The nitty gritty of Fejer's Theorem

Peter Haggstrom
www.gotohaggstrom.com
mathsatbondibeach@gmail.com

June 19, 2016

1 Introduction

The subject of Fourier theory starts with the outrageous proposition that:

$$S_n(f,t) = \sum_{r=-n}^{n} \hat{f}(r)e^{irt}\,dt \to f(t) \text{ as } n \to \infty \tag{1}$$

where $\hat{f}(r)$ are the Fourier coefficients defined by:

$$\hat{f}(r) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-irt}\,dt \tag{2}$$

Note: In what follows the integrals will be on $[-\pi, \pi]$ but one can work more generally on the circle $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ ie the real line mod 2π.

Fourier's basic claim was that (1) worked for any function (although several of his contemporaries disputed that he had actually proved such a claim). Notwithstanding the doubts about convergence Fourier used the theory to great practical advantage in the theory of heat conduction. Dirichlet proved that if f is continuous and has a bounded continuous derivative except possibly at a finite number of points then (1) does hold. However Du Bois-Reymond constructed a continuous function whose partial sums as defined by (1) blew up at $t = 0$ ie $\lim_{n\to\infty} \sup S_n(f,0) = \infty$.

At the ripe old age of 19 Fejér proved that if f is a continuous function from $\mathbb{T}$ to $\mathbb{C}$ then if we are given the Fourier coefficients $\hat{f}(r)$ $(r \in \mathbb{Z})$ of f, we can find $f(t)$ for $t \in \mathbb{T}$. Fejér went on to become thesis supervisor for John von Neumann, Paul Erdös, and George Polya and several other influential Hungarian mathematicians. Fejér's theorem was a surprise because the standard approach to partial sums did not always give rise to well behaved results. In other words if $s_n = \sum_{k=1}^{n} a_k$ is the usual partial sum then the sequence $s_0, s_1, \dots$ may not be well behaved. Cesàro had already shown that averages of the form $s_0, \frac{s_0+s_1}{2}, \frac{s_0+s_1+s_2}{3}, \ldots$ have better behaviour than $s_0, s_1, \ldots$. To understand why see [1] for a detailed explanation.

In what follows I will follow the first two chapters of Tom Kӧrner's book on Fourier analysis (see [4]), however, I will expand steps that he compresses. As you will see there are one or two steps that are a bit like icebergs whose small tips hide a large mass of tedious calculation below the surface.

2 Lemma on Cesàro convergence of partial sums

The importance of the Cesàro limit is that if this limit exists it will equal the usual limit whenever that usual limit exists and may exist even if the usual limit does not exist. This is such an important result it needs to be proved.

In the context of Fourier theory it was Fejér who showed that although the partial sums $S_n(f,t) = \sum_{r=-n}^{n} \hat{f}(r)e^{irt}$ might fail to converge, the Cesàro sum might behave better: $C_n(f,t) = \frac{1}{n+1}\sum_{j=0}^{n} S_j(f,t)$.

Lemma: If $s_n \to s$ then $\frac{1}{n+1}\sum_{j=0}^{n} s_j \to s$

As usual, take $\epsilon > 0$ and note that because $s_n \to s$ we can find an $N(\epsilon)$ such that $|s_n - s| < \frac{\epsilon}{2}$ for $n \geq N(\epsilon)$. We write $N(\epsilon)$ to emphasise the dependence of N on ϵ. Now let $Q = \sum_{j=0}^{N(\epsilon)} |s_n - s|$. Then:

$$\begin{aligned}\Big|\frac{1}{n+1}\sum_{j=0}^{n} s_j - s\Big| &= \frac{1}{n+1}\Big|\sum_{j=0}^{n}(s_j - s)\Big| \leq \frac{1}{n+1}\sum_{j=0}^{n}|s_j - s| \\ &= \frac{1}{n+1}\Big(\sum_{j=0}^{N(\epsilon)}|s_j - s| + \sum_{j=N(\epsilon)+1}^{n}|s_j - s|\Big) < \frac{1}{n+1}\Big(Q + (n - N(\epsilon))\frac{\epsilon}{2}\Big)\end{aligned} \tag{3}$$

Now it is certainly true that $n - N(\epsilon) \leq n+1$ and if we choose $M(\epsilon) \geq N(\epsilon)$ such that $M(\epsilon) \geq \frac{2Q}{\epsilon}$ it follows that $Q \leq \frac{\epsilon M(\epsilon)}{2}$. Thus the last line of (3) is :

$$\frac{1}{n+1}\Big(Q + (n - N(\epsilon))\frac{\epsilon}{2}\Big) \leq \frac{1}{n+1}\Big((n+1)\frac{\epsilon}{2} + (n+1)\frac{\epsilon}{2}\Big) = \epsilon \tag{4}$$

This establishes that $\frac{1}{n+1}\sum_{j=0}^{n} s_j \to s$ as $n \to \infty$. For a slightly different proof see [2], page 30.

Thus, although the partial sums $S_n(f,t) = \sum_{r=-n}^{n} \hat{f}(r)e^{irt}$ might fail to converge, Fejér's insight was that their averages defined as follows might behave better. In essence the Cesàro limit supplants the usual limit.

$$\text{Cesàro average} = \sigma_n(f,t) = \frac{1}{n+1}\sum_{j=0}^{n} S_j(f,t) \tag{5}$$

By expanding (5) we get:

$$\begin{aligned}
\sigma_n(f,t) =& \frac{1}{n+1}\sum_{j=0}^{n} S_j(f,t) \\
=& \frac{1}{n+1}\Big(S_0(f,t) + S_1(f,t) + \cdots + S_n(f,t)\Big) \\
=& \frac{1}{n+1}\Big(\hat{f}(0) + \sum_{r=-1}^{n} \hat{f}(r)e^{irt} + \cdots + \sum_{r=-n}^{n} \hat{f}(r)e^{irt}\Big) \\
=& \frac{1}{n+1}\Big((n+1)\hat{f}(0) + n\hat{f}(1)e^{it} + (n-1)\hat{f}(2)e^{2it} + \cdots + 1\hat{f}(n)\, e^{nit} \\
& + n\hat{f}(-1)e^{-it} + (n-1)\hat{f}(-2)e^{-2it} + \cdots + 1\hat{f}(-n)\, e^{-nit}\Big) \\
=& \sum_{r=-n}^{n} \left(\frac{n+1-|r|}{n+1}\right) \hat{f}(r)\, e^{irt}
\end{aligned} \tag{6}$$

The representation $\sigma_n(f,t) = \sum_{r=-n}^{n} \left(\frac{n+1-|r|}{n+1}\right) \hat{f}(r)\, e^{irt}$ is obviously oscillatory but it is not immediately clear whether it oscillates between positive and negative values. As we shall see, one of the remarkable facts about the Fejér kernel is its positivity.

