Fourier Analysis Notes

Connor Leipelt

Contents

1 Introduction

These notes follow through chapters of *Fourier Analysis* by Javier Duoandikoetxea. [1]

2 Chapter 1 Fourier Series and Integrals

2.1 Fourier coefficients and series

We start by looking at the problem of representing a function f, defined on (an interval of) $\mathbb{R}$, by a trigonometric series of the form

$$f(x) = \sum_{k=0}^{\infty} a_k \cos(kx) + b_k \sin(kx). \tag{1}$$

Note that since sine and cosine are functions of period 2π, we see that $\sin(kx)$ and $\cos(kx)$ have period $2\pi/k$. This is because if a function $f(x)$ has period θ, then $f(\lambda x)$ has period θ/λ. Note: A function has period θ if $f(x) = f(x+\theta)$ for all x.

Proof. Given a function $f(x)$ with period θ we investigate $f(\lambda x)$. Let $g(x) = f(\lambda x)$.

$$g\left(x + \frac{\theta}{\lambda}\right) = f\left(\lambda\left(x + \frac{\theta}{\lambda}\right)\right) = f(\lambda x + \theta).$$

Since $f(x)$ has period θ we see that $f(\lambda x + \theta) = f(\lambda x) = g(x)$.
Therefore $g(x) = g(x + \frac{\theta}{\lambda})$ so $g(x) = f(\lambda x)$ has period θ/λ. □

This means that $f(x)$ is composed of functions with periods of $2\pi/k$ for $k \in \mathbb{N}$. The largest of these periods is 2π which happens when $k = 1$ (note that when $k = 0$ the term of the sum is a_0 which is a constant value that doesn't change the period of $f(x)$). From here we prove that functions of period θ, meaning $f(x+\theta) = f(x)$, also have the fact that $f(x+np) = f(x)\ \forall n \in \mathbb{Z}$. This happens because

$$\begin{aligned} f(x+2p) &= f((x+p)+p) = f(x+p) = f(x) \text{ and} \\ f(x) &= f((x-p)+p) = f(x-p) \\ f(x-2p) &= f((x-p)-p) = f(x-p) = f(x) \end{aligned}$$

One can achieve this equality for any $n \in \mathbb{Z}$ by following this method.
This means that all terms with $2\pi/k$ also have period 2π making each term of the summation have period 2π. Now that we've shown each term of the summation has period 2π and sums of functions that share the same period also have that period,

Given $f(x)$ and $g(x)$ with period θ, let $h(x) = f(x) + g(x)$:

$$\begin{aligned} h(x+\theta) &= f(x+\theta) + g(x+\theta) = f(x) + g(x) \\ &= h(x) \end{aligned}$$

Therefore $h(x)$ also has period θ.

the function f must also have this property.

Now we will use Euler's identity:

From calculus we know that for $z \in \mathbb{C}$ we have

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!},$$

which has an infinite radius of convergence. This can be shown by the ratio test:

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \left| \frac{z^{n+1}}{(n+1)!} \cdot \frac{n!}{z^n} \right| = \lim_{n\to\infty} \left| \frac{z}{n} \right|,$$

since z is a fixed value independent of n we see that this limit goes to 0 as n goes to ∞ for any $z \in \mathbb{C}$. This means that this power series has an infinite radius of convergence (it also tells us that it is absolutely convergent). From here we look at

$$e^{ikx} = \sum_{n=0}^{\infty} \frac{(ikx)^n}{n!} = 1+ikx+\frac{(ikx)^2}{2!}+\frac{(ikx)^3}{3!}+\frac{(ikx)^4}{4!}+\frac{(ikx)^5}{5!}+\frac{(ikx)^6}{6!}+\frac{(ikx)^7}{7!}+\frac{(ikx)^8}{8!}+\dots$$

By multiplying out i we obtain the following:

$$e^{ikx} = 1 + ikx - \frac{(kx)^2}{2!} - \frac{i(kx)^3}{3!} + \frac{(kx)^4}{4!} + \frac{i(kx)^5}{5!} - \frac{(kx)^6}{6!} - \frac{i(kx)^7}{7!} + \frac{(kx)^8}{8!} + \dots$$

which then, since the series is absolutely convergent, we can group the terms containing i and factor i out to obtain:

$$e^{ikx} = \left(1 - \frac{(kx)^2}{2!} + \frac{(kx)^4}{4!} - \frac{(kx)^6}{6!} + \frac{(kx)^8}{8!} - \dots\right) + i\left(kx - \frac{(kx)^3}{3!} + \frac{(kx)^5}{5!} - \frac{(kx)^7}{7!} + \dots\right),$$

which are the series for sine and cosine giving us Euler's identity:

$$e^{ikx} = \cos(kx) + i\sin(kx).$$

Now that we have Euler's Identity, we look at the following:

$$e^{ikx} - e^{-ikx} = (\cos(kx) + i\sin(kx)) - (\cos(-kx) + i\sin(-kx)).$$

By definitions of even and odd functions, we know that since cosine is even we have $\cos(x) = \cos(-x)$ and since sine is odd we have $\sin(x) = -\sin(-x)$. Therefore, we can simplify to obtain:

$$\begin{aligned} e^{ikx} - e^{-ikx} &= \cos(kx) + i\sin(kx) - \cos(kx) + isin(kx) \\ &= 2i\sin(kx) \\ \frac{e^{ikx} - e^{-ikx}}{2i} &= \sin(kx). \end{aligned} \tag{2}$$

