Contents

1 Convergence results for Fourier series

1.1 A historical perspective

Fourier series made their first appearance in mathematics through the way of physics, as is the norm with a lot of classical analysis. In particular, they turned up in the analysis of the following differential equation:

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0 \tag{1}$$

on the unit disc, with the following boundary conditions:

$$u(1, \theta) = f(\theta)$$

where f is an arbitrary continous function on $[-\pi, \pi]$ and $f(-\pi) = f(\pi)$.

Equation 1 is what is called the Laplace equation, which describes the temperature u on a conducting surface at steady state. This naturally raises the question: What would the temperature profile on a disk look like at steady state if the boundary was maintained at some temperature $f(\theta)$. That is precisely what the described differential equation along with the boundary conditions tries to answer.

Once one makes the simplifying assumption that the solution is separable, i.e.

$$u(r, \theta) = a(r) \cdot b(\theta)$$

equation 1 is easy to solve, and one sees that it has solutions of the form:

$$u_n(r, \theta) = r^n e^{in\theta}$$

for all integers n. And finite linear combinations of the solutions u_n are also solutions to the differential equation 1. But there is one problem. Not every continuous function f on $[-\pi, \pi]$ is a finite linear combination of $e^{in\theta}$, e.g. $|x|$. Leaving that problem aside for now, one notes that if f is of the form

$$f(\theta) = \sum_{n=-N}^{N} \widehat{f}(n) e^{in\theta}$$

then $\widehat{f}(n)$ is can be determined in the following manner:

$$\widehat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt \tag{2}$$

Here $\widehat{f}(n)$ is called the n^{th} Fourier coefficient of f. Since we assumed that f is a finite linear combination of $e^{-in\theta}$, that means it has only finitely many non-zero coefficients. Contrariwise, continuous functions that are not finte linear combinations of $e^{in\theta}$ must have infinitely many non-zero Fourier coefficients. Could it perhaps be that f is an *infinite* linear combination of $e^{in\theta}$? To be more precise, does the following sequence converge to f (in an appropriate sense) as N goes to ∞?

$$S_N(\theta) = \sum_{n=-N}^{N} \widehat{f}(n) e^{in\theta}$$

The best we can hope for is for the sequence S_N to converge uniformly to the function f, but as we shall see, that is not actually true, i.e. there are continuous functions whose Fourier series diverges for some point in the domain. The next step would be either weakening the mode of convergence, i.e. instead of expecting convergence in the L^∞ norm, one could expect convergence in weaker norms such as L^1 or L^2, or even Cesàro convergence. Another possible way to go about would be to strengthen the conditions on the function, i.e. forcing the function to be C^1, or even absolutely continuous ensures that the sequence S_N converges uniformly to f.

1.2 Fourier series of continuous functions are Cesàro summable

Before we go on to show that Fourier series converge under weaker notions of convergence, we'll develop a technique that can be used more generally to show certain sequences of functions converge uniformly. Furthermore, in all the sections that are to follow, we'll be working with continuous complex valued functions on the circle T, which is defined as the space $\mathbb{R}$ quotiented with the equivalence relation $\sim$ defined as

$$a \sim b \iff (a-b) = 2n\pi,\ n \in \mathbb{Z}$$

1.2.1 Convolution

For two continuous functions f and g on T, we can define a binary operation $*$:

$$(f*g)(x) = \int_{-\pi}^{\pi} f(x-t)g(t)dt$$

It follows that the operation is commutative (substitution of variables), associative (again substitution of variable), and distributive (integration is a linear operator). It also follows that $f*g$ is continuous since f is uniformly continuous.

It's not too hard to show that this operation has no identity element.

Proposition 1.1. *There exists no continuous function g on T such that for all continuous functions f*

$$f*g = f$$

Proof. We will prove the result by showing that if such a g existed, then $g(x) = 0$ for $x \neq 0$. But that would mean g is non-zero on a set of measure 0, hence $\int_T f(x-t)g(t)dt = 0$ for all f, which means g is not the identity, hence a contradiction.