At this point it is worth noting the relationship between the Dirichlet kernel and the Fejér kernel. Full details of the Dirichlet kernel can be found in [2]], however, we can quickly recount the basic properties here. The Dirichlet kernel can be defined this way:

$$D_n(x) = \sum_{k=-n}^{n} e^{ikx} = 1 + 2\sum_{k=1}^{n}\cos(kx) = \frac{\sin[(n+\frac{1}{2})x]}{\sin(\frac{x}{2})} \tag{7}$$

The RHS of (7) is obtained as follows. First note that:

$$2\cos u\, \sin v = \sin(u+v) - \sin(u-v) \tag{8}$$

Then with $u = kx$ and $v = \frac{x}{2}$ we have:

$$2\cos(kx) = \frac{\sin[(k+\frac{1}{2})x] - \sin[(k-\frac{1}{2})x]}{\sin(\frac{x}{2})} \tag{9}$$

Hence:

$$
\begin{aligned}
2\sum_{k=1}^{n}\cos(kx) =& \sum_{k=1}^{n}\frac{\sin[(k+\frac{1}{2})x]-\sin[(k-\frac{1}{2})x]}{\sin(\frac{x}{2})}\\
=&\frac{\sin[(n+\frac{1}{2})x]-\sin(\frac{x}{2})}{\sin(\frac{x}{2})}\\
=&\frac{\sin[(n+\frac{1}{2})x]}{\sin(\frac{x}{2})}-1\\
\therefore \quad D_n(x) = 1+2\sum_{k=1}^{n}\cos(kx) =& \frac{\sin[(n+\frac{1}{2})x]}{\sin(\frac{x}{2})}
\end{aligned}
\tag{10}
$$

With this definition the Fejér kernel is related to the Dirichet kernel as follows:

$$
\begin{aligned}
K_n(x) =& \frac{1}{n+1}\sum_{k=0}^{n}D_k(x)\\
=&\frac{1}{n+1}\sum_{k=0}^{n}\sum_{j=-k}^{k}e^{ijx}\\
=&\frac{1}{n+1}\Big[\sum_{j=0}^{0}e^{ijx}+\sum_{j=-1}^{1}e^{ijx}+\sum_{j=-2}^{2}e^{ijx}+\cdots+\sum_{j=-n}^{n}e^{ijx}\Big]\\
=&\frac{1}{n+1}\sum_{k=-n}^{n}(n+1-|j|)\,e^{ijx}
\end{aligned}
\tag{11}
$$

3 Positivity of the Fejér kernel

The Dirichlet kernel oscillates between positive and negative values whereas the Fejér kernel is non-negative. At first blush it may seem unlikely that the Fejér kernel is non-negative. There are various ways of proving this result. Typically with complex exponentials one tries taking the real or imaginary part of the complex expression as appropriate or use a difference of sines or cosines and then use telescoping at some point. I will present both approaches below and also give a third proof based on applying a general result to the Fejér kernel.

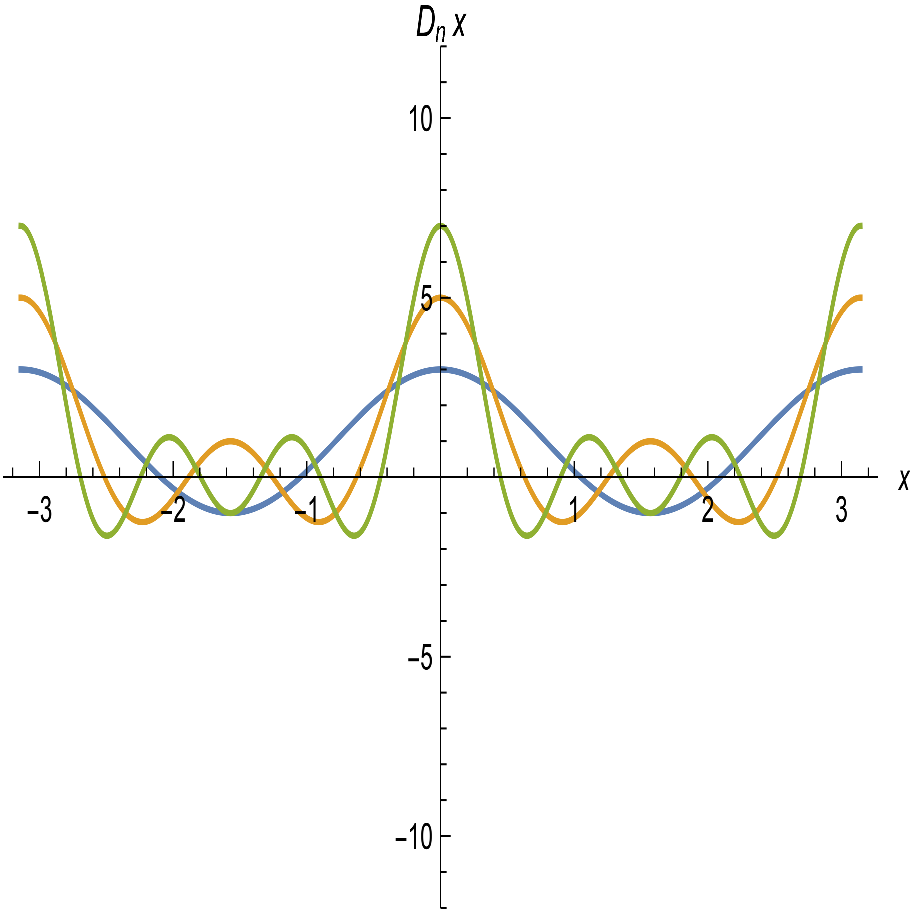
$$
\begin{aligned}
(n+1)K_n(x) =& \sum_{k=0}^{n} D_k(x) \\
=& \sum_{k=0}^{n} \frac{\sin[(k+\frac{1}{2})x]}{\sin(\frac{x}{2})} \\
=& \frac{1}{\sin(\frac{x}{2})} \mathfrak{Im}\Big\{ \sum_{k=0}^{n} e^{i(k+\frac{1}{2})x} \Big\} \\
=& \frac{1}{\sin(\frac{x}{2})} \mathfrak{Im}\Big\{ e^{\frac{ix}{2}} \sum_{k=0}^{n} e^{ikx} \Big\} \\
=& \frac{1}{\sin(\frac{x}{2})} \mathfrak{Im}\Big\{ e^{\frac{ix}{2}} \frac{\big(e^{i(n+1)x} - 1\big)}{e^{ix} - 1} \Big\} \\
=& \frac{1}{\sin(\frac{x}{2})} \mathfrak{Im}\left\{ e^{\frac{ix}{2}} \frac{\big(e^{i(n+1)x} - 1\big)}{e^{\frac{ix}{2}} \big(e^{\frac{ix}{2}} - e^{\frac{-ix}{2}}\big)} \right\} \\
=& \frac{1}{\sin(\frac{x}{2})} \mathfrak{Im}\left\{ \frac{\cos[(n+1)x] - 1 + i\sin[(n+1)x]}{2i\sin(\frac{x}{2})} \right\} \\
=& \frac{1 - \cos[(n+1)x]}{2\sin^2(\frac{x}{2})} \\
=& \frac{2\sin^2[(n+1)\frac{x}{2}]}{2\sin^2(\frac{x}{2})} \\
=& \frac{\sin^2[(n+1)\frac{x}{2}]}{\sin^2(\frac{x}{2})}
\end{aligned}
\tag{12}
$$

