Chapter 1

Fourier Series

Consider the Hilbert space $L^2[0,T]$, $T>0$, with inner product

$$\langle f,g\rangle = \frac{1}{T}\int_0^T f(t)\overline{g(t)}\,dt, \qquad f,g\in L^2[0,T].$$

The space $L^2[0,T]$, or equivalently $L^2[-T/2,T/2]$, is a model for T-periodic signals with finite energy (on each interval of length T).

To simplify the notation we take $T=2\pi$. The general case can be obtained just by rescaling.

Lemma 1. *The system $\{e^{int}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2[0,2\pi]$.*

Proof. The orthonormality is clear, by the 2π-periodicity of the exponentials: letting $e_n(t)=e^{int}$,

$$\langle e_n,e_m\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-m)t}\,dt = \begin{cases} 0 & \text{if } n\neq m \\ 1 & \text{if } n=m. \end{cases}$$

To see the completeness we identify the interval $[0,2\pi]$ with the boundary of the unit disk $\mathbb{D}$ in $\mathbb{C}$: each $t\in[0,2\pi]$ yields $e^{it}\in\partial\mathbb{D}$ and reciprocally. Then a given function $f\in L^2[0,2\pi]$ gives rise to a function $g\in L^2(\partial\mathbb{D})$ in the obvious way, $g(e^{it})=f(t)$.

It is enough to see that the exponentials $e_n(t)$, $n\in\mathbb{Z}$, are dense in the space of continuous functions of compact support $\mathcal{C}_c[0,2\pi]$, since this subspace is dense in the L^2 norm in $L^2[0,2\pi]$ (see the Annex 1.4 at the end of the chapter in case you haven't seen this).

Let thus $f\in\mathcal{C}_c[0,2\pi]$ and consider its identification $g\in\mathcal{C}_c(\partial\mathbb{D})$. Since $\partial\mathbb{D}$ is a compact set in $\mathbb{C}\simeq\mathbb{R}^2$ we can use Weierstrass approximation theorem to obtain a sequence of polynomials in two variables $\{p_k(x,y)\}_{k\geq 1}$ converging uniformly on $\partial\mathbb{D}$. Then, taking (x,y) with $x+iy=e^{it}\in\partial\mathbb{D}$ (i.e. $x=\cos t$, $y=\sin t$), we have

$$\lim_{k\to\infty} p_k(x,y) = g(e^{it}) = f(t)$$

uniformly in t. Writing the polinomials p_k in terms of $x = \frac{z+\bar{z}}{2}$ and $y = \frac{z-\bar{z}}{2i}$ (or equivalently, using that $\cos t = (e^{it} + e^{-it})/2, \sin t = (e^{it} - e^{-it})/(2i)$) we obtain polinomials $\tilde{p}_k(z, \bar{z})$ such that

$$\lim_{k\to\infty} \tilde{p}_k(e^{it}, e^{-it}) = g(e^{it}) = f(t)$$

uniformly in $t \in [0, 2\pi]$. In particular

$$\lim_{k\to\infty} \|\tilde{p}_k - f\|_2^2 = \lim_{k\to\infty} \frac{1}{2\pi} \int_0^{2\pi} \left|\tilde{p}_k(e^{it}, e^{-it}) - f(t)\right|^2 dt = 0,$$

that is, $\tilde{p}_k(e^{it}, e^{-it})$, which is obviously generated by the $\{e_n\}_{n\in\mathbb{Z}}$, tends to f in $L^2[0, 2\pi]$. □

As a result of this lemma any function $f \in L^2[0, 2\pi]$ can be represented in the following way

$$f(t) = \sum_{n\in\mathbb{Z}} \langle f, e_n\rangle e_n(t).$$

Definition 1. Given $n \in \mathbb{Z}$, the n^{th} *Fourier coefficient* of a function $f \in L^2[0, 2\pi]$ is

$$\hat{f}(n) = \langle f, e_n\rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)e^{-int}dt.$$

Observe that, in this language, the previous identity takes the form

$$f(t) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{int}. \tag{1.1}$$

The series on the right hand side of this equality is called the *Fourier series of* f, and is usually denoted by $Sf(t)$.

Remark 1. Identity (1.1) is in $L^2[0, 2\pi]$, meaning that

$$\lim_{N\to\infty} \int_0^{2\pi} \Big|f(t) - \sum_{|n|\le N} \hat{f}(n)e^{int}\Big|^2 dt = 0.$$

In particular, given a $t_0 \in [0, 2\pi]$ we don't know anything about the pointwise convergence of the series $\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{int_0}$.

Also, as a consequence of the general theory of Hilbert spaces, from (1.1) we have the following identity, which is nothing but Parseval's identity in this setting.