Assume we have a continuous function g such that for all continuous f, $f*g = f$. Pick any non-zero x_0. We claim that $g(x_0)$ must be 0. If it's not 0, then without loss of generality, let $g(x_0) = \varepsilon > 0$. That means there exists a δ_0 such that for all $x \in (x_0 - \delta_0, x+\delta_0)$, $g(x) > \frac{\varepsilon}{2}$. Let $\delta = \min(\delta_0, |x_0|)$.

Now define a continuous function f in the following manner:

$$f(-x) = \begin{cases} 1 & x \in (x_0 - \frac{\delta}{2}, x_0 + \frac{\delta}{2}) \\ \frac{2(x-(x_0-\delta))}{\delta} & x \in \left(x_0 - \delta, x_0 - \frac{\delta}{2}\right] \\ \frac{2((x_0+\delta)-x)}{\delta} & x \in \left[x_0 + \frac{\delta}{2}, x_0 + \delta\right) \\ 0 & \text{otherwise} \end{cases}$$

It's clear that $f(0) = 0$.

$$\begin{aligned} f(0) &= (f * g)(0) \\ &= \int_{-\pi}^{\pi} f(-t)g(t)dt \\ &\geq \int_{x-\frac{\delta}{2}}^{x+\frac{\delta}{2}} f(-t)g(t)dt \\ &\geq \frac{\varepsilon\delta}{2} \\ &> 0 \end{aligned}$$

But $f(0) = 0$, and $0 \not> 0$. We have a contradiction, which means $g(x) = 0$ for all $x \neq 0$. This completes the proof. □

1.2.2 Dirac sequences

We saw that there is no continuous function which acts as an identity for $*$ operation. However, there do exist sequences of functions, such that the sequence $\{f * g_n\}$ converges uniformly to f as n goes to infinity. Such a sequence $\{g_n\}$ is called a Dirac sequence. The formal definition of a Dirac sequence is the following:

Definition 1.1. A sequence of a continuous functions $\{g_n\}$ is called a Dirac sequence if it satisfies the following conditions:

1. $g_n(x) \geq 0$ for all $n \in \mathbb{N}$ and all $x \in T$.

2. $g_n(x) = g_n(-x)$ for all $n \in \mathbb{N}$ and all $x \in T$.

3. $\displaystyle\int_{-\pi}^{\pi} g_n(t)dt = 1$ for all $n \in \mathbb{N}$.

4. For all $\varepsilon > 0$ and $\delta > 0$, there exists an N such that for all $n > N$,

$$\int_{-\pi}^{-\delta} g_n(t)dt + \int_{\delta}^{\pi} g_n(t)dt < \varepsilon$$

Following from the definition, we get this very useful theorem:

Theorem 1.2. *If $\{g_n\}$ is a Dirac sequence, then for all continuous functions f, the sequence $\{f * g_n\}$ converges uniformly to f.*

Proof. Pick any $\varepsilon > 0$. We need to show there exists an N such that for all $n > N$, $\|(f * g_n) - f\|_\infty < \varepsilon$. Pick a particular continuous function f. Let M be the maximum of $|f|$ on T. Pick $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that

$$\varepsilon_1\varepsilon_2 + \varepsilon_2 + 2M\varepsilon_1 < \varepsilon \tag{3}$$

Since f is uniformly continuous on T, pick a δ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon_2$.