Therefore:

$$
K_n(x) = \frac{1}{n+1} \left(\frac{\sin(n+1)\frac{x}{2}}{\sin(\frac{x}{2})} \right)^2 \geq 0 \tag{13}
$$

$$
\begin{aligned}
K_n(0) =& \sum_{r=-n}^{n} \frac{n+1-|r|}{n+1} \\
=& \sum_{r=-n}^{n} 1 - \frac{1}{n+1} \sum_{r=-n}^{n} |r| \\
=& 2n + 1 - \frac{1}{n+1} \times 2 \times \frac{n}{2}(n+1) \\
=& 2n + 1 - n \\
=& n + 1
\end{aligned}
\tag{14}
$$

The Dirichlet kernels have the following form:



A more tedious derivation is as follows:

$$
\begin{aligned}
\sum_{k=-n}^{n} \frac{n+1-|k|}{n+1} e^{ikx} =& 1 + \sum_{k=1}^{n} \frac{n+1-|k|}{n+1} e^{-ikx} + \sum_{k=1}^{n} \frac{n+1-|k|}{n+1} e^{ikx} \\
=& 1 + \frac{1}{n+1} \sum_{k=1}^{n} (n+1-k)\, 2\cos(kx) \\
=& 1 + \frac{1}{n+1} \sum_{k=1}^{n} (n+1-k) \Big(\frac{\sin[(k+\frac{1}{2})x] - \sin[(k-\frac{1}{2})x]}{\sin(\frac{x}{2})} \Big) \\
=& 1 + \frac{1}{(n+1)\sin(\frac{x}{2})} \Bigg\{ n\big[\sin(\frac{3x}{2}) - \sin(\frac{x}{2})\big] + (n-1)\big[\sin(\frac{5x}{2}) - \sin(\frac{3x}{2})\big] \\
& (n-2)\big[\sin(\frac{7x}{2}) - \sin(\frac{5x}{2})\big] + \cdots + 3\big[\sin(n-\frac{3}{2})x - \sin(n-\frac{5}{2})x)\big] \\
& + 2\big[\sin(n-\frac{1}{2})x - \sin(n-\frac{3}{2})x)\big] + 1\big[\sin(n+\frac{1}{2})x - \sin(n-\frac{1}{2})x\big] \Bigg\} \\
=& 1 + \frac{1}{(n+1)\sin(\frac{x}{2})} \Big\{ -n\sin(\frac{x}{2}) + \sum_{k=1}^{n} \sin[(k+\frac{1}{2})x] \Big\} \\
=& 1 + \frac{1}{(n+1)\sin(\frac{x}{2})} \Big\{ -n\sin(\frac{x}{2}) + \sum_{k=1}^{n} \frac{\cos(kx) - \cos(k+1)x}{2\sin(\frac{x}{2})} \Big\} [*] \\
=& 1 - \frac{n}{n+1} + \frac{1}{2(n+1)\sin^2(\frac{x}{2})} \Big[\cos x - \cos(n+1)x \Big] \\
=& \frac{2\sin^2(\frac{x}{2}) + \cos x - \cos(n+1)x}{2(n+1)\sin^2(\frac{x}{2})} \\
=& \frac{1 - \cos x + \cos x - \cos(n+1)x}{2(n+1)\sin^2(\frac{x}{2})} \\
=& \frac{1 - \cos(n+1)x}{2(n+1)\sin^2(\frac{x}{2})} \\
=& \frac{1 - \big(1 - 2\sin^2(\frac{(n+1)x}{2})\big)}{2(n+1)\sin^2(\frac{x}{2})} \\
=& \frac{1}{n+1} \left(\frac{\sin\big(\frac{(n+1)x}{2}\big)}{\sin(\frac{x}{2})} \right)^2
\end{aligned}
\tag{15}
$$

In the step [∗] the following manipulations were employed:

$$
\begin{aligned}
\cos(k+1)x &= \cos(k+\frac{1}{2}+\frac{1}{2})x \\
&= \cos(k+\frac{1}{2})x\,\cos\frac{x}{2} - \sin(k+\frac{1}{2})x\,\sin\frac{x}{2} \\
\therefore \sin(k+\frac{1}{2})x\,\sin\frac{x}{2} &= \cos(k+\frac{1}{2})x\,\cos\frac{x}{2} - \cos(k+1)x
\end{aligned} \tag{16}
$$

Also,

$$
\begin{aligned}
\cos kx &= \cos(k+\frac{1}{2}-\frac{1}{2})x \\
&= \cos(k+\frac{1}{2})x\,\cos\frac{x}{2} + \sin(k+\frac{1}{2})x\,\sin\frac{x}{2} \\
\therefore \sin(k+\frac{1}{2})x\,\sin\frac{x}{2} &= \cos kx - \cos(k+\frac{1}{2})x\,\cos\frac{x}{2}
\end{aligned} \tag{17}
$$

Adding (16) and (17) we then have:

$$
\sin(k+\frac{1}{2})x = \frac{\cos kx - \cos(k+1)x}{2\sin\frac{x}{2}} \tag{18}
$$

In his textbook Körner proves ([4], page 7) the positivity of the Fejèr kernel by a third method that is more indirect than the two methods already given. He starts the derivation with the following proposition:

$$
\sum_{r=-n}^{n}(n+1-|r|)e^{irx} = \Big(\sum_{k=0}^{n} e^{i(k-\frac{n}{2})x}\Big)^2 \tag{19}
$$

That this is so is not immediately obvious and, typically, he gives no hint. It is undoubtedly one of the many little mathematical 'engines' that Cambridge Tripos students (Körner 's audience) would be familiar with. Let us consider the following polynomial:

$$
p_n(x) = \Big(\sum_{k=0}^{n} x^k\Big)^2 \tag{20}
$$

We can play with some low values of n to get the following format which looks like Pascal's Triangle:

$$
\begin{aligned}
(x^0)^2 &= \boxed{1} \\
(1+x)^2 &= \boxed{1} + \boxed{2}x + \boxed{1}x^2 \\
(1+x+x^2)^2 &= \boxed{1} + \boxed{2}x + \boxed{3}x^2 + \boxed{2}x^3 + \boxed{1}x^4 \\
(1+x+x^2+x^3)^2 &= \boxed{1} + \boxed{2}x + \boxed{3}x^2 + \boxed{4}x^3 + \boxed{3}x^4 + \boxed{2}x^5 + \boxed{1}x^6
\end{aligned} \tag{21}
$$

Thus we guess that for $n \geq 1$:

$$
p_n(x) = \sum_{k=0}^{n} (k+1)x^k + \sum_{k=n+1}^{2n} (2n-k+1)x^k \tag{22}
$$

This is proved inductively as follows. The proposition is true for $n = 1$:

$p_1(x) = 1 + 2x + x^2$ while $\sum_{k=0}^{1}(k+1)x^k + \sum_{k=2}^{2}(2-k+1)x^k = 1 + 2x + x^2$. Assuming the proposition is true for any n we have:

$$
\begin{aligned}
p_{n+1}(x) =& \Big(\sum_{k=0}^{n+1} x^k\Big)^2 \\
=& \Big(\sum_{k=0}^{n} x^k + x^{n+1}\Big)^2 \\
=& \Big(\sum_{k=0}^{n} x^k\Big)^2 + 2x^{n+1} \sum_{k=0}^{n} x^k + x^{2n+2} \\
=& \underbrace{\sum_{k=0}^{n}(k+1)x^k + \sum_{k=n+1}^{2n} (2n-k+1)x^k}_{\text{using the induction hypothesis}} + 2 \sum_{k=0}^{n} x^{k+n+1} + x^{2n+2} \\
=& \sum_{k=0}^{n}(k+1)x^k + \sum_{k=n+2}^{2n+2} (2n-k+1)x^k + nx^{n+1} + x^{2n+2} + 2x^{n+1} + 2 \sum_{k=1}^{n} x^{k+n+1} + x^{2n+2} \\
=& \sum_{k=0}^{n+1}(k+1)x^k + \sum_{k=n+2}^{2n+2} (2n-k+1)x^k + 2 \sum_{k=1}^{n} x^{k+n+1} + 2x^{2n+2} \\
=& \sum_{k=0}^{n+1}(k+1)x^k + \sum_{k=n+2}^{2n+2} (2n-k+1)x^k + 2 \underbrace{\sum_{k=n+2}^{2n+1}}_{k+n+1\to k} x^k + 2x^{2n+2} \\
=& \sum_{k=0}^{n+1}(k+1)x^k + \sum_{k=n+2}^{2n+2} (2n-k+1)x^k + 2 \sum_{k=n+2}^{2n+2} x^k \\
=& \sum_{k=0}^{n+1}(k+1)x^k + \sum_{k=n+2}^{2n+2} [2(n+1)-k+1]x^k
\end{aligned}
\tag{23}
$$

Hence the proposition is true for $n+1$. We can now apply this general result to $\left(\sum_{k=0}^{n} e^{i(k-\frac{n}{2})x}\right)^2$:

$$
\begin{aligned}
\Big(\sum_{k=0}^{n} e^{i(k-\frac{n}{2})x}\Big)^2 =& e^{-inx}\Big(\sum_{k=0}^{n} e^{ikx}\Big)^2 \\
=& e^{-inx}\Big(\sum_{k=0}^{n}(k+1)e^{ikx} + \sum_{k=n+1}^{2n}[2n-k+1]e^{ikx}\Big) \\
=& \underbrace{\sum_{k=0}^{n}(k+1)e^{i(k-n)x}}_{\text{low}} + \underbrace{\sum_{k=n+1}^{2n}[2n-k+1]e^{i(k-n)x}}_{\text{high}}
\end{aligned}
\tag{24}
$$

The coefficient of e^{irx} for $1 \leq r \leq n$ comes from the "high" representation in (24) so that $k-n=r$. Hence the coefficient is $2n-k+1 = 2n-(r+n)+1 = n+1-r$. Similarly the coefficient of e^{irx} for $-n \leq r \leq -1$ comes from the "low representation" so that $k-n=r$ and so $k+1 = n+r+1 = n+1-|r|$. When $r=0$, $k=n$ and the coefficient of e^0 is $n+1 = n+1-r$. So in all cases the coefficient of e^{irx} is of the form $n+1-|r|$.

Hence $\sum_{r=-n}^{n}(n+1-|r|)e^{irx} = \Big(\sum_{k=0}^{n} e^{i(k-\frac{n}{2})x}\Big)^2$. We can now derive the desired result as follows:

$$
\begin{aligned}
\sum_{r=-n}^{n}\frac{(n+1-|r|)}{n+1}e^{irx} =& \frac{1}{n+1}\Big(\sum_{k=0}^{n} e^{i(k-\frac{n}{2})x}\Big)^2 \\
=& \frac{1}{n+1}\Big(e^{-\frac{inx}{2}}\sum_{k=0}^{n} e^{ikx}\Big)^2 \\
=& \frac{1}{n+1}\Big(e^{-\frac{inx}{2}}\frac{(1-e^{i(n+1)x})}{1-e^{ix}}\Big)^2 \\
=& \frac{1}{n+1}\Big(e^{-\frac{inx}{2}}\frac{(e^{-\frac{i(n+1)x}{2}}e^{\frac{i(n+1)x}{2}} - e^{\frac{i(n+1)x}{2}}e^{\frac{i(n+1)x}{2}})}{(e^{\frac{ix}{2}}e^{-\frac{ix}{2}} - e^{\frac{ix}{2}}e^{\frac{ix}{2}})}\Big)^2 \\
=& \frac{1}{n+1}\left(e^{-\frac{inx}{2}}e^{\frac{i(n+1)x}{2}}\frac{(e^{-\frac{i(n+1)x}{2}} - e^{\frac{i(n+1)x}{2}})}{e^{\frac{ix}{2}}(e^{-\frac{ix}{2}} - e^{\frac{ix}{2}})}\right)^2 \\
=& \frac{1}{n+1}\left(\frac{e^{-\frac{i(n+1)x}{2}} - e^{\frac{i(n+1)x}{2}}}{e^{-\frac{ix}{2}} - e^{\frac{ix}{2}}}\right)^2 \\
=& \frac{1}{n+1}\left(\frac{-2i\sin(\frac{(n+1)x}{2})}{-2i\sin(\frac{x}{2})}\right)^2 \\
=& \frac{1}{n+1}\left(\frac{\sin(\frac{(n+1)x}{2})}{\sin(\frac{x}{2})}\right)^2
\end{aligned}
\tag{25}
$$

4 Fejér 's Theorem

(1) If $f : \mathbb{T} \to \mathbb{C}$ is Riemann integrable then, if f is continuous at t

$$\sigma_n(f,t) = \sum_{r=-n}^{n} \left(\frac{n+1-|r|}{n+1} \right) \hat{f}(r)\, e^{irt} \to f(t) \tag{26}$$

(2) If $f : \mathbb{T} \to \mathbb{C}$ is continuous then

$$\sigma_n(f,t) = \sum_{r=-n}^{n} \left(\frac{n+1-|r|}{n+1} \right) \hat{f}(r)\, e^{irt} \to f(t) \quad \text{uniformly on } \mathbb{T} \tag{27}$$

5 Proof of Fejér 's Theorem

In what follows we will work on $[-\pi, \pi]$ rather than $\mathbb{T}$. Given that f is Riemann integrable we have:

$$\begin{aligned}\sigma_n(f,t) = \sum_{r=-n}^{n} \left(\frac{n+1-|r|}{n+1} \right) \hat{f}(r)\, e^{irt} &= \sum_{r=-n}^{n} \left(\frac{n+1-|r|}{n+1} \right) \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} f(x) e^{-irx}\, dx \right) e^{irt} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \sum_{r=-n}^{n} \left(\frac{n+1-|r|}{n+1} \right) e^{ir(t-x)}\, dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) K_n(t-x)\, dx \end{aligned} \tag{28}$$

The Fejér kernel is:

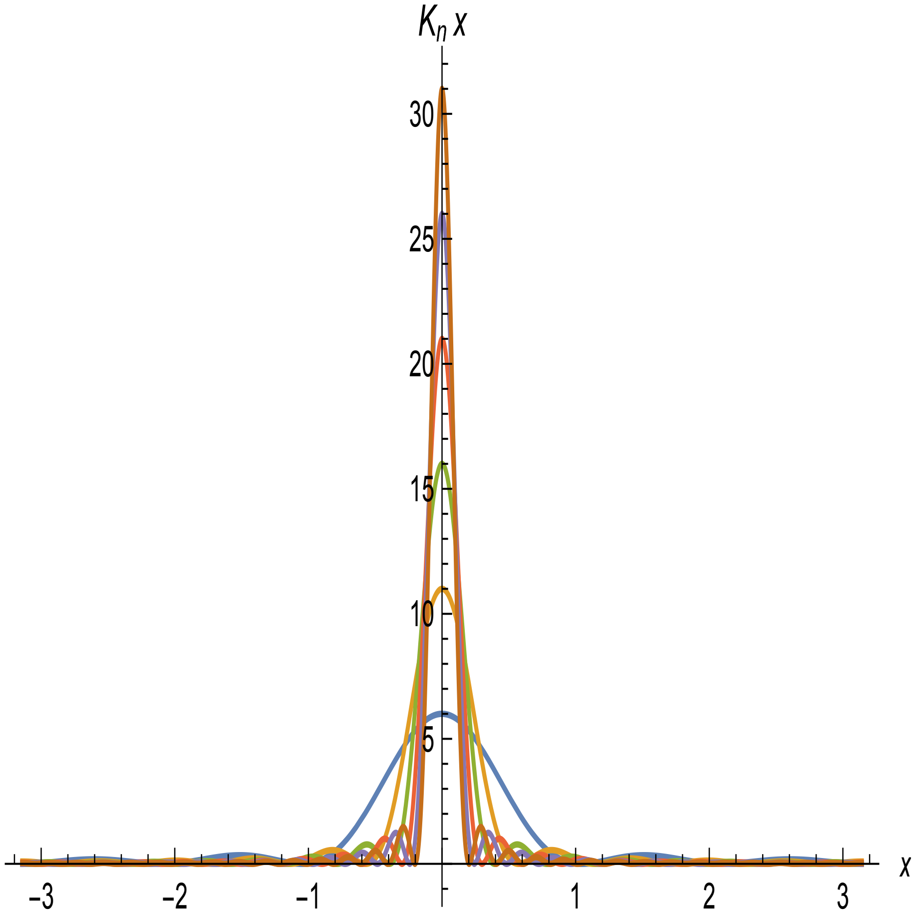
$$K_n(x) = \sum_{r=-n}^{n} \left(\frac{n+1-|r|}{n+1} \right) e^{irx} \tag{29}$$

Note that (28) is a convolution with the Fejér kernel and the function f.

By making the substitution $y = t - x$ we can show that:

$$\begin{aligned}\sigma_n(f,t) =&\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)K_n(t-x)\,dx \\ =&\frac{-1}{2\pi}\int_{t+\pi}^{t-\pi} f(t-y)K_n(y)\,dy \\ =&\frac{1}{2\pi}\int_{t-\pi}^{t+\pi} f(t-y)K_n(y)\,dy \\ =&\frac{1}{2\pi}\int_{\mathbb{T}} f(t-y)K_n(y)\,dy\end{aligned} \tag{30}$$

Hence as noted earlier we are justfiied in working with $[-\pi, \pi]$. The graphs of $K_5(x), K(x), K15(x), K_{20}(x)$, $K_{25}(x), K_{30}(x)$ are as follows.



As n increases the $K_n(x)$ peak more around 0 ie the product $K_n(x)f(x)$ tends to localise to $f(0)$. Visually, the decay of $K_n(x)$ for n large suggests that $\sigma_n(f,t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-y)K_n(y)\,dy \approx \frac{1}{2\pi}\int_{-\delta}^{\delta} f(t-y)K_n(y)\,dy$ where δ is some small positive number. Because f is continuous in a neighbourhood of t and δ is small, f is essentially constant in $[-\delta, \delta]$ and so we can make the following estimate:

$$\begin{aligned}
\sigma_n(f,t) &\approx \frac{1}{2\pi}\int_{-\delta}^{\delta} f(t-y)K_n(y)\,dy \\
&= f(t)\left(\frac{1}{2\pi}\int_{-\delta}^{\delta} K_n(y)\,dy\right) \\
&\approx f(t)\underbrace{\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(y)\,dy\right)}_{=1} \\
\therefore \sigma_n(f,t) &\approx f(t)
\end{aligned} \tag{31}$$

In (28) the penultimate step is justified because most of the mass of $K_n(x)$ is concentrated in $[-\delta, \delta]$. The formal proof essentially replicates these steps with the usual deference to ϵ, δ, N.

5.1 Properties of a 'good' kernel

The Fejér kernel has three properties of a good kernel (see [5] pages 48-51 for a general discussion):

1.

$$K_n(x) \geq 0 \quad \text{for all } x \in [-\pi, \pi] \tag{32}$$

We have already established this much.

2.

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)\,dx = 1 \tag{33}$$

This was used in (32) and is proved as follows:

$$
\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)\,dx =& \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{r=-n}^{n}\frac{(n+1-|r|)}{n+1}\,e^{irx}\,dx \\
=& \sum_{r=-n}^{n}\frac{(n+1-|r|)}{n+1}\,\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{irx}\,dx \\
=& \underbrace{\frac{n+1}{n+1}\,\frac{1}{2\pi}\int_{-\pi}^{\pi} dx}_{=1} + \sum_{r=-n,\,n\neq 0}^{n}\frac{(n+1-|r|)}{n+1}\,\frac{1}{2\pi}\left[\frac{e^{irx}}{ir}\right]_{-\pi}^{\pi} \\
=& 1 + \underbrace{\sum_{r=-n,\,n\neq 0}^{n}\frac{(n+1-|r|)}{n+1}\,\frac{1}{2\pi}\times\frac{2i\sin\pi r}{ir}}_{=0} \\
=& 1
\end{aligned}
\tag{34}
$$

Following ([5], page 48) the more general property of a 'good' kernel is that:

$$
\exists M > 0 \quad \text{such that for all } n \geq 1\ ,\ \int_{-\pi}^{\pi} |K_n(x)|\,dx \leq M \tag{35}
$$

Because of the positivity of $K_n(x)$ and (29), (32) clearly holds. However, (34) is violated by the Dirichlet kernel since $\int_{-\pi}^{\pi}|D_n(x)|\,dx \geq c\log N$ as $n \to \infty$ (see [2]).

3.