Plancherel identity. *For* $f \in L^2[0, 2\pi]$

$$\|f\|^2_{L^2[0,2\pi]} = \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2.$$

Remark 2. Analogous definitions and formulas hold for the space $L^2[0,T]$ mentioned at the beginning of this chapter. The system $\{e^{in\frac{2\pi}{T}t}\}_{n\in\mathbb{Z}}$ is an orthonormal basis and the corres Fourier coefficients ar given by

$$\hat{f}(n) = \frac{1}{T}\int_0^T f(t)\, e^{-in\frac{2\pi}{T}t} dt.$$

Then any $f \in L^2[0,T]$ can be expressed as $f(t) = \sum_{n\in\mathbb{Z}} \hat{f}(n)\, e^{in\frac{2\pi}{T}t}$ and

$$\|f\|^2_{L^2[0,T]} = \frac{1}{T}\int_0^T |f(t)|^2 dt = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2.$$

Exercise 1. Reorganise the Fourier series in (1.1), using $e^{int} = \cos(nt) + i\sin(nt)$, to prove that

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(nt) + i \sum_{n=1}^{\infty} b_n \sin(nt),$$

where

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_0^{2\pi} f(t)\, \cos(nt)\, dt, \qquad n \geq 1 \\ b_n &= \frac{1}{\pi}\int_0^{2\pi} f(t)\, \sin(nt)\, dt, \qquad n \geq 1 \\ a_0 &= \frac{1}{2\pi}\int_0^{2\pi} f(t)\, dt. \end{aligned}$$

1.1 Extension and pointwise convergence of the Fourier series

As mentioned before (Remark 1), the identity (1.1) is to be understood in the L^2 sense. In this situation at least two fundamental questions arise:

1) Does the Fourier series make sense for functions f other than in $L^2[0,2\pi]$? If so, how is the series related to f?

2) What conditions on f ensure that, given a $t_0 \in [0,2\pi]$, the Fourier series $Sf(t_0)$ converges pointwise and coincides with $f(t_0)$?

These are (very) delicate matters that we shall not really discuss. We shall only state a couple of elementary results.

1.1.1 Fourier coefficients for L^1 functions

Let

$$L^1[0,2\pi] = \big\{ f : [0,2\pi] \longrightarrow \mathbb{C} : \|f\|_1 := \frac{1}{2\pi}\int_0^{2\pi} |f(t)|\, dt < \infty \big\},$$

and observe that an immediate application of the Cauchy-Schwartz inequality shows that $L^2[0,2\pi] \subset L^1[0,2\pi]$ and $\|f\|_1 \le \|f\|_2$:

$$\frac{1}{2\pi}\int_0^{2\pi} |f(t)|\,dt \le \left(\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\,dt\right)^{1/2} \left(\frac{1}{2\pi}\int_0^{2\pi} 1^2 dt\right)^{1/2} = \|f\|_2.$$

A straightforward estimate shows that $\hat{f}(n)$ is well defined for $f \in L^1[0,2\pi]$. Actually,

Lemma 2. *Let $f \in L^1[0,2\pi]$; then:*

(a) $|\hat{f}(n)| \le \|f\|_1$ for all $n \in \mathbb{Z}$,

(b) $\lim\limits_{|n|\to\infty} |\hat{f}(n)| = 0$ (Riemann-Lebesgue lemma).

Proof. (a) Directly

$$|\hat{f}(n)| = \left|\frac{1}{2\pi}\int_0^{2\pi} f(t)\, e^{-int} dt\right| \le \frac{1}{2\pi}\int_0^{2\pi} |f(t)|\,dt = \|f\|_1.$$

(b) Assume first that $f \in \mathcal{C}^1(0,2\pi)$ and $f(0) = f(2\pi)$, that is, that f seen as a 2π-periodic function in $\mathbb{R}$ is continuous everywhere. An integration by parts yields

$$\hat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(t) e^{-int} dt = \frac{1}{2\pi}\left[f(t)\frac{e^{-int}}{-in}\right]_0^{2\pi} + \frac{1}{in}\frac{1}{2\pi}\int_0^{2\pi} f'(t)e^{-int}dt = \frac{\hat{f'}(n)}{in}.$$

Then, by (a) applied to f',

$$|\hat{f}(n)| \le \frac{\|f'\|_1}{|n|} \overset{|n|\to\infty}{\longrightarrow} 0.$$