Now pick an N such that for all $n > N$

$$\int_{-\pi}^{-\delta} g_n(t)dt + \int_{\delta}^{\pi} g_n(t)dt < \varepsilon_1$$

This would imply for all $n > N$

$$\int_{-\delta}^{\delta} g_n(t)dt > 1 - \varepsilon_1$$

Now consider $(f * g_n)(x_0)$ for some $x \in T$.

$$\begin{aligned}(f * g_n)(x_0) &= \int_{-\pi}^{\pi} f(x_0 - t)g_n(t)dt \\ &= \left(\int_{-\pi}^{-\delta} f(x_0 - t)g_n(t)dt + \int_{\delta}^{\pi} f(x_0 - t)g_n(t)dt\right) + \left(\int_{-\delta}^{\delta} f(x_0 - t)g_n(t)dt\right)\end{aligned}$$

Let's analyze the two terms separately. Since $-M \leq f(x_0 - t) \leq M$, we can bound the first term as

$$-M\varepsilon_1 \leq \int_{-\pi}^{-\delta} f(x_0 - t)g_n(t)dt + \int_{\delta}^{\pi} f(x_0 - t)g_n(t)dt \leq M\varepsilon_1$$

Similarly, for all x in the interval $(x_0 - \delta, x_0 + \delta)$

$$f(x_0) - \varepsilon_2 \leq f(x) \leq f(x_0) + \varepsilon_2$$

This lets us bound the second term in the following manner:

$$(1 - \varepsilon_1)(f(x_0) - \varepsilon_2) \leq \int_{-\delta}^{\delta} f(x_0 - t)g_n(t)dt \leq f(x_0) + \varepsilon_2$$

This gives us a complete bound on $(f * g_n)(x_0)$.

$$-M\varepsilon_1 + (1 - \varepsilon_1)(f(x_0) - \varepsilon_2) \leq (f * g_n)(x_0) \leq M\varepsilon_1 + (f(x_0) + \varepsilon_2)$$

Using the triangle inequality and the fact that $|f(x_0)| < M$, we get

$$f(x_0) - (\varepsilon_1\varepsilon_2 + \varepsilon_2 + 2M\varepsilon_1) \leq (f * g_n)(x_0) \leq f(x_0) + (\varepsilon_1\varepsilon_2 + \varepsilon_2 + 2M\varepsilon_1)$$

But from inequality 3, we get

$$f(x_0) - \varepsilon \leq (f * g_n)(x_0) \leq f(x_0) + \varepsilon$$

This shows that $\{f * g_n\}$ converges uniformly to f for all continuous f. $\square$

This is a useful result because this will let us deal with the question of convergence of the partial Fourier series of some continuous function f. One can write the n^{th} partial Fourier series as the convolution of f with some function g_n, called a kernel, and if one shows that the kernels g_n form a Dirac sequence, then the Fourier series also converges.

1.2.3 Cesàro summability

We mentioned that for general continuous functions, their Fourier series need not converge uniformly, or even pointwise, to the function; weakening the notion of convergence lets the conjecture go through. A weaker notion of convergence for infinite sums is the notion of Cesàro summability.

Definition 1.2. A sequence $\{x_n\}$ is said to be Cesàro summable if the following sequence $\{\sigma_n\}$ converges:

$$\sigma_n = \frac{\sum_{k=1}^n s_k}{n}$$

where s_k is the k^{th} partial sum of the sequence $\{x_n\}$. If the sequence $\{\sigma_n\}$ converges to L, then L is called the Cesàro sum of $\{x_n\}$.

It's easy to see that if the series $\displaystyle\sum_{k=1}^{n} x_n$ converges, then the Cesàro sums also converge to the same limit. To show it is a strictly weaker notion of convergence, consider the sequence $x_n = (-1)^n$. Clearly, the partial sums of x_n do not converge, but the Cesàro sums do converge to $\frac{1}{2}$.