Similarly for a relation to cosine we investigate $e^{ikx} + e^{-ikx}$:

$$\begin{aligned}
e^{ikx} + e^{-ikx} &= (\cos(kx) + i\sin(kx)) + (\cos(-kx) + i\sin(-kx)) \\
&= \cos(kx) + i\sin(kx) + \cos(kx) - i\sin(kx) \\
&= 2\cos(kx) \\
\frac{e^{ikx} + e^{-ikx}}{2} &= \cos(kx). \qquad (3)
\end{aligned}$$

From here we look at (1) and substitute in (2) and (3) (note that $\frac{1}{i} = -i$)

$$\begin{aligned}
f(x) &= \sum_{k=0}^{\infty} a_k \left(\frac{e^{ikx} + e^{-ikx}}{2}\right) + b_k \left(\frac{e^{ikx} - e^{-ikx}}{2i}\right) \\
&= \sum_{k=0}^{\infty} \frac{a_k}{2} e^{ikx} + \frac{a_k}{2} e^{-ikx} + \frac{b_k}{2i} e^{ikx} - \frac{b_k}{2i} e^{-ikx} \\
&= \sum_{k=0}^{\infty} \frac{a_k}{2} e^{ikx} + \frac{a_k}{2} e^{-ikx} - \frac{b_k i}{2} e^{ikx} + \frac{b_k i}{2} e^{-ikx} \\
&= \sum_{k=0}^{\infty} \left(\frac{a_k - b_k i}{2}\right) e^{ikx} + \left(\frac{a_k + b_k i}{2}\right) e^{-ikx}.
\end{aligned}$$

Now if we define $c_k = \frac{a_k - b_k i}{2}$ and $c_{-k} = \frac{a_k + b_k i}{2}$ we can rewrite this as:

$$f(x) = \sum_{k=0}^{\infty} c_k e^{ikx} + c_{-k} e^{-ikx}. \qquad (4)$$

Note that for the first term of (4) ($k = 0$) we get the following

$$\begin{aligned}
c_0 e^0 + c_{-0} e^{-0} &= c_0 + c_{-0} \\
&= \left(\frac{a_0 - b_0 i}{2}\right) + \left(\frac{a_0 + b_0 i}{2}\right) \\
&= a_0
\end{aligned}$$

From (4) we can re-index the second term of this sum by seeing that:

$$\sum_{k=1}^{\infty} c_{-k} e^{-ikx} = \sum_{k=-\infty}^{-1} c_k e^{ikx}.$$

Therefore (4) becomes:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx} \qquad (5)$$

where

$$c_k = \begin{cases} \frac{a_k - b_k i}{2}, & k > 0 \\ a_k, & k = 0 \\ \frac{a_k + b_k i}{2}, & k < 0 \end{cases}$$

We wish to investigate functions with period 1 instead of 2π so we change our system of functions, $\{e^{ikx} : k \in \mathbb{Z}\}$, to the system of functions, $\{e^{2\pi ikx} : k \in \mathbb{Z}\}$. This is because if a function $f(x)$ has period θ, then $f(\lambda x)$ has period θ/λ. This was proven at the beginning of this section.

Hence, our problem is transformed into studying the representation of f by

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi ikx}. \tag{6}$$

If we assume that (6) converges uniformly, then we can solve for the c_k coefficients by multiplying by $e^{-2\pi imx}$ and integrating term-by-term on $(0, 1)$:

$$\begin{aligned} e^{-2\pi imx} \cdot f(x) &= e^{-2\pi imx} \cdot \sum_{-\infty}^{\infty} c_k e^{2\pi ikx} \\ \int_0^1 f(x)e^{-2\pi imx}\, dx &= \int_0^1 e^{-2\pi imx} \cdot \sum_{-\infty}^{\infty} c_k e^{2\pi ikx}\, dx \end{aligned} \tag{7}$$

Note that because of our assumption that the summation converges uniformly we are allowed to say the following about the RHS:

$$\begin{aligned} \int_0^1 \sum_{k=-\infty}^{\infty} e^{-2\pi imx} c_k e^{2\pi ikx}\, dx &= \sum_{k=-\infty}^{\infty} \int_0^1 e^{-2\pi imx} c_k e^{2\pi ikx}\, dx \\ &= \sum_{k=-\infty}^{\infty} \int_0^1 c_k e^{2\pi i(k-m)x}\, dx \end{aligned}$$

For the integral there are two cases,
Case 1: $k = m$

$$\int_0^1 c_k e^{2\pi i(k-m)x}\, dx = \int_0^1 c_m e^0\, dx = \int_0^1 c_m\, dx = c_m x\Big|_0^1 = c_m$$

Case 2: $k \neq m$ (Note that $k - m \in \mathbb{Z}$)

$$\begin{aligned} \int_0^1 c_k e^{2\pi i(k-m)x}\, dx &= \frac{c_k}{2\pi i(k-m)} e^{2\pi i(k-m)x}\Big|_0^1 \\ &= \frac{c_k}{2\pi i(k-m)} e^{2\pi i(k-m)} - \frac{c_k}{2\pi i(k-m)} \\ &= \frac{c_k}{2\pi i(k-m)} - \frac{c_k}{2\pi i(k-m)} \\ &= 0 \end{aligned}$$

Case 2 uses the fact that $e^{2\pi in} = 1$ for $n \in \mathbb{Z}$, we see this because

$$\begin{aligned} e^{\pi i} &= -1 \\ \left(e^{\pi i}\right)^2 = e^{2\pi i} &= (-1)^2 = 1 \\ \left(e^{2\pi i}\right)^n = e^{2\pi in} &= (1)^n = 1 \end{aligned}$$

Therefore if we go back to (7) we see that for a chosen m the only term that does not evaluate to 0 is the term where $k = m$. This means that we have:

$$\int_0^1 f(x)e^{-2\pi imx}\,dx = c_m \tag{8}$$

The reason for making our Fourier series have period 1 is because of the following: denote the additive group of the reals modulo 1 (aka $\mathbb{R}/\mathbb{Z}$) by $\mathbb{T}$, the one-dimensional torus. This can be identified with the unit circle, S^1. This takes any real valued number and divides it by 1, whatever the remainder is is now the new value of the number and exists in $[0, 1]$. This effectively maps the whole real number line into $[0, 1]$. Therefore, saying that a function is defined on $\mathbb{T}$ is equivalent to saying that it is defined on $\mathbb{R}$ and has period 1.