For general $f \in L^1[0,2\pi]$ let us see that for every $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $|\hat{f}(n)| < \epsilon$ for $|n| \ge n_0$. We do this by taking approximations by functions of the kind just considered. Let $g_\epsilon \in \mathcal{C}^1[0,2\pi]$ with $g_\epsilon(0) = g_\epsilon(2\pi)$ and $\|f - g_\epsilon\|_1 < \epsilon$. Then, by (a),

$$|\hat{f}(n)| \le |\hat{f}(n) - \widehat{g_\epsilon}(n)| + |\widehat{g_\epsilon}(n)| \le \|f - g_\epsilon\|_1 + |\widehat{g_\epsilon}(n)| \le \epsilon + |\widehat{g_\epsilon}(n)|.$$

Since $\{\widehat{g_\epsilon}(n)\}_n$ tends to 0 as $|n|$ goes to infinity, there exists $n_0 \in \mathbb{N}$ such that $|\widehat{g_\epsilon}(n)| < \epsilon$ for $|n| \ge n_0$, and therefore $|\hat{f}(n)| \le 2\epsilon$. □

Exercise 2. Let $f \in \mathcal{C}^k[0,2\pi]$ be such that it is also of class $\mathcal{C}^k$ once extended 2π-peridiocally to $\mathbb{R}$ (that is, such that $f^{(j)}(0) = f^{(j)}(2\pi)$ for all $j = 0,\dots,k$). Obtain an expression of $\widehat{f^{(k)}}(n)$ in terms of $\hat{f}(n)$ and deduce that $|\hat{f}(n)| = O(1/|n^k|)$ as $|n| \to \infty$.

1.1.2 Pointwise convergence of Fourier series

Given $f \in L^1[0, 2\pi]$ (or in any other space), when does its Fourier series converge? And when it does, what does it converge to? We shall not discuss in detail these questions. We just mention a couple of results.

Theorem 1 (Dirichlet). *Let $f \in L^1[0, 2\pi]$ and let $t \in (0, 2\pi)$ be such that $f(t^+)$, $f(t^-)$ exist, as well as $f'(t^+)$, $f'(t^-)$. Then*

$$Sf(t) := \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{int} = \frac{f(t^+) + f(t^-)}{2}.$$

In particular, if $f \in \mathcal{C}^1[0, 2\pi]$ the Fourier series $Sf(t)$ converges to $f(t)$ for all $t \in (0, 2\pi)$.

Remark. After proving this result Dirichlet conjectured that for for all continuous functions there is everywhere convergence of the Fourier series. This was disproved by Paul DuBois-Reymond in 1876, providing an example of a continuous function whose Fourier series diverges at a single point. Some years later, in 1923, Andrei N. Kolmogorov gave an example of an L^1 function for which its Fourier series diverges almost everywhere. This was refined in 1926 by the same Kolmogorov, who produced a L^1 whose Fourier series diverges everywhere. On the other hand Lennart Carleson proved in 1966 that for L^2 functions the Fourier series converges pointwise to the function almost everywhere. In summary, pointwise converge is a delicate matter (that we will avoid almost completely).

1.2 Convolution and the Dirichlet and Fejér kernels

In order to understand better the convergence (or not) of the partial sums of the Fourier series

$$S_N f(t) = \sum_{|n|\le N} \hat{f}(n)e^{int}, \quad N \in \mathbb{N},$$

as N tends to infinity write

$$\begin{aligned} S_N f(t) &= \sum_{|n|\le N} \hat{f}(n)e^{int} = \sum_{|n|\le N} \big(\frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}dx\big)e^{int} \\ &= \frac{1}{2\pi}\int_0^{2\pi} f(x)\big(\sum_{|n|\le N} e^{in(t-x)}\big)\, dx. \end{aligned}$$

Definition 2. The function

$$\sum_{|n|\le N} e^{int} = \frac{e^{-iNt} - e^{iNt}e^{it}}{1 - e^{it}} = \frac{e^{-i(N+1/2)t} - e^{i(N+1/2)t}}{e^{-it/2} - e^{it/2}} = \frac{\sin\big((N + 1/2)t\big)}{\sin(t/2)}$$

is called the *Dirichlet kernel*.

Definition 3. Given $f, g \in L^1[0, 2\pi]$, and identifying these functions with their 2π-periodic extensions to $\mathbb{R}$, the *convolution of f and g* is definied as the function

$$(f * g)(t) = \frac{1}{2\pi} \int_0^{2\pi} f(x)g(t-x)\,dx.$$

Observe that with this notation and terminology the above computation takes the form

$$S_N f(t) = (f * D_N)(t).$$

Thus, the convergence of the Fourier series is encoded not only on the properties of f but also in the behaviour of $D_N(t)$ as N tends to infinity which, as we shall point out now, is not very good.