The third property of the Fejér kernel that we need is a uniform convergence property defined as follows:

$$
K_n(x) \to 0 \text{ uniformly outside } [-\delta, \delta] \text{ for all } \delta > 0 \tag{36}
$$

In a more general context, where the kernel can oscillate between positive and negative values we replace $K_n(x)$ by $|K_n(x)|$. Since $\delta \leq |x| \leq \pi$ we have that $\frac{\delta}{2} \leq \frac{|x|}{2} \leq \frac{\pi}{2}$ and so $\sin\frac{\delta}{2} \leq \sin\frac{x}{2}$ (draw a diagram to convince yourself) . Thus we have the following estimate:

$$
\begin{aligned}
K_n(x) =& \frac{1}{n+1}\left(\frac{\sin\left(\frac{(n+1)x}{2}\right)}{\sin(\frac{x}{2})}\right)^2 \\
\leq& \frac{1}{n+1}\left(\frac{1}{\sin(\frac{x}{2})}\right)^2 \\
\leq& \frac{1}{n+1}\left(\frac{1}{\sin(\frac{\delta}{2})}\right)^2 \to 0 \text{ a } n \to \infty
\end{aligned} \tag{37}
$$

With these prelimiinaries we can now prove Fejér's main result which is that if f is Riemann integrable on $[-\pi, \pi]$ and continuous at $t \in [-\pi, \pi]$ then $\sigma(f, t) \to f(t)$ as $n \to \infty$. As a general rule, the way to approaach such proofs is to construct estimates of the "hump" and the "tails" ie $\int_{-\delta}^{\delta} G(x)dx + \int_{x \notin [-\delta, \delta]} G(x)dx$. Usually uniform continuity/convergence and continuity are used to show that the second integral (the "tails") is small while continuity and boundedness figure in showing that the first integral (the "hump") is small.

Since f is Riemann integrable on $[-\pi, \pi]$ it must be bounded:

$$\exists M > 0 \text{ such that } |f(x)| \leq M \text{ for all } x \in [-\pi, \pi] \tag{38}$$

Continuity of f at t means that for any $\epsilon > 0$ we can find a $\delta > 0$ (which depends on t and ϵ) such that:

$$|f(x) - f(t)| \leq \frac{\epsilon}{2} \text{ for } |x - t| < \delta \tag{39}$$

The uniform convergence property of $K_n(x)$ ensures that we can find an N such that:

$$|K_n(x)| \leq \frac{\epsilon}{4M} \text{ for all } x \in [-\delta, \delta] \text{ and all } n \geq N. \tag{40}$$

Note that N will depend on t and ϵ

We also know that because of the positivity of $K_n(x)$:

$$\Big| \int_{-\delta}^{\delta} K_n(x)\, dx \Big| \leq \int_{-\delta}^{\delta} |K_n(x)|\, dx = \int_{-\delta}^{\delta} K_n(x)\, dx \leq \int_{-\pi}^{\pi} K_n(x)\, dx \tag{41}$$

The fact that $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)\, dx = 1$ is used below in the first and last lines.

Using the convolution formula for $K_n(x)$ we have:

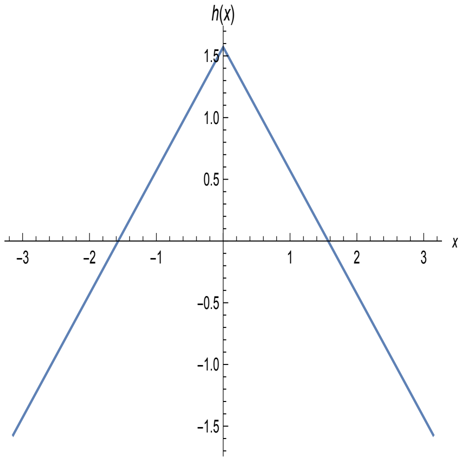
$$
\begin{aligned}
|\sigma_n(f,t)-f(t)| =& \Big|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-x)K_n(x)\,dx - \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)f(t)\,dx\Big| \\
=& \Big|\frac{1}{2\pi}\int_{-\pi}^{\pi}\Big(f(t-x)-f(t)\Big)K_n(x)\,dx\Big| \\
=& \underbrace{\Big|\frac{1}{2\pi}\int_{-\delta}^{\delta}\Big(f(t-x)-f(x)\Big)K_n(x)\,dx\Big|}_{\text{hump}} + \underbrace{\Big|\frac{1}{2\pi}\int_{x\notin[-\delta,\delta]}\Big(f(t-x)-f(x)\Big)K_n(x)\,dx\Big|}_{\text{tails}} \\
\leq & \underbrace{\frac{\epsilon}{2}\frac{1}{2\pi}\int_{-\delta}^{\delta}\Big|K_n(x)\Big|\,dx}_{\text{using continuity assumption}} + \underbrace{\frac{2M}{2\pi}\int_{x\notin[-\delta,\delta]}\Big|K_n(x)\Big|\,dx}_{|f(t-x)-f(t)|\leq|f(t-x)|+|f(t)|\leq 2M} \\
\leq & \frac{\epsilon}{2}\underbrace{\frac{1}{2\pi}\int_{-\delta}^{\delta}K_n(x)}_{\leq 1}\;dx + \frac{2M}{2\pi}\underbrace{\int_{x\notin[-\delta,\delta]}\frac{\epsilon}{4M}\,dx}_{\text{uniform convergence property}} \\
\leq & \frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon
\end{aligned} \tag{42}
$$

The final part of Fejér's theorem is to prove that if f is continuous on $[-\pi,\pi]$ then $\sigma_n(f,t)\to f(t)$ on $[-\pi,\pi]$. The only difference in the proof of this and the proof just given is that we move from a pointwise point of view to a global point of view. Because f is continuous on the compact interval $[-\pi,\pi]$ it is uniformly continuous on that interval. This means that where in the earlier proof the choices of δ and N were dependent on both t and ϵ, but because of the uniform continuity the dependence on t no longer applies - δ and N only depend on ϵ.

6 Rates of convergence

A classic beginner problem in Fourier theory is to calcuate the Fourier series for a plucked string. The plucked string can be represented as follows:

$$
h(x)=\frac{\pi}{2}-|x| \quad 0\leq|x|\leq\pi \tag{43}
$$



The Fourier coefficients are given by:

$$\begin{aligned}\hat{h}(r) =&\frac{1}{2\pi}\int_{-\pi}^{\pi} h(t)e^{-irt}\,dt\\ =&\frac{1}{2\pi}\Big\{\int_{-\pi}^{0} h(t)e^{-irt}\,dt+\int_{0}^{\pi} h(t)e^{-irt}\,dt\Big\}\\ =&\frac{1}{2\pi}\int_{0}^{\pi} h(t)(e^{irt}+e^{-irt})\,dt\\ =&\frac{1}{\pi}\int_{0}^{\pi} h(t)\cos rt\,dt\\ =&\frac{1}{\pi}\int_{0}^{\pi}(\frac{\pi}{2}-t)\cos rt\,dt\\ =&\frac{1}{\pi}\Big[(\frac{\pi}{2}-t)\frac{\sin rt}{r}\Big]_0^{\pi}+\frac{1}{\pi}\int_0^{\pi}\frac{\sin rt}{r}\,dt\\ =&\frac{1}{\pi r}\int_0^{\pi}\sin rt\,dt\\ =&\frac{1}{\pi r^2}\Big[-\cos rt\Big]_0^{\pi}\\ =&\frac{1}{\pi r^2}[-\cos r\pi+1]=\begin{cases}0 & \text{if r is even but} \neq 0\\ \frac{2}{\pi r^2} & \text{if r is odd}\end{cases}\end{aligned} \tag{44}$$

Note that $\hat{h}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} h(t)\,dt = 0$ due to the fact that $h(t)$ is even. Clearly then $\sum_{r=-n}^{n}|\hat{h}(r)|$ converges and it is a fundamental result of Fourier theory (see [4] page 32) that $S_n(x) \to h(x)$ uniformly on $[-\pi, \pi]$.