1.2.4 The Fejér kernel

As outlined in a previous subsection, one would like to find functions g_N such that

$$(f * g_N)(\theta) = \sum_{n=-N}^{N} \widehat{f}(n)e^{in\theta}$$

where $\widehat{f}(n)$ is the n^{th} Fourier coefficient of f. Rewriting the above sum, we get

$$\sum_{n=-N}^{N} \widehat{f}(n)e^{in\theta} = \sum_{n=-N}^{N} \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt\right) e^{in\theta}$$

Since this is a finite sum, we can exchange the sum and the integral to get

$$\sum_{n=-N}^{N} \widehat{f}(n)e^{in\theta} = \int_{-\pi}^{\pi} \left(\frac{1}{2\pi}\sum_{n=-N}^{N} e^{in(\theta-t)}\right) f(t)dt$$

Summing up the geometric series, we get

$$\frac{1}{2\pi}\sum_{n=-N}^{N} e^{int} = \frac{1}{2\pi}\frac{\sin\left(Nt+\frac{t}{2}\right)}{\sin\left(\frac{t}{2}\right)}$$

The function g_N we wanted was is this function

$$g_N(t) = \frac{1}{2\pi}\frac{\sin\left(Nt+\frac{t}{2}\right)}{\sin\left(\frac{t}{2}\right)}$$

g_N certainly is even, and its integral over T is 1, but it is not non-negative everywhere, hence it fails to form a Dirac sequence. Let's look at the Cesàro partial sums instead, and concentrates on the kernel corresponding to those, which are called Fejér kernels. Clearly, the kernel corresponding to the N^{th} Cesàro sum would be the following:

$$F_N = \frac{\sum_{n=1}^{N} g_n}{N}$$

Summing up the geometric progression once again, we get the following closed form expression for F_N:

$$\begin{aligned} F_N(t) &= \frac{1}{2N\pi} \frac{\sin^2\left(\frac{Nt}{2}\right)}{\sin^2\left(\frac{t}{2}\right)} \\ &= \frac{1}{2N\pi} \frac{1-\cos(nt)}{1-\cos(t)} \end{aligned}$$

It follows that F_N is even and its integral over T is 1. Furthermore, it is non-negative, and finally we have the following inequality for $0 < \delta < |t| \leq \pi$

$$\begin{aligned} \frac{1}{2N\pi} \frac{1-\cos(nt)}{1-\cos(t)} &\leq \frac{1}{2N\pi} \frac{2}{1-\cos(t)} \\ &\leq \frac{1}{N\pi} \frac{1}{1-\cos(\delta)} \end{aligned}$$

This shows that $\{F_N\}$ satisfies all the conditions required to be a Dirac sequence. This easily leads to the following result.

Theorem 1.3 (Fejér's Theorem)**.** *The Fourier series of a continuous function is Cesàro summable, and the partial Cesàro sums converge uniformly to f.*

Proof. The Fejér kernels form a Dirac sequence. Now use the result of theorem 1.2. □

This result gives us two things: it tells us that the exponential polynomials (linear combinations of $e^{in\theta}$) are dense in $C(T)$, and for a given f in $C(T)$, it gives us an *explicit* sequence of polynomials which converge uniformly to f. The first result is not that impressive, since it also follows from Stone-Weierstrass theorem (exponential polynomials form a subalgebra of $C(T)$ containing a non-zero constant function and it separates points), but the Stone-Weierstrass theorem does not give an explicit sequence, which this result does.

1.3 Sufficient conditions for convergence of Fourier series

Now that we've established the Fourier series of continuous functions are Cesàro summable, which is a notion of convergence strictly weaker than uniform convergence, we'll proceed in the other direction, which involves strengthening conditions on the function f.

Clearly, if the Fourier coefficients of f are absolutely convergent, then that would mean the Fourier series also converges uniformly (triangle inequality). While this is rather strong condition, this does give us an idea to look at the size of Fourier coefficients in order to determine convergence. This leads us to a lemma:

Lemma 1.4. *If f is a continuous function, then*

$$\lim_{|n|\to\infty} \widehat{f}(n) = 0$$

Proof. First step in the proof will be showing that if f and g are two continuous functions such that $\|f-g\|_\infty < \varepsilon$, then $|\widehat{f-g}(n)| < 2\pi\varepsilon$. This follows from the definition of $\widehat{f-g}(n)$:

$$\begin{aligned}
|\widehat{f-g}(n)| &= \left|\int_{-\pi}^{\pi} (f-g)(t)e^{-int}dt\right| \\
&\le \int_{-\pi}^{\pi} |(f-g)(t)| \cdot |e^{-int}|dt \\
&< \varepsilon \int_{-\pi}^{\pi} |e^{-int}|dt \\
&= 2\pi\varepsilon
\end{aligned}$$

Pick an $\varepsilon > 0$, and using theorem 1.3, get an exponential polynomial g such that $\|f-g\|_\infty < \frac{\varepsilon}{2\pi}$. This means for all n, $|\widehat{f-g}(n)| < \varepsilon$. Let the degree of g be k. Then for $|n| > k$, $\widehat{g}(n) = 0$. That means for $|n| > k$, $|\widehat{f}(n)| = |\widehat{f-g}(n)| < \varepsilon$. This completes the proof. □

This lemma tells us the that the Fourier coefficient $\widehat{f}(n)$ is $o(1)$. This is a much weaker condition than the Fourier sums being absolutely convergent, since there exist $o(1)$ sequences whose sums are not absolutely convergent, or even convergent, e.g. the harmonic series.

Furthermore, we can show that Fourier coefficients of smoother functions decay faster.

Lemma 1.5. *If $f \in C^1(T)$, then $\widehat{f}(n)$ is in $o\left(\frac{1}{n}\right)$.*

Proof. Consider the Fourier coefficient of f', which is a continous function.

$$\begin{aligned}
\widehat{f'}(n) &= \int_{-\pi}^{\pi} f'(t)e^{-int}dt \\
&= \left[e^{-int}f(t)\right]_{-\pi}^{\pi} + in\int_{-\pi}^{\pi} f(t)e^{-int}dt && \text{(Integration by parts)} \\
&= in\widehat{f}(n)
\end{aligned}$$

This means

$$\left|\widehat{f}(n)\right| = \left|\frac{\widehat{f'}(n)}{n}\right|$$

Since $\widehat{f'}(n)$ is in $o(1)$, that means $\widehat{f}(n)$ is in $o\left(\frac{1}{n}\right)$. □

This leads to the following corollary.

Corollary 1.6. *If $f \in C^k(T)$, then $\widehat{f}(n)$ is in $o\left(\frac{1}{n^k}\right)$.*

From the above result, we get that if $f \in C^2(T)$, then $\widehat{f}(n)$ is in $o\left(\frac{1}{n^2}\right)$, which means the Fourier sums are absolutely convergent, and the Fourier series converges uniformly to f. We have the following theorem, as a consequence.

Theorem 1.7. *If $f \in C^2(T)$, then the Fourier series of f converges uniformly to f.*

Proof. $\widehat{f}(n)$ is in $o\left(\frac{1}{n^2}\right)$. $\sum \frac{1}{n^2}$ is absolutely convergent. Hence the Fourier series converges. □

The final result of this section will be that if f is continuous and $\widehat{f}(n)$ is $O\left(\frac{1}{n}\right)$, the Fourier series converges uniformly to f. In particular, this will give us that the Fourier series of C^1 converge uniformly since the Fourier coefficients are in $o\left(\frac{1}{n}\right)$.