Proposition 1. *Let $D_N(t)$ denote the Dirichlet kernel. Then,*

(a) $\displaystyle \frac{1}{2\pi} \int_0^{2\pi} D_N(t)\,dt = 1$

(b) $\displaystyle \frac{1}{2\pi} \int_0^{2\pi} |D_N(t)|\,dt \gtrsim \log N$ *as* $N \to \infty$.

Property (b) is at the root of many convergence issues with Fourier series.

Proof. (a) is straightforward if we write the Dirichlet kernel as a sum of exponentials.

(b) Since $\sin(t/2) \le t/2$ for $t \in [0, \pi]$ one has

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} |D_N(t)|\,dt &\ge \frac{1}{2\pi} \int_0^{2\pi} \frac{|2\sin\big((N+1/2)t\big)|}{t}\,dt \ge \frac{1}{2\pi} \int_0^{\pi} \frac{|\sin\big(2N+1)t\big)|}{t}\,dt \\ &= \frac{1}{2\pi} \int_0^{(2N+1)\pi} \frac{|\sin u|}{u}\,du \gtrsim \sum_{j=0}^{2N} \int_{j\pi}^{(j+1)\pi} \frac{|\sin u|}{u}\,du \\ &\ge \sum_{j=0}^{2N} \frac{1}{(j+1)\pi} \int_{j\pi}^{(j+1)\pi} |\sin u|\,du. \end{aligned}$$

Since, by periodicity of of the sine function,

$$\int_{j\pi}^{(j+1)\pi} |\sin u|\,du = 2$$

independently of j, we finally get

$$\frac{1}{2\pi} \int_0^{2\pi} |D_N(t)|\,dt \gtrsim \sum_{j=0}^{2N} \frac{1}{j+1} \ge \sum_{j=1}^{2N} \frac{1}{j} \simeq \log N.$$

$\square$

A standard way of improving the convergence is to change the summation method, taking averages of the sums previously considered. This is sometines mentioned as the *Cesàro summation method*. Consider

$$\sigma_N f = \frac{S_0 f + \cdots + S_N f}{N+1}.$$

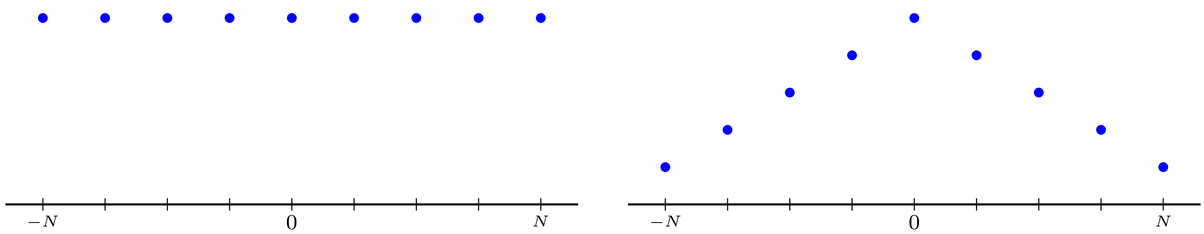
It is immediate to check that $\sigma_N f(t) = (f * K_N)(t)$, where

$$\begin{aligned} K_N(t) &= \frac{D_0(t) + \cdots D_N(t)}{N+1} = \sum_{|n|\le N} \Big(1 - \frac{|n|}{N+1}\Big)\, e^{int} = \frac{1}{N+1}\, \frac{\sin^2\big(\frac{N+1}{2}t\big)}{\sin^2(t/2)} \\ &= \frac{1}{N+1}\, \frac{1-\cos\big((N+1)t\big)}{1-\cos t} \end{aligned}$$

is the so-called *Fejér kernel*.

Lipót Fejér, born Lipót Weisz in 1880, was the thesis advisor of John von Neumann, Paul Erdős, György Pólya, Tibor Radó, Mihály Fekete, Marcel Riesz...

Observe that $K_N(t)$ charges the same exponetials as the Dirichlet kernel, but with different, more regular weights. Graphically:



Dirichlet kernel Fejér kernel

Proposition 2. *(a) For $f \in L^1[0,2\pi]$ the Fejér sums tend to f in the L^1-norm, i.e*

$$\lim_{N\to\infty} \|\sigma_N f - f\|_1 = 0.$$

(b) If $f \in \mathcal{C}[0,2\pi]$ (as a 2π-periodic function in $\mathbb{R}$, i.e., with $f(0) = f(2\pi)$) then

$$\lim_{N\to\infty} \sigma_N f(t) = f(t)$$

uniformly in $t \in [0,2\pi]$.

As we see, the behaviour of the summation formula depends on the properties of the corresponding convolution kernel.