To each function $f \in L^1(\mathbb{T})$ we associate the sequence $\{\hat{f}(k)\}$ of Fourier coefficients of f, which are defined by

$$\hat{f}(k) = \int_0^1 f(x)e^{-2\pi ikx}\,dx \tag{9}$$

Our trigonometric series defined earlier combined with this gives us the Fourier series of f

$$\sum_{k=-\infty}^{\infty} \hat{f}(k)e^{2\pi ikx}. \tag{10}$$

Our problem is now to determine when and in what sense does the series (10) represent the function f.

2.2 Criteria for pointwise convergence

We start by denoting the N-th symmetrical partial sum of (10) by $S_N f(x)$; that is,

$$S_N f(x) = \sum_{k=-N}^{N} \hat{f}(k)e^{2\pi ikx}.$$

Note that this is also the N-th partial sum of (1).

Our first approach to this problem of representing f by its Fourier series is to determine whether $\lim S_N f(x)$ exists for each x, and if so, whether it is equal to $f(x)$. The first result was proven by P.G. L. Dirichlet (1829), who proved that if f is bounded, piecewise continuous, and has a finite number of maxima and minima, then $\lim S_N f(x)$ exists and is equal to

$\frac{1}{2}[f(x+) + f(x-)]$. $f(x+)$ denotes the limit approaching x from the right: $\lim_{a\to x^+} f(a)$ and $f(x-)$ denotes the limit approaching x from the left: $\lim_{a\to x^-} f(a)$.
In order to study $S_N f(x)$ we need a more manageable expression of it. Dirichlet wrote the partial sums as follows:

$$\begin{aligned}
S_N f(x) &= \sum_{k=-N}^{N} \hat{f}(k) e^{2\pi i k x} \\
&= \sum_{k=-N}^{N} \int_0^1 f(t) e^{-2\pi i k t}\, dt \cdot e^{2\pi i k x} \\
&= \sum_{k=-N}^{N} \int_0^1 f(t) e^{2\pi i k (x-t)}\, dt \\
&= \int_0^1 f(t) \sum_{k=-N}^{N} e^{2\pi i k (x-t)}\, dt
\end{aligned}$$

From here we will stop to define and implement the Dirichlet kernel, D_N. Note that this is a finite sum so we are able to expand it out and group the $e^{2\pi i k(x-t)}$ terms back together since each term shares the same integral and bounds.

Proof. We shall prove the Dirichlet kernel:

$$\begin{aligned}
D_N(t) &= \sum_{k=-N}^{N} e^{2\pi i k t} \\
&= e^{2\pi i(-N)t} + e^{2\pi i(-N+1)t} + \cdots + e^{2\pi i N t} \\
&= e^{2\pi i(-N)t}\left(1 + e^{2\pi i t} + \cdots + e^{2\pi i (2N) t}\right)
\end{aligned}$$

Now recall the finite ratio sum:

$$\begin{aligned}
s_N &= \sum_{k=1}^{N} a r^{k-1} = a + ar + \cdots + ar^{N-1} \\
s_N - r s_N &= \left(a + ar + \cdots + ar^{N-1}\right) - \left(ar + \cdots + ar^{N-1} + ar^N\right) \\
s_N(1-r) &= a\left(1 - r^N\right) \\
s_N &= \frac{a\left(1-r^N\right)}{1-r}
\end{aligned}$$

This means that we can rewrite the last expression ($a = 1$)

$$\begin{aligned}
D_N(t) &= e^{2\pi i(-N)t} \sum_{k=1}^{2N+1} \left(e^{2\pi i t}\right)^{k-1} \\
&= e^{2\pi i(-N)t} \cdot \frac{1 - \left(e^{2\pi i t}\right)^{2N+1}}{1 - e^{2\pi i t}} \\
&= \frac{e^{2\pi i(-N)t} - e^{2\pi i(2N+1-N)t}}{1 - e^{2\pi i t}}
\end{aligned}$$

Now we multiply by $\frac{e^{-\pi it}}{e^{-\pi it}} = 1$ and then by $\frac{1/2i}{1/2i} = 1$

$$\begin{aligned} D_N(t) &= \frac{e^{\pi i(-2N)t} - e^{\pi i2(N+1)t}}{1 - e^{2\pi it}} \cdot \frac{e^{-\pi it}}{e^{-\pi it}} \\ &= \frac{e^{\pi i(-(2N+1))t} - e^{\pi i(2N+1)t}}{e^{-\pi it} - e^{\pi it}} \\ &= \frac{(-1) \cdot \frac{e^{\pi i(2N+1)t} - e^{\pi i(-(2N+1))t}}{2i}}{(-1) \cdot \frac{e^{\pi it} - e^{-\pi it}}{2i}} \end{aligned}$$

which using the sine equality,(2), from section 1 gives our desired result

$$D_N(t) = \frac{\sin(\pi(2N+1)t)}{\sin(\pi t)} \tag{11}$$

□

Note that

$$\int_0^1 D_N(t)\, dt = 1 \tag{12}$$

$$|D_N(t)| \leq \frac{1}{sin(\pi\delta)} \text{ for } \delta \leq |t| \leq 1/2 \tag{13}$$

We will now prove (12) and (13).