Theorem 1.8. *If f is a continuous function and $\widehat{f}(n)$ is $O\left(\frac{1}{n}\right)$, then the partial Fourier sums $S_n(f)$ converge uniformly to f on T. [1]*

Proof. What we need to show in the proof is that the partial Fourier sums $S_n(f)$, defined as

$$S_n(f)(\theta) = \sum_{k=-n}^{n} \widehat{f}(k) e^{ik\theta}$$

converge uniformly to the function f, if f is continuous, and $\widehat{f}(k)$ is $O\left(\frac{1}{k}\right)$. What we do know is the the partial Cesàro sums $\sigma_n(f)$, defined as

$$\sigma_n(f)(\theta) = \frac{\sum_{k=1}^{n} S_k(\theta)}{n}$$

do converge uniformly to f. What we will do in this proof is create an intermediate series $I_{(n,k)}$, between S_n and σ_n, such that $I_{(n,k)}(f)$ converges uniformly to f as n goes to ∞ for large enough k, and the distance between $I_{(n,k)}(f)$ and $S_m(f)$, where $kn \leq m < (k+1)n$, is less than $\frac{A}{k}$ for some constant A.

Define the series $I_{(n,k)}$ as

$$I_{n,k}(f)(\theta) = \frac{S_{kn+1}(f)(\theta) + \ldots + S_{(k+1)n}(f)(\theta)}{n}$$

This can be equivalently written as

$$\begin{aligned} I_{(n,k)}(f)(\theta) &= \frac{((k+1)n+1)\sigma_{(k+1)n+1}(f)(\theta) - (kn+1)\sigma_{kn+1}(f)(\theta)}{n} && (4) \\ &= \left(k+1+\frac{1}{n}\right)\sigma_{(k+1)n+1}(f)(\theta) - \left(k+\frac{1}{n}\right)\sigma_{kn+1}(f)(\theta) && (5) \end{aligned}$$

Since we know that $\sigma_n(f)$ converges uniformly to f, expression 5 will also converge to f as n goes to ∞, for all k, although the rate of convergence will depend on k.

Now consider another way of writing $I_{(n,k)}(f)(\theta)$

$$I_{(n,k)}(f)(\theta) = \sum_{q=-kn}^{kn} \widehat{f}(q) e^{iq\theta} + \sum_{kn<|q|\leq(k+1)n} \frac{(k+1)n+1-|q|}{n} \widehat{f}(q)(n) e^{iq\theta}$$

Since $\widehat{f}(k)$ is $O\left(\frac{1}{k}\right)$, it is bounded above by $\frac{A}{k}$ for some constant A.

Writing $S_m(f)(\theta)$ similarly

$$S_m(f)(\theta) = \sum_{q=-kn}^{kn} \widehat{f}(q)e^{iq\theta} + \sum_{kn<|q|\leq m} \widehat{f}(q)(n)e^{iq\theta}$$

where $kn \leq m < (k+1)n$. Taking their difference, we get

$$\begin{aligned} |I_{(n,k)}(f)(\theta) - S_m(f)(\theta)| &\leq \sum_{kn<|q|\leq(k+1)n} |\widehat{f}(q)| \\ &\leq 2n \cdot \frac{A}{kn} \\ &= \frac{2A}{k} \end{aligned}$$

For $\varepsilon > 0$, pick a k such that $\frac{2A}{k} < \frac{\varepsilon}{2}$. Pick an $N > k$ such that for all $n > N$,

$$\left\| I_{(n,k)}(f) - f \right\|_\infty < \frac{\varepsilon}{2} \tag{6}$$

Now pick any $m > kN$. Since $N > k$, we have the following inequality

$$k(n+1) < (k+1)n$$

where $n > N$. That means for every $m > kN$, there exists $n > N$, such that

$$kn \leq m < (k+1)n$$

That means for all $m > kN$

$$\left\| S_m(f) - I_{(n,k)}(f) \right\|_\infty < \frac{\varepsilon}{2} \tag{7}$$

Combining inequalities 6 and 7, we get

$$\|S_m(f) - f\|_\infty < \varepsilon$$

This shows that the partial Fourier sums converge uniformly the function f. $\square$

1.4 A continuous function whose Fourier series diverges at a point

Blah

2 Weyl's equidistribution theorem

Blah

3 Roth's theorem

3.1 Fourier analysis on finite cyclic groups

Blah

References

[1] Rajendra Bhatia, *Fourier series*, Mathematical Association of America, 2005.