Remark 3. An immediate consequence of part (a) is the uniqueness theorem for Fourier series of L^1 functions: if $f \in L^1[0,2\pi]$ is such that $\hat{f}(n) = 0$ for all $n \in Z$, then $f = 0$.

We conclude this section by pointing out that the action of the convolution is simplified at the Fourier side.

Theorem 2. *Let $f, g \in L^1[0, 2\pi]$, extended as 2π periodic functions in $\mathbb{R}$. Then*

(a) $f * g \in L^1[0, 2\pi]$ *and* $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$

(b) For all $n \in \mathbb{Z}$, $\widehat{(f * g)}(n) = \hat{f}(n)\, \hat{g}(n)$.

Proof. (a) By definition and a change in the order of integration

$$\begin{aligned}
\|f * g\|_1 &= \frac{1}{2\pi} \int_0^{2\pi} |(f * g)(t)| dt = \frac{1}{2\pi} \int_0^{2\pi} \left| \frac{1}{2\pi} \int_0^{2\pi} f(s)\, g(s-t)\, ds \right| dt \\
&\leq \frac{1}{2\pi} \int_0^{2\pi} |f(s)| \left(\frac{1}{2\pi} \int_0^{2\pi} |g(s-t)|\, dt \right) ds \\
&= \frac{1}{2\pi} \int_0^{2\pi} |f(s)| \left(\frac{1}{2\pi} \int_0^{2\pi} |g(u)|\, du \right) ds = \|f\|_1 \|g\|_1.
\end{aligned}$$

(b) Similarly

$$\begin{aligned}
\widehat{(f * g)}(n) &= \frac{1}{2\pi} \int_0^{2\pi} (f * g)(t)\, e^{-int} dt = \frac{1}{2\pi} \int_0^{2\pi} \big(\frac{1}{2\pi} \int_0^{2\pi} f(s) g(t-s)\, ds \big)\, e^{-int} dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} f(s) \big(\frac{1}{2\pi} \int_0^{2\pi} g(t-s) e^{-in(t-s)} dt \big)\, e^{-ins} ds \\
&= \frac{1}{2\pi} \int_0^{2\pi} f(s) (\hat{g}(n))\, e^{-ins} ds = \hat{g}(n)\, \hat{f}(n).
\end{aligned}$$

□

Remark 4. In case we consider $L^2[0, T]$ for a general interval $[0, T]$, $T > 0$, (i.e., in case we work with T-periodic functions) we follow the same arguments, starting with the orthonormal basis $\{e_n\}_{n \in \mathbb{Z}}$, defined as $e_n(t) = e^{in\frac{2\pi}{T}t}$.

1.3 Annex. The Haar system.

Let us see here a very different orthonormal basis of $L^2[0, 1]$, which will be relevant in the chapter devoted to wavelets.

Define the constant function $\chi \equiv 1$, which obviously belongs to $L^2[0, 1]$. Let also

$$\psi_0(t) = \begin{cases} -1 & \text{if } 0 \leq t < 1/2 \\ 1 & \text{if } 1/2 \leq t < 1. \end{cases}$$

Observe that the functions which are constant on each half of the interval $[0,1]$ are a linear combination of χ and ψ_0, since $\chi_{[0,1/2)} = \frac{1}{2}(\chi - \psi_0)$ and $\chi_{[1/2,1)} = \frac{1}{2}(\chi + \psi_0)$.

In each half of $[0,1]$ consider a function with the shape of ψ_0, rescaled so that its L^2 norm is 1, and do so for each dyadic interval of all generations. Explicitly, take, for all $n \geq 0$ and $k = 0, \ldots, 2^n - 1$, the functions

$$\psi_{n,k}(t) = 2^{n/2}\psi_0(2^n t - k) = \begin{cases} -2^{n/2} & \text{if } t \in [\frac{k}{2^n}, \frac{k+1/2}{2^n}) \\ 2^{n/2} & \text{if } t \in [\frac{k+1/2}{2^n}, \frac{k+1}{2^n}). \end{cases}$$

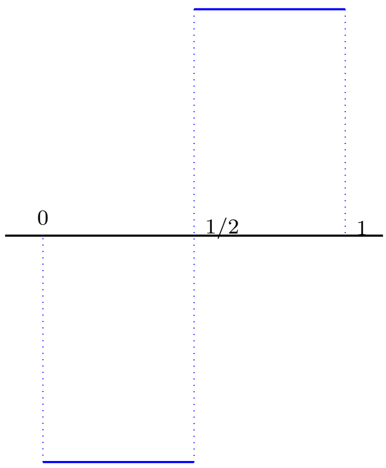


Figure 1.1: $\psi_0(t)$

Theorem 3. *Let $n \geq 0$ and $k = 0, \ldots, 2^n - 1$.*

(a) Each $\psi_{n,k}$ is supported in the dyadic interval on the n^{th} generation $I_{n,k} = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)$.