Proof. We will prove (12):

$$\begin{aligned} \int_0^1 D_N(t)\, dt = \int_0^1 \sum_{k=-N}^{N} e^{2\pi ikt}\, dt &= \int_0^1 \sum_{k=-N}^{-1} e^{2\pi ikt} + 1 + \sum_{k=1}^{N} e^{2\pi ikt}\, dt \\ &= \left(\sum_{k=-N}^{-1} \frac{1}{2\pi ik} e^{2\pi ikt} + t + \sum_{k=1}^{N} \frac{1}{2\pi ik} e^{2\pi ikt} \right) \Big|_0^1 \\ &= \left(\sum_{k=-N}^{-1} \frac{1}{2\pi ik} e^{2\pi ik} + 1 + \sum_{k=1}^{N} \frac{1}{2\pi ik} e^{2\pi ik} \right) - \left(\sum_{k=-N}^{-1} \frac{1}{2\pi ik} e^{0} + 0 + \sum_{k=1}^{N} \frac{1}{2\pi ik} e^{0} \right) \end{aligned}$$

From earlier we showed that $e^{2\pi in} = 1$ for $n \in \mathbb{Z}$ therefore we are left with

$$\begin{aligned} \int_0^1 D_N(t) &= \left(\sum_{k=-N}^{-1} \frac{1}{2\pi ik} + 1 + \sum_{k=1}^{N} \frac{1}{2\pi ik} \right) - \left(\sum_{k=-N}^{-1} \frac{1}{2\pi ik} + \sum_{k=1}^{N} \frac{1}{2\pi ik} \right) \\ &= 1. \end{aligned}$$

Concluding the proof of (12). □

Proof. We will prove (13):
To prove (13) we must prove

$$|D_N(t)| = \left| \frac{\sin(\pi(2N+1)t)}{\sin(\pi t)} \right| \leq \frac{1}{sin(\pi\delta)} \text{ for } \delta \leq |t| \leq 1/2$$

We know that the following holds $\forall x \in \mathbb{R}$.

$$0 \leq |sin(x)| \leq 1$$

Also note that for $0 \leq a \leq b$ and $c \geq 0$ we have that $0 \leq a \cdot c \leq b \cdot c$.
This means that for $0 \leq a \leq b$ and $0 \leq c \leq d$ we have $0 \leq a \cdot c \leq b \cdot d$.
This comes directly:
Given $0 \leq c \leq d$ and $b \geq 0$ we get $0 \leq c \cdot b \leq d \cdot b$.
Given $0 \leq a \leq b$ and $c \geq 0$ we get $0 \leq a \cdot c \leq b \cdot c$.
Therefore since multiplication is commutative we have $0 \leq a \cdot c \leq b \cdot c \leq b \cdot d$ showing that $0 \leq a \cdot c \leq b \cdot d$ as wanted.
From this and the sine inequality (which holds also for $sin(\pi(2N+1)x)$ since it holds for all $x \in \mathbb{R}$) we only need to show:

$$0 \leq \frac{1}{|sin(\pi t)|} \leq \frac{1}{sin(\pi\delta)} \text{ for } \delta \leq |t| \leq 1/2$$

Since then

$$0 \leq |sin(x)| \leq 1 \text{ and } 0 \leq \frac{1}{|sin(\pi t)|} \leq \frac{1}{sin(\pi\delta)} \Rightarrow \quad \left|\frac{\sin(\pi(2N+1)t)}{\sin(\pi t)}\right| \leq \frac{1}{sin(\pi\delta)} \text{ for } \delta \leq |t| \leq 1/2$$

Now to prove $0 \leq \frac{1}{|sin(\pi t)|} \leq \frac{1}{sin(\pi\delta)}$ for $\delta \leq |t| \leq 1/2$ we will first examine $\sin(\pi t)$. The derivative of $\sin(\pi t)$ is $\pi\cos(\pi t)$ which equals 0 for $t = \frac{1}{2} + n$ for $n \in \mathbb{Z}$. This means that $\pi\cos(\pi t)$ is either positive $\forall t \in \left(-\frac{1}{2}, \frac{1}{2}\right)$ or negative $\forall t \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. Since $\pi\cos(\pi \cdot 0) = 1$ we see that $\pi\cos(\pi t)$ is positive $\forall t \in \left(-\frac{1}{2}, \frac{1}{2}\right)$ meaning that $\sin(\pi t)$ is strictly increasing on $\left(-\frac{1}{2}, \frac{1}{2}\right)$.
Also note that for $t = -1/2$ and $t = 1/2$ we see from the first and second derivatives that these points are minimum and maximum points respectively (second derivative of $\sin(\pi t)$ is $-\pi^2\sin(\pi t)$)

$$\pi\cos\left(\pi \cdot \left(-\frac{1}{2}\right)\right) = 0 \text{ and } -\pi^2\sin\left(\pi \cdot \left(-\frac{1}{2}\right)\right) = \pi^2 > 0 \Rightarrow \ t = -\frac{1}{2} \text{ is a minimum.}$$
$$\pi\cos\left(\pi \cdot \frac{1}{2}\right) = 0 \text{ and } -\pi^2\sin\left(\pi \cdot \frac{1}{2}\right) = -\pi^2 < 0 \Rightarrow \ t = \frac{1}{2} \text{ is a maximum.}$$