If we approximate $h(x)$ by $S_n(h, x) = \sum_{r=-n}^{n} \hat{h}(r)\, e^{irt}\,dt$ (see (1)), how good is that approximation? If, instead, we use the Cesàro sum, is the approximation better or worse? Chapter 10 of Körner's book [4] gives the detailed derivations surrounding rates of convergence of $S_n(h, x)$. He shows that:

$$
\begin{aligned}
|h(x) - S_n(h,x)| &\le \frac{2}{\pi(n-1)} \\
h(0) - S_n(h,0) &\ge \frac{2}{\pi(n+2)}
\end{aligned} \tag{45}
$$

The moral of the story is that while it may be possible to get reasonable accuracy for low values of n, to get very high orders of accuracy can involve hundreds of thousands of terms in the series since the convergence can be slow for many functions. In addition, as Körner shows in Chapter 33 of his book ([4], page 37) "no trigonometrical polynomial of degree n, of any form, can be a very good uniform approximation to h".

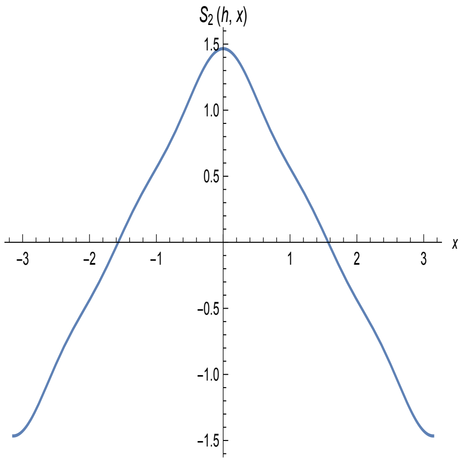
We know from Fejér's theorem that the Cesàro sum converges uniformly on $[-\pi, \pi]$ ie $\sigma_n(h,x) \to h(x)$ and it is useful to compare the convergence rates of the two forms of summation. The usual sum is given by:

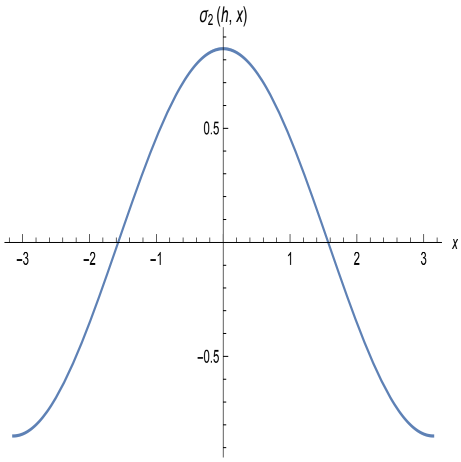
$$
\begin{aligned}
S_n(h,x) &= \sum_{r=-n}^{n} \hat{h}(r)\, e^{irx} \\
&= \sum_{r=-n}^{-1} \hat{h}(r)\, e^{irx} + \sum_{r=1}^{n} \hat{h}(r)\, e^{irx} \\
&= \sum_{r=1}^{n} \hat{h}(-r)\, e^{-irx} + \sum_{r=1}^{n} \hat{h}(r)\, e^{irx} \\
&= \sum_{k=0}^{n} \frac{2}{\pi(2k+1)^2}\, e^{-i(2k+1)x} + \sum_{k=0}^{n} \frac{2}{\pi(2k+1)^2}\, e^{i(2k+1)x} \\
&= \sum_{k=0}^{n} \frac{4\cos(2k+1)x}{\pi(2k+1)^2}
\end{aligned} \tag{46}
$$

The Cesàro sum is given by (see (27)):

$$
\begin{aligned}
\sigma_n(h,x) =& \frac{1}{2\pi} \int_{-\pi}^{\pi} h(t) K_n(x-t)\, dt \\
=& \frac{1}{2\pi} \int_{-\pi}^{\pi} \big(\frac{\pi}{2} - |t|\big)\, \frac{1}{n+1} \left(\frac{\sin(n+1)\frac{(x-t)}{2}}{\sin(\frac{x-t}{2})} \right)^2 dt
\end{aligned} \tag{47}
$$

Using *Mathematica* we can graph both $S_n(h,x)$ and $\sigma_n(h,x)$ for various values of n on $[-\pi, \pi]$ and compare with $h(x)$. For the low value of $n = 2$ the two graphs look like this:





Clearly $\sigma_2(h, x)$ is a poorer approximation at this low order of n. How poor can be seen from the following graphs which show the absolute value of the deviations between the sums and $h(x)$. Thus:

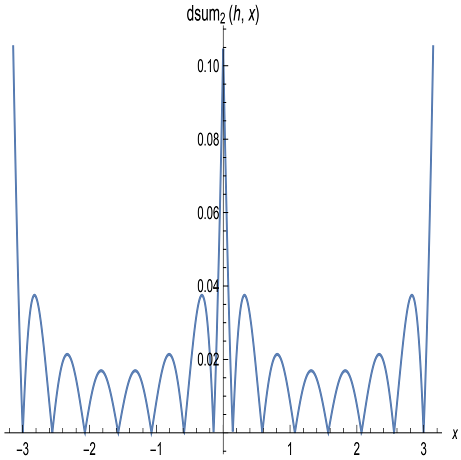
$$dsum_n(h, x) = \left| \frac{\pi}{2} - |x| - \sum_{k=0}^{n} \frac{4\cos(2k+1)x}{\pi(2k+1)^2} \right| \tag{48}$$

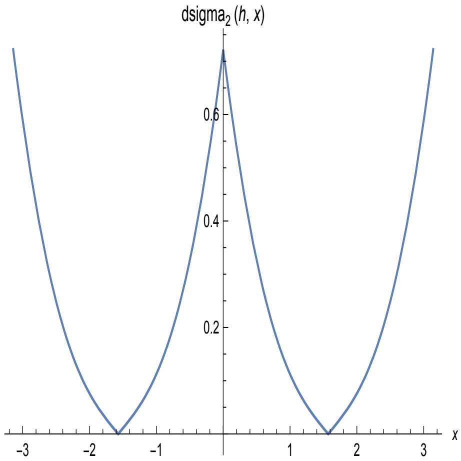
and

$$dsigma_n(h, x) \;\; = \left| \frac{\pi}{2} - |x| - \sigma_n(h, x) \right| \tag{49}$$

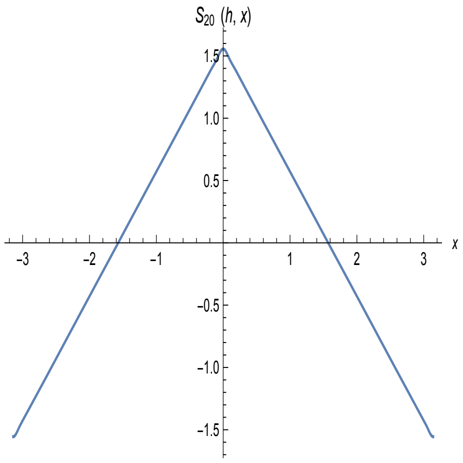
The corresponding deviation graphs for $n = 2$ are as follows:

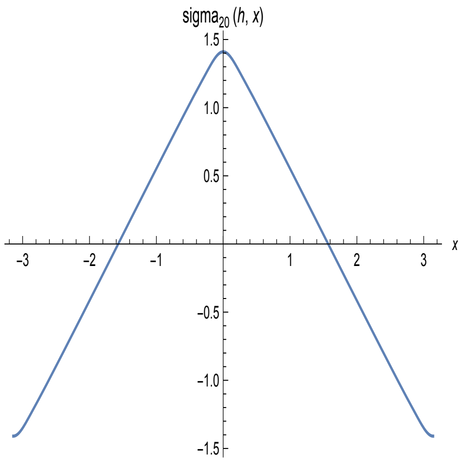
It is clear that the Cesàro sum is converging more slowly.

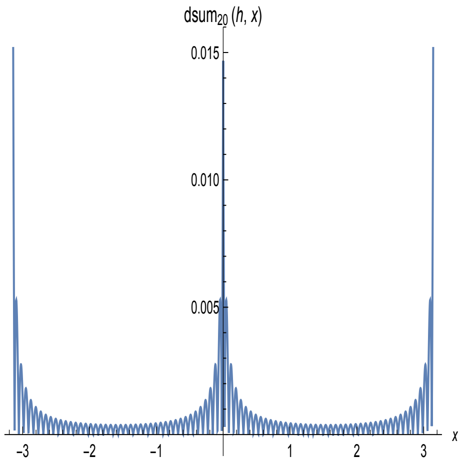

dsum$_2$ (h, x)
0.10
0.08
0.06
0.04
0.02
x
−3
−2
−1
1
2
3

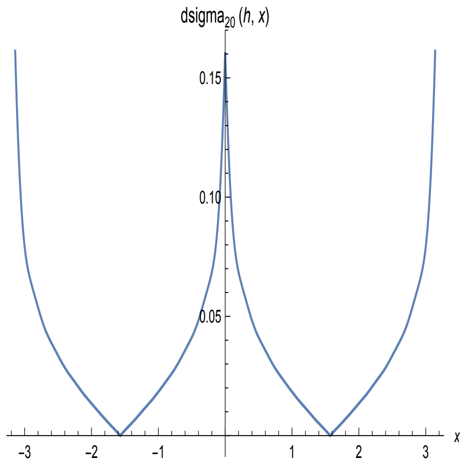

dsigma$_2$ (h, x)
0.6
0.4
0.2
x
−3
−2
−1
1
2
3

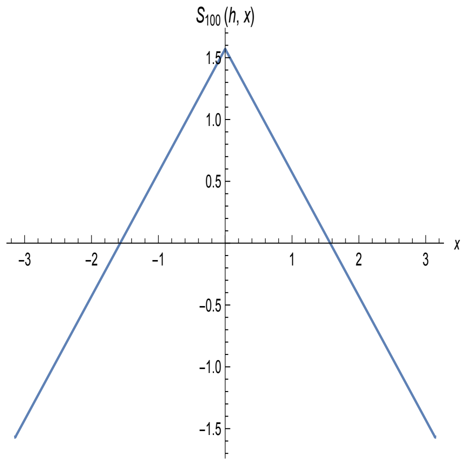
For $n = 20$ and $n = 100$ the comparative graphs are as follows:

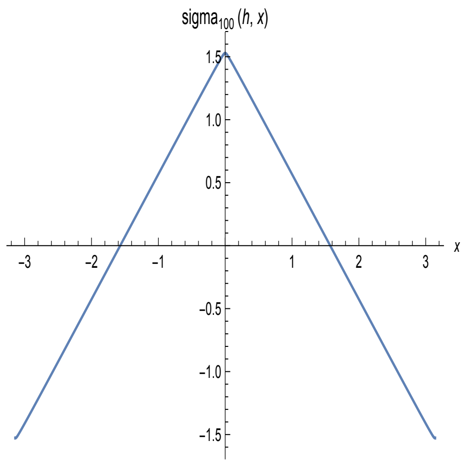


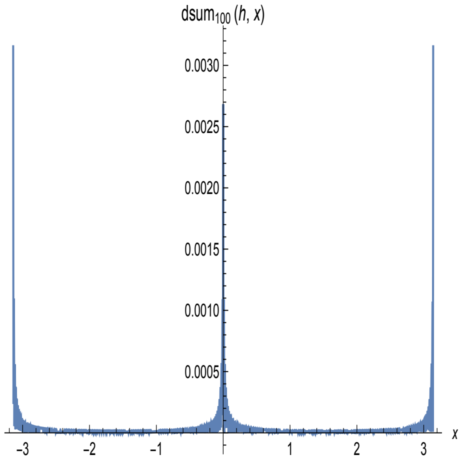


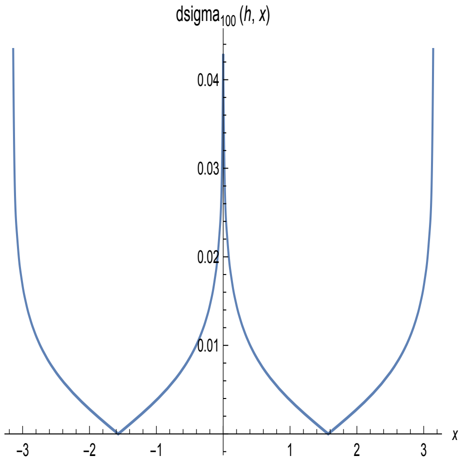



dsigma20 (h, x)
0.15
0.10
0.05
x
−3
−2
−1
1
2
3


S100 (h, x)
1.5
1.0
0.5
−0.5
−1.0
−1.5
x
−3
−2
−1
1
2
3


sigma100 (h, x)
1.5
1.0
0.5
−0.5
−1.0
−1.5
x
−3
−2
−1
1
2
3


dsum$_{100}$(h, x)
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
x
−3
−2
−1
1
2
3


dsigma$_{100}$(h, x)
0.04
0.03
0.02
0.01
x
−3
−2
−1
1
2
3

7 References

[1] Peter Haggstrom *The basics of Cesàro summability* : http://www.gotohaggstrom.com/The%20basics%20of%20Cesaro%20summability.pdf

[2] Peter Haggstrom *The good, the bad and the ugly of kernels: why the Dirichlet kernel is not a good kernel* :http://www.gotohaggstrom.com/The%20good,%20the%20bad,%20and%20the%20ugly%20of%20kernels.pdf

[3] Peter Haggstrom *Basic Fourier integrals*:http://www.gotohaggstrom.com/Basic%20Fourier%20integrals.pdf

[4] T W Körner, *Fourier Analysis*, Cambridge University Press, 1990.

[5] Elias M Stein and Rami Shakarchi, *Fourier Analysis - An Introduction*, Princeton University Press, 2003.

8 History

Created 19 June 2016