(b) $\int_0^1 \psi_{n,k}(t)\,dt = 0$ for all $n \geq 0$ and $k = 0, \ldots, 2^n - 1$.

(c) The system formed by χ and all $\psi_{n,k}$ is an orthonormal basis of $L^2[0,1]$.

Proof. (a) and (b) are immediate from the definition.

In order to prove (c) observe first that it is also immediate that $\|\psi_{n,k}\|_{L^2 0,1]} = 1$ for all n and k and that $\langle \chi, \psi_{n,k} \rangle = 0$ for all n and k. Therefore, it will be enough to prove that $\langle \psi_{n,k}, \psi_{m,j} \rangle = 0$ for $(n,k) \neq (m,j)$ and that $\{\chi, \psi_{n.k}\}_{n,k}$ spans the whole $L^2[0,1]$.

Let us see first the orthogonality $\langle \psi_{n,k}, \psi_{m,j} \rangle = 0$ for $(n,k) \neq (m,j)$. This is immediate in the cases $I_{n,k} \cap I_{m,j} = \emptyset$, since the respective supports of $\psi_{n,k}$ and $\psi_{m,j}$ are then disjoint. By construction of the dyadic intervals, if $I_{n,k} \cap I_{m,j} \neq \emptyset$ then either $I_{n,k} \subset I_{m,j}$ or $I_{m,j} \subset I_{n,k}$ (since $I_{n,k} \cap I_{n,j} = \emptyset$ if $j \neq k$).

Assume that $I_{n,k} \subset I_{m,j}$, that is, that $n > m$. The other case is treated analogously. In this situation $I_{n,k}$ is contained in one of the two halves of $I_{m,j}$, and therefore $\psi_{m,j}$ is constant on the support of

$\psi_{n,k}$ (either $-2^{m/2}$ or $2^{m/2}$, depending on which of the two halves of $I_{m,j}$ the interval $I_{n,k}$ is contained in). Then, by (b)

$$\langle \psi_{n,k}, \psi_{m,j} \rangle = \pm 2^{m/2} \int_0^1 \psi_{n,k}(t)\, dt = 0.$$

It remains to see that the system $\{\chi, \psi_{n,k}\}_{n,k}$ is complete. As we know, it is enough to see that this system is complete, in the L^2 norm, in the subspace of continuous functions in $[0, 1]$. Let us see next that it is enough to approximate functions which are constant on dyadic intervals of the same generation.

Lemma 3. *Let* $f \in \mathcal{C}[0, 1]$. *For all* $\epsilon > 0$ *there exist* $n \geq 1$ *and* $g_n = \sum_{k=0}^{2^n-1} \alpha_k \chi_{I_{n,k}}$, $\alpha_k \in \mathbb{C}$, *such that*

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

In particular $\|f - g_n\|_{L^2[0,1]} < \epsilon$.

We may think that g_n is an approximation of f at a "resolution" smaller than ϵ.

Proof. By the uniform continuity of f on $[0, 1]$, given $\epsilon > 0$ there exists $\delta > 0$ such that

$$|t - t'| < \delta \quad \Longrightarrow \quad |f(t) - f(t')| < \epsilon.$$

Take $n \geq 1$ such that $2^{-n} < \delta$ and let $\alpha_k = f(t_k)$, where t_k is any point in $I_{n,k}$. Define $g_n = \sum_{k=0}^{2^n-1} \alpha_k \chi_{I_{n,k}}$.

Given $t \in [0, 1]$ let k be such that $t \in I_{n,k}$, so that $|t - t_k| < 2^{-n} < \delta$ and therefore

$$|f(t) - g_n(t)| = |f(t) - f(t_k)| < \epsilon,$$

as desired. □

It remains to see that the system $\{\chi, \psi_{n,k}\}_{n,k}$ generates the functions which are constant on dyadic intervals of the same generation. For that it is enough to observe at scale n what we observed at level 1: functions which are constant on each half of $I_{n,k} = [k/2^n, (k + 1)/2^n)$ are a linear combination of $\psi_{n-1,j}$ and $\psi_{n,k}$, where $I_{n-1,j}$ is the "parent" of $I_{n,k}$ (that is, the interval at level $n - 1$ such that $I_{n,k} \subset I_{n-1,j}$). □