Now that we know that $\sin(\pi t)$ is increasing for $t \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, we look at $0 < \delta \leq t \leq 1/2$

$$0 < \delta \leq t \leq 1/2 \Rightarrow \ 0 < \sin(\pi\delta) \leq \sin(\pi t) \leq \sin(\pi/2)$$

Since $\sin(\pi\delta) > 0$ and $\sin(\pi t) > 0$ we have

$$0 < \sin(\pi\delta) \leq \sin(\pi t)$$
$$0 < 1 \leq \frac{\sin(\pi t)}{\sin(\pi\delta)}$$
$$0 < \frac{1}{\sin(\pi t)} \leq \frac{1}{\sin(\pi\delta)}$$

Since both expressions are greater than 0, we have shown that $0 \leq \frac{1}{|sin(\pi t)|} \leq \frac{1}{sin(\pi\delta)}$ for $0 < \delta \leq t \leq 1/2$.
We must now prove this inequality for $-1/2 \leq t \leq -\delta < 0$.
Since $\sin(\pi t)$ is increasing for $t \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ we have

$$\sin(-\pi/2) \leq \sin(\pi t) \leq \sin(-\pi\delta) < 0$$

Note that sine is an odd function we have $\sin(t) = -\sin(-t)$. From this we can write $\sin(-\pi\delta) = -\sin(\pi\delta)$ and $\sin(-\pi/2) = -\sin(\pi/2)$ giving us

$$-\sin(\pi/2) \leq \sin(\pi t) \leq -\sin(\pi\delta) < 0$$

We multiply each expression by -1 to flip the inequality signs

$$0 < \sin(\pi\delta) \leq -\sin(\pi t) \leq \sin(\pi/2)$$

Lastly, since we have $0 < -\sin(\pi t)$ this means that taking absolute values does not change the inherent value of the expression so similarly we do the following

$$\begin{aligned}
&0 < \sin(\pi\delta) \leq |\sin(\pi t)| \\
&0 < 1 \leq \frac{|\sin(\pi t)|}{\sin(\pi\delta)} \\
&0 < \frac{1}{|\sin(\pi t)|} \leq \frac{1}{\sin(\pi\delta)}
\end{aligned}$$

Hence we have proven that $0 < \frac{1}{|\sin(\pi t)|} \leq \frac{1}{\sin(\pi\delta)}$ for $\delta \leq |t| \leq 1/2$ meaning that we have truly proven that

$$|D_N(t) = \left|\frac{\sin(\pi(2N+1)t)}{\sin(\pi t)}\right| \leq \frac{1}{sin(\pi\delta)} \text{ for } \delta \leq |t| \leq 1/2$$

□

Now that we have proven the Dirichlet kernel and some facts about it, we can continue with writing the partial sums from the beginning of this section

$$\begin{aligned}
S_N f(x) &= \int_0^1 f(t) \sum_{k=-N}^{N} e^{2\pi i k(x-t)}\, dt \\
&= \int_0^1 f(t) D_N(x-t)\, dt
\end{aligned}$$

From here we do a change of variables, let $v = x - t$, then $dv = -dt$. This changes our bounds from $1 \Rightarrow\ x - 1$ and $0 \Rightarrow\ x$, therefore we have

$$S_N f(x) = \int_x^{x-1} f(x-v) D_N(v) \cdot (-1)\, dv$$

From here we flip our bounds on the integral leaving the negative signs to multiply to a positive 1

$$S_N f(x) = \int_{x-1}^{x} f(x-v) D_N(v)\, dv \tag{14}$$

Now we enact the property of periodic functions with period p

$$\int_a^{a+p} \phi(x)\, dx = \int_0^p \phi(x)\, dx \text{ for any } a \in \mathbb{R}. \tag{15}$$

Proof. We will prove (15):
First we choose an $n \in \mathbb{Z}$ such that $a \leq np < a+p$ holds. This is always possible due to the fact that the distance between a and $a+p$ is p and the expression np is allowed to be equal to a.
From here we split the integral apart

$$\int_a^{a+p} \phi(x)\, dx = \int_a^{np} \phi(x)\, dx + \int_{np}^{a+p} \phi(x)\, dx.$$

For the first term on the RHS we do the following change of variables $x = t + (n-1)p$ with $dx = dt$ with bounds changing by $t = x - (n-1)p$. For the second term we change the variables as follows $x = t + np$ and $dx = dt$ with bounds changing by $t = x - np$. This changes the integral to the following

$$\int_a^{a+p} \phi(x)\, dx = \int_{a-(n-1)p}^{np-(n-1)p} \phi(t + (n-1)p)\, dt + \int_{np-np}^{a+p-np} \phi(t+np)\, dt$$

From here we use the fact that for functions of period p we have $\phi(x+np) = \phi(x)$ $\forall n \in \mathbb{Z}$. Which we proved at the beginning of this chapter.
From this we have $\phi(t + (n-1)p) = \phi(t+np) = \phi(t)$ giving us, along with cleaning up the bounds,

$$\begin{aligned}
\int_a^{a+p} \phi(x)\, dx &= \int_{a-(n-1)p}^{np-(n-1)p} \phi(t)\, dt + \int_{np-np}^{a+p-np} \phi(t)\, dt \\
&= \int_{a+(1-n)p}^{p} \phi(t)\, dt + \int_0^{a+(1-n)p} \phi(t)\, dt \\
&= \int_0^p \phi(t)\, dt
\end{aligned}$$

If we wish we can change our variable from t to x which then shows

$$\int_a^{a+p} \phi(x)\, dx = \int_0^p \phi(x)\, dx \text{ for any } a \in \mathbb{R},$$

concluding the proof. ☐

$S_N f(x)$ is a function with period 1 since it is the sum of functions with period 1. This means that in (14) we have (letting $a = x - 1$)

$$S_N f(x) = \int_{x-1}^{x} f(x-v) D_N(v)\, dv = \int_{(x-1)}^{(x-1)+1} f(x-v) D_N(v)\, dv = \int_{0}^{1} f(x-v) D_N(v)\, dv$$

Relabeling v as t gives us the end result:

$$S_N f(x) = \int_0^1 f(x-t) D_N(t)\, dt \tag{16}$$

From here we will now prove two criteria for pointwise convergence.

