|\lambda f\|_2 = |\lambda| \|f\|_2$.

If $f, h \in \mathcal{L}^2([0,1])$, then

$$|f(x)\overline{h(x)}| \leq \frac{1}{2}(|f(x)|^2 + |h(x)|^2)$$

and it follows that $f(x)\overline{h(x)}$ is integrable. As

$$|f(x) + h(x)|^2 = (f(x) + h(x))\overline{(f(x) + h(x))} = |f(x)|^2 + f(x)\overline{h(x)} + h(x)\overline{f(x)} + |h(x)|^2$$

it follows then that $|f + h|^2$ is integrable and $f + h \in \mathcal{L}^2([0,1])$. So $\mathcal{L}^2([0,1])$ is a vector space.

The subset $\{f \in \mathcal{L}^2([0,1]) : \|f\|_2 = 0\}$ coincides with those $f \in \mathcal{L}^2([0,1])$ such that $f(x) = 0$ almost everywhere and it can be checked that this is a vector subspace. So $L^2[0,1]$ is a vector space quotient, hence a vector space.

We also have observed that the integral giving $\langle f, h\rangle$ makes sense for $f, h \in \mathcal{L}^2([0,1])$.

It is easy to check that the function $\langle \cdot, \cdot\rangle$ satisfies the properties of an inner product, except that $\langle f, f\rangle = 0$ only implies that $f(x) = 0$ almost everywhere. It then follows as in the proof of Lemma A.3.2 that $\|f\|_2 = \sqrt{\langle f, f\rangle}$ defines a seminorm on $\mathcal{L}^2([0,1])$ and that the analogue of the Cauchy-Schwarz inequality $|\langle f, h\rangle| \leq \|f\|_2\|h\|_2$ holds for $f, h \in \mathcal{L}^2([0,1])$.

If $f(x) = f_1(x)$ almost everywhere and $h(x) = h_1(x)$ almost everywhere, then $f(x)\overline{h(x)} = f_1(x)\overline{h_1(x)}$ almost everywhere and so $\langle f, h\rangle$ can be defined for almost everywhere equivalence classes of $f, h \in \mathcal{L}^2([0,1])$, that is for $f, h \in L^2([0,1])$. Then $(L^2([0,1]), \langle \cdot, \cdot\rangle)$ is an inner product space and $\|f\|_2 = \sqrt{\langle f, f\rangle}$ is a norm on $L^2[0,1]$. □

A.3.4 Theorem. *$(L^2[0,1], \|\cdot\|_2)$ is a Banach space.*

Proof. Following the lines of proof of Theorem A.2.4 we take an absolutely convergent series $\sum_{n=1}^\infty f_n$, where $f_n \in L^2[0.1]$ for each n and $\sum_{n=1}^\infty \|f_n\|_2 < \infty$.

Let

$$h(x) = \sum_{n=1}^\infty |f_n(x)| \in [0, \infty].$$

By the monotone convergence theorem

$$\begin{aligned}
\int_{[0,1]} h(x)^2\, d\mu(x) &= \lim_{N\to\infty} \int_{[0,1]} \left(\sum_{n=1}^N |f_n(x)|\right)^2 d\mu(x) \\
&= \lim_{N\to\infty} \left\| \sum_{n=1}^N |f_n| \right\|_2^2 \\
&\leq \lim_{N\to\infty} \left(\sum_{n=1}^N \|f_n\|_2\right)^2
\end{aligned}$$

where we have used the triangle inequality for $\|\cdot\|_2$ plus $\||f_n|\|_2 = \|f_n\|_2$.

Now proceed almost exactly as in proof of Theorem A.2.4 by taking $f(x) = \sum_{n=1}^\infty f_n(x)$ almost everywhere. From $|f(x)| \leq h(x)$ we get $f \in \mathcal{L}^2[0,1]$ and then use the dominated convergence theorem to show $\lim_{N\to\infty} \left\| f - \sum_{n=1}^N f_n \right\|_2^2 = 0$. □

A.3.5 Corollary. *$L^2(\mathbb{T})$ is a Banach space.*

A.3.6 Theorem. *$CP[0,1]$ is dense in $L^2[0,1]$.*

Proof. This follows from Lusin's theorem in a way that is very similar to the proof of Corollary A.2.7 □

A.3.7 Corollary. *$C(\mathbb{T}) \subseteq L^2(\mathbb{T})$ is dense (or, more formally, the equivalence classes in $L^2(\mathbb{T})$ which have a continuous element are dense).*

$L^2(\mathbb{T})$ with the distance arising from $\|\cdot\|_2$ is the completion of $C(\mathbb{T})$ with the same distance.

Proof. This is really just a restatement of Corollary A.3.6 and Theorem A.3.4. □

We can also show, as in Theorem A.2.11, that there is are similar results for functions on $\mathbb{R}$.

A.3.8 Theorem. *$(L^2(\mathbb{R}), \|\cdot\|_2)$ is a Banach space.*

$C_c(\mathbb{R})$ is dense in $L^2(\mathbb{R})$ and $L^2(\mathbb{R})$ is the completion of $C_c(\mathbb{R})$ (in the distance arising from the norm $\|\cdot\|_2$).

Changes 2/11/2017: Fix typos in Definition A.3.1. Add remark in proof of Lemma A.3.2. Remove obsolete Remark before Proposition 2.4.7.

Changes 27/11/2017: Add word Banach in Definition 2.4.1. Add remark in Definition 2.4.3. Fix comment following Definition 2.5.1. Clarify proof of Proposition 2.5.2*2 Fix typo in Corollary 2.6.5. Fix typo in Definition A.1.7.
Richard M. Timoney (November 27, 2017)