Remark. It is also possible to see directly that a given function g_n as in the previous lemma is a linear combination of χ and the finite $\psi_{m,j}$ with $m < n$, that is, that

$$g_n = \sum_{k=0}^{2^n-1} \alpha_k \chi_{I_{n,k}} = \langle g_n \chi \rangle \chi + \sum_{m=0}^{n-1} \sum_{j=0}^{2^m-1} \langle g_n, \psi_{m,j} \rangle \psi_{m,j}. \tag{1.2}$$

This is a matter of linear algebra, and is readily verified. If we look at this as a system of equations, one for each interval $I_{n,k}$, we have 2^n given values at the left hand side (the α_k) and

$$1+\sum_{m=0}^{n-1} 2^m = 1+(2+\cdots+2^{n-1}) = 1+\frac{2^n-1}{2-1} = 2^n$$

unknowns on the right hand side (the $\langle g_n, \chi\rangle$, $\langle g_n, \psi_{m,j}\rangle$, $m<n$, $j=0,\ldots,2^m-1$).

Observe that $I_{n,k}$ has exactly one ancestor at each level $m<n$ (that is, there is exactly one $j \in \{0,\ldots,2^m-1\}$ such that $I_{n,k}\subset I_{m,j}$. For that $I_{m,j}$,

$$\langle \chi_{I_{n,k}}, \psi_{m,j}\rangle\psi_{m,j} = \big(\int_{I_{n,k}} (\pm 2^{m/2})dt\big)(\pm 2^{m/2})\chi_{m,j} = 2^{m-n}\chi_{m,j},$$

while for the other $\psi_{m,l}$ one has $\langle \chi_{I_{n,k}}, \psi_{m,l}\rangle = 0$ (because $I_{n,k}\cap I_{m,j}=\emptyset$).

Thus, for example, for $t\in I_{n,k}$, the right-hand side of (1.2) is

$$2^{-n}+\sum_{m=0}^{n-1} 2^{m-n} = 2^{-n}+2^{-n}(1+\cdots+2^{n-1}) = 2^{-n}+2^{-n}\frac{2^n-1}{2-1} = 1,$$

so the identity (1.2) is checked for such t.

1.4 Annex. Urysohn's lemma

The subspace $\mathcal{C}_c[0,2\pi]$ is dense in $L^2[0,2\pi]$; in other words, for any function $f\in L^2[0,2\pi]$ and any $\epsilon>0$ there exists $g_\epsilon\in\mathcal{C}_c[0,2\pi]$ such that

$$\|f-g_\epsilon\|_2^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(t)-g_\epsilon(t)|^2\,dt < \epsilon.$$

Since by the general Lebesgue theory f is the limit of simple functions, it is enough to assume that f is simple, and since any simple function is a linear combination of characteristic functions, it is enough to assume that f is a characteristic function.

Let thus $f=\chi_E$, where $E\subset(0,2\pi)$ is a measurable set. Since for every $\delta>0$ there exists a compact $K\subset E$ such that $|E\setminus K|<\delta$, it is enough to consider $f=\chi_K$, with $K\subset(0,2\pi)$ compact.

Let's recall the following general result.

Urysohn's lemma. *Let $K\subset\mathbb{R}^n$ be a compact set and let $V\subset\mathbb{R}^n$ be an open set such that $K\subset V$. There exists $f\in\mathcal{C}_c(\mathbb{R}^n)$ such that $\chi_K\le f\le\chi_V$.*

By this result applied to K and to any open set V such that $K\subset V$ with $|V\setminus K|<\epsilon$, there exists $h\in\mathcal{C}_c[0,2\pi]$ such that $\chi_K\le h\le\chi_V$.

Then

$$\|\chi_K - h\|_2^2 = \frac{1}{2\pi}\int_0^{2\pi} |\chi_K(t) - h(t)|^2 dt \leq \int_{V\setminus K} |h(t)|^2 dt \leq |V \setminus K| < \epsilon,$$

as desired.

Proof of Urysohn's lemma. Let us see first the following elementary result.

Lemma 4. *Let $K \subset \mathbb{R}^n$ be compact and let $V \subset \mathbb{R}^n$ be an open set such that $K \subset V$. Then there exists an open set U such that $K \subset U \subset \overline{U} \subset V$.*

Proof. Consider the closed set $V^c = \mathbb{R}^n \setminus V$ and consider $d(x, V^c) = \inf_{y\in V^c} \|x - y\|$. Since V^c is closed $d(x, V^c) = 0$ if and only if $x \in V^c$. In particular, by the hypothesis, for any $x \in K$ there exists an $\epsilon(x) > 0$ such that $d(x, V^c) \geq \epsilon(x)$; in particular $B\big(x, \epsilon(x)\big) \subset V$.