Theorem 2.1 (Dini's Criterion). *If for some x there exists $\delta > 0$ such that*

$$\int_{|t|<\delta} \left| \frac{f(x+t) - f(x)}{t} \right| dt < \infty$$

then

$$\lim_{N\to\infty} S_N f(x) = f(x).$$

Theorem 2.2 (Jordan's Criterion). *If f is a function of bounded variation in a neighborhood of x, then*

$$\lim_{N\to\infty} S_N f(x) = \frac{1}{2}[f(x+) + f(x-)].$$

It is important to realize that the convergence of a Fourier series is effectively a local property even though a slight change to the function results to a change in the Fourier coefficients. If the modifications are made outside of a neighborhood of x, then the behavior of the series at x does not change. This is made precise by the following result.

Theorem 2.3 (Riemann Localization Principle). *If f is zero in a neighborhood of x, then*

$$\lim_{N\to\infty} S_N f(x) = 0.$$

An equivalent formulation of this result is to say that if two functions agree in the neighborhood of x, then their Fourier series behave in the same way at x. This is due to the following:

Goal: Show that if two functions agree in the neighborhood of x, then their Fourier series behave in the same way at x (assuming the Riemann Localization Principle). Assume $f(x) - g(x) = 0$ on $(x-\delta, x+\delta)$. From this we want to show $\lim_{N\to\infty} (S_N f(x) - S_N g(x)) = 0$. Define $h = f - g$ meaning $h(x) = 0$ for $(x-\delta, x+\delta)$. From the Riemann Localization Principle we then have $\lim_{N\to\infty} S_N h(x) = 0$. Now we proceed:

$$\begin{aligned}
0 = \lim_{N\to\infty} S_N h(x) &= \lim_{N\to\infty} S_N(f-g)(x) \\
&= \lim_{N\to\infty} \sum_{k=-N}^{N} (f-g)\widehat{\ }(k) e^{2\pi i k x} \\
&= \lim_{N\to\infty} \sum_{k=-N}^{N} \int_0^1 (f(x)-g(x)) e^{-2\pi i k x}\, dx \cdot e^{2\pi i k x} \\
&= \lim_{N\to\infty} \sum_{k=-N}^{N} \left[\int_0^1 f(x) e^{-2\pi i k x}\, dx \cdot e^{2\pi i k x} - \int_0^1 g(x) e^{-2\pi i k x}\, dx \cdot e^{2\pi i k x} \right] \\
&= \lim_{N\to\infty} \sum_{k=-N}^{N} \int_0^1 f(x) e^{-2\pi i k x}\, dx \cdot e^{2\pi i k x} - \sum_{k=-N}^{N} \int_0^1 g(x) e^{-2\pi i k x}\, dx \cdot e^{2\pi i k x} \\
0 &= \lim_{N\to\infty} S_N f(x) - S_N g(x)
\end{aligned}$$

Showing that these statements are equivalent.
From the definition of Fourier coefficients (9) we see that

$$|\hat{f}(k)| \le \|f\|_1,$$

Proof. We first use the Triangle inequality

$$\begin{aligned}
|\hat{f}(k)| = \left| \int_0^1 f(x) e^{-2\pi i k x}\, dx \right| &\le \int_0^1 \left| f(x) \cdot e^{-2\pi i k x} \right| \, dx \\
&= \int_0^1 |f(x)| \left| e^{-2\pi i k x} \right| \, dx
\end{aligned}$$

Now remember Euler's identity and what it means to take the $|z|$ for $z = a + bi \in \mathbb{C}$

$$e^{i\theta} = \cos(\theta) + i \sin(\theta) \text{ and } |z| = \sqrt{a^2 + b^2}$$

Therefore:

$$\begin{aligned}
e^{i(-2\pi k x)} &= \cos(-2\pi k x) + i \sin(-2\pi k x) \\
\left| e^{i(-2\pi k x)} \right| &= \sqrt{(\cos(-2\pi k x))^2 + (\sin(-2\pi k x))^2} \\
&= \sqrt{1} \\
&= 1
\end{aligned}$$

Therefore $\left| e^{i(-2\pi k x)} \right| = 1$ so we have

$$|\hat{f}(k)| \le \int_0^1 |f(x)|\, dx = \|f\|_1$$

as desired. □

This estimate is good, but a sharper estimate is true and we will use it to prove the past theorems.

Lemma 2.4 (Riemann-Lebesgue)**.** *If* $f \in L^1(\mathbb{T})$ *then*

$$\lim_{|k|\to\infty} \hat{f}(k) = 0.$$

Proof of Lemma 2.4. Since f and $e^{-2\pi ix}$ have period 1 and $-e^{-\pi i} = -(\cos(-\pi)+i\sin(-\pi)) = 1$, we have

$$\begin{aligned}
\hat{f}(k) &= \int_0^1 f(x)e^{-2\pi ikx} \cdot 1\,dx \\
&= \int_0^1 f(x)e^{-2\pi ikx} \cdot -e^{-\pi i}\,dx \\
&= -\int_0^1 f(x)e^{-2\pi ik(x+1/2k)}\,dx
\end{aligned}$$

We then do a change of variables, $v = x + 1/2k$ and $dv = dx$

$$\hat{f}(k) = -\int_{1/2k}^{1+1/2k} f(v - 1/2k)e^{-2\pi iv}\,dv$$

Since both f and $e^{-2\pi ix}$ are period 1 and the multiplication of functions with the same period produce a function with the same period:
Given $f(x)$ and $g(x)$ have period T, then $h(x)$ also has period T.

$$\begin{aligned}
h(x) &= f(x)g(x) \\
h(x+T) &= f(x+T)g(x+T) = f(x)g(x) = h(x)
\end{aligned}$$

Therefore the RHS also had period 1, so with relabeling $v = x$

$$\begin{aligned}
\hat{f}(k) &= -\int_{1/2k}^{1+1/2k} f(v - 1/2k)e^{-2\pi iv}\,dv \\
&= -\int_0^1 f(v - 1/2k)e^{-2\pi iv}\,dv \\
&= -\int_0^1 f(x - 1/2k)e^{-2\pi ix}\,dx
\end{aligned}$$

So

$$\begin{aligned}
2\hat{f}(k) &= \int_0^1 f(x)e^{-2\pi ikx}\,dx - \int_0^1 f(x - 1/2k)e^{-2\pi ix}\,dx \\
&= \int_0^1 [f(x) - f(x - 1/2k)]e^{-2\pi ikx}\,dx \\
\hat{f}(k) &= \frac{1}{2}\int_0^1 [f(x) - f(x - 1/2k)]e^{-2\pi ikx}\,dx
\end{aligned}$$

Now, if f is continuous, it follows by taking $\lim_{|k|\to\infty}$ that

$$\lim_{|k|\to\infty} \hat{f}(k) = \lim_{|k|\to\infty} \frac{1}{2}\int_0^1 [f(x) - f(x-1/2k)]e^{-2\pi ikx}\, dx \ = 0$$

The second equality happens because we are assuming that f is continuous and $f \in L^1(\mathbb{T})$ meaning that f is bounded. From this we see that $\lim_{|k|\to\infty} f(x-1/2k) = f(x)$ for a.e. x and that it is bounded by an integrable function (due to being bounded) therefore by the Dominated Convergence Theorem we can pull the limit inside of the integral and evaluate as follows (also note that $\lim_{|k|\to\infty} e^{-2\pi ikx}$ oscillates between 1 and -1 therefore combined with $\lim_{|k|\to\infty}[f(x) - f(x+1/2k)]$ going to 0 makes this limit approaches 0)

$$\lim_{|k|\to\infty} \frac{1}{2}\int_0^1 [f(x) - f(x-1/2k)]e^{-2\pi ikx}\, dx = \frac{1}{2}\int_0^1 \lim_{|k|\to\infty}[f(x) - f(x-1/2k)]e^{-2\pi ikx}\, dx = 0$$

For arbitrary $f \in L^1(\mathbb{T})$, given $\epsilon > 0$, choose g continuous such that $\|f-g\|_1 < \epsilon/2$ and choose k sufficiently large that $|\hat{g}(k)| < \epsilon/2$ (which is possible because g is continuous). Then by triangle inequality and the fact that $|\hat{f}(k)| \le \|f\|_1$ we have

$$\begin{aligned}
|\hat{f}(k)| &= \left|\int_0^1 [f(x) - g(x)]e^{-2\pi ikx}\, dx + \int_0^1 g(x)e^{-2\pi ikx}\, dx\right| \\
&\le \left|\int_0^1 [f(x) - g(x)]e^{-2\pi ikx}\, dx\right| + \left|\int_0^1 g(x)e^{-2\pi ikx}\, dx\right| \\
&= |(f-g)\widehat{\ }(k)| + |\hat{g}(k)| \le \|f-g\|_1 + |\hat{g}(k)| \\
&< \epsilon/2 + \epsilon/2 = \epsilon
\end{aligned}$$

Note that we use the fact that continuous functions are dense in $L^1(\mathbb{T})$. This is due to the fact that simple functions can approximate functions in $L^1(\mathbb{T})$ and step functions can be approximated by simple functions and continuous functions can be approximated by step functions. $\square$

We will now prove Theorem 2.3

Proof of Theorem 2.3. We start by assuming that $f(t) = 0$ on $(x-\delta, x+\delta)$. Then since $S_N f(x)$ has period 1

$$\begin{aligned}
S_N f(x) = \int_0^1 f(x-t)D_N(t)\, dt &= \int_{-1/2}^{1/2} f(x-t)D_N(t)\, dt \\
&= \int_{\delta\le|t|<1/2} f(x-t)D_N(t)\, dt
\end{aligned}$$