From the covering $K \subset \cup_{x\in K} B\big(x, \epsilon(x)\big)$ extract a finite subcovering $K \subset \cup_{i=1}^N B\big(x_i, \epsilon_i\big)$, where $\epsilon_i = \epsilon(x_i)$, and define

$$U = \cup_{i=1}^N B\big(x_i, \epsilon_i\big),$$

which satisfies the required conditions. □

Applying this lemma twice we obtain open sets V_0, V_1, such that

$$K \subset V_1 \subset \overline{V_1} \subset V_0 \subset \overline{V_0} \subset V.$$

Let $r_0 = 0$, $r_1 = 1$ and $\{r_n\}_{n\geq 3}$ be an enumertation of $\mathbb{Q} \cap (0, 1)$. Let us construct open sets V_{r_n} recursivelly, in the following way. Suppose that $n \geq 2$ and that $V_{r_0}, \ldots, V_{r_n}$ have already been chosen in a way that $r_i < r_j$ implies $\overline{V_{r_j}} \subset V_{r_i}$. Define

$$r_i = \max_{k=1,\ldots,n} \{r_k : r_k < r_{n+1}\}, \quad r_j = \min_{k=1,\ldots,n} \{r_k : r_k > r_{n+1}\}.$$

By Lemma 4 there exists an open set $V_{r_{n+1}}$ such that

$$\overline{V_{r_j}} \subset V_{r_{n+1}} \subset \overline{V_{r_{n+1}}} \subset V_{r_i}.$$

This produces a nested collection of open sets $\{V_r\}_{r\in\mathbb{Q}\cap[0,1]}$ such that $K \subset V_1$, $\overline{V_0} \subset V$, each $\overline{V_r}$ is compact and for $s > r$ the inclusion $\overline{V_s} \subset V_r$ holds.

Define the functions

$$f(x) = \sup_r f_r(x), \qquad g(x) = \inf_s g_s(x),$$

where for $r, s \in \mathbb{Q} \cap [0, 1]$

$$f_r(x) = r\chi_{V_r} = \begin{cases} 1 & \text{if } x \in V \\ 0 & \text{if } x \notin V \end{cases} \qquad \text{and} \qquad g_s(x) = \begin{cases} 1 & \text{if } x \in \overline{V_s} \\ s & \text{if } x \notin \overline{V_s}. \end{cases}$$

By construction $0 \leq f \leq 1$ and $f(x) = 1$ for all $x \in K$. Also $\operatorname{supp} f \subset \overline{V_0} \subset V$.

We shall be done as soon as we show that f is continuous. To see this we prove first that $f(x) = g(x)$ for all $x \in \mathbb{R}^n$.

If $f(x) > g(x)$ for some $x \in \mathbb{R}^n$ there exist $r, s \in \mathbb{Q} \cap [0,1]$ such that $f_r(x) > g_s(x)$. This is possible only if $r > s$, $x \in V_r$ and $x \notin \overline{V_s}$. But $r > s$ implies $V_r \subset V_s$, by construction.

Similarly, if $f(x) < g(x)$ for some $x \in \mathbb{R}^n$ take $r, s \in \mathbb{Q} \cap [0,1]$ such that $f(x) < r < s < g(x)$. Since $f(x) < r$ we have $x \notin \overline{V_r}$, and since $g(x) > s$ we also have $x \in V_s$. This contradicts the inclusion $V_s \subset V_r$.

Finally, in order to see that $f = g$ is continuous let us see that for any $a < b$ the set $f^{-1}(a,b)$ is open. Observe that

$$f^{-1}(a,b) = \{x \in \mathbb{R}^n : f(x) > a\} \cap \{x \in \mathbb{R}^n : g(x) < b\},$$

so it is enough to see that these two sets are open.

On the one hand

$$\{x \in \mathbb{R}^n : f(x) > a\} = \cup_r \{x \in \mathbb{R}^n : f_r(x) > a\}.$$

Since

$$\{x \in \mathbb{R}^n : f_r(x) > a\} = \begin{cases} \mathbb{R}^n & \text{if } a < 0 \\ V_r & \text{if } 0 \leq a < r \\ \emptyset & \text{if } r \leq a \end{cases}$$

are all open sets, also is $\{x \in \mathbb{R}^n : f(x) > a\}$.

Similarly

$$\{x \in \mathbb{R}^n : g(x) < b\} = \cup_s \{x \in \mathbb{R}^n : g_s(x) < b\}$$

and

$$\{x \in \mathbb{R}^n : g(x) < b\} = \begin{cases} \emptyset & \text{if } b \leq s \\ \mathbb{R}^n \setminus \overline{V_s} & \text{if } s < b \leq 1 \\ \mathbb{R}^n & \text{if } b > 1. \end{cases}$$

□