From here we write out $D_N(t)$ and use the sine equality (2)

$$\begin{aligned}
S_N f(x) &= \int_{\delta\le|t|<1/2} f(x-t)D_N(t)\,dt \\
&= \int_{\delta\le|t|<1/2} f(x-t)\cdot\frac{\sin(\pi(2N+1)t)}{\sin(\pi t)}\,dt \\
&= \int_{\delta\le|t|<1/2} \frac{f(x-t)}{\sin(\pi t)}\cdot\frac{e^{i\pi(2N+1)t}-e^{-i\pi(2N+1)t}}{2i}\,dt \\
&= \int_{\delta\le|t|<1/2} \frac{f(x-t)}{2i\sin(\pi t)}\cdot[e^{i\pi 2Nt}\cdot e^{i\pi t}-e^{-i\pi 2Nt}\cdot e^{-i\pi t}]\,dt \\
&= \int_{\delta\le|t|<1/2} \left(\frac{f(x-t)}{2i\sin(\pi t)}e^{i\pi t}\right)e^{-2\pi i(-N)t}\,dt - \int_{\delta\le|t|<1/2} \left(\frac{f(x-t)}{2i\sin(\pi t)}e^{-i\pi t}\right)e^{-2\pi i(N)t}\,dt \\
&= (g\cdot e^{i\pi})\widehat{\ }(-N) - (g\cdot e^{-i\pi})\widehat{\ }(N)
\end{aligned}$$

where

$$g(t) = \frac{f(x-t)}{2i\sin(\pi t)}\chi_{\{\delta\le|t|<1/2\}}(t)$$

Note that this is integrable due to the fact that $\delta > 0$ meaning that the denominator is bounded by a strict nonzero output. Also note that $\sin(\pi t)$ has period 1 with zeros at $t\in\mathbb{Z}$ which are not allowed for the given interval. Again note that $e^{i\pi t}$ and $e^{-2\pi i(\pm N)t}$ both oscillate between 1 and -1 making the integrands integrable.
From here we see that the Riemann-Lebesgue Lemma allows us to take the limit as $N\to\infty$, since $f\in L^1(\mathbb{T})$, hence we achieve $\lim_{N\to\infty} S_N f(x) = 0$

$$\begin{aligned}
\lim_{N\to\infty} S_N f(x) &= \lim_{N\to\infty}\left[(g\cdot e^{i\pi})\widehat{\ }(-N) - (g\cdot e^{-i\pi})\widehat{\ }(N)\right] \\
&= 0.
\end{aligned}$$

$\square$

Now we will prove Theorem 2.1

Proof of Theorem 2.1. From (12) we know that $\int_0^1 D_N(t)\,dt = 1$ and $D_N(t)$ has period 1. This is because

$$D_N(t) = \sum_{k=-N}^{N} e^{2\pi ikt} \Rightarrow D_N(t+1) = \sum_{k=-N}^{N} e^{2\pi ik(t+1)} = \sum_{k=-N}^{N} e^{2\pi ikt}\cdot e^{2\pi ik}$$

Since $e^{2\pi ik} = 1$ for $k\in\mathbb{Z}$, as we proved in case 2 of computing Fourier coefficients, we see

$$D_N(t+1) = \sum_{k=-N}^{N} e^{2\pi ikt}\cdot e^{2\pi ik} = \sum_{k=-N}^{N} e^{2\pi ikt}\cdot 1 = D_N(t),$$

proving that $D_N(t)$ has period 1.
Therefore we can write the following

$$\begin{aligned}
S_N f(x) - f(x) &= \int_{-1/2}^{1/2} f(x-t) D_N(t)\,dt - f(x) \cdot \int_{-1/2}^{1/2} D_N(t)\,dt \\
&= \int_{-1/2}^{1/2} [f(x-t) - f(x)] \cdot \frac{\sin(\pi(2N+1)t)}{\sin(\pi t)}\,dt \\
&= \int_{|t|<\delta} [f(x-t) - f(x)] \cdot \frac{\sin(\pi(2N+1)t)}{\sin(\pi t)}\,dt \\
&\quad + \int_{\delta \le |t| < 1/2} [f(x-t) - f(x)] \cdot \frac{\sin(\pi(2N+1)t)}{\sin(\pi t)}\,dt
\end{aligned}$$

The second integral follows very similarly to the proof of Theorem 2.3

$$\begin{aligned}
&\int_{\delta \le |t| < 1/2} [f(x-t) - f(x)] \cdot \frac{\sin(\pi(2N+1)t)}{\sin(\pi t)}\,dt = \int_{\delta \le |t| < 1/2} \frac{f(x-t) - f(x)}{2i\sin(\pi t)} \cdot [e^{i\pi 2Nt} \cdot e^{i\pi t} - e^{-i\pi 2Nt} \cdot e^{-i\pi t}]\,dt \\
&= \int_{\delta \le |t| < 1/2} \left(\frac{f(x-t) - f(x)}{2i\sin(\pi t)} e^{i\pi t}\right) e^{-2\pi i(-N)t}\,dt - \int_{\delta \le |t| < 1/2} \left(\frac{f(x-t) - f(x)}{2i\sin(\pi t)} e^{-i\pi t}\right) e^{-2\pi i(N)t}\,dt \\
&= (h \cdot e^{i\pi})\widehat{\ }(-N) - (h \cdot e^{-i\pi})\widehat{\ }(N)
\end{aligned}$$

where

$$h(t) = \frac{f(x-t) - f(x)}{2i\sin(\pi t)} \chi_{\{\delta \le |t| < 1/2\}}(t).$$

Therefore by the Riemann-Lebesgue Lemma (since $f \in L^1(\mathbb{T})$) and by the fact that our denominator is bounded by δ avoiding the problem of dividing by 0, we see that this integral evaluates to 0 as we take the limit as $N \to \infty$.
Now for the first integral we start similarly

$$\begin{aligned}
&\int_{|t|<\delta} [f(x-t) - f(x)] \cdot \frac{\sin(\pi(2N+1)t)}{\sin(\pi t)}\,dt = \int_{|t|<\delta} \frac{f(x-t) - f(x)}{2i\sin(\pi t)} \cdot [e^{i\pi 2Nt} \cdot e^{i\pi t} - e^{-i\pi 2Nt} \cdot e^{-i\pi t}]\,dt \\
&= \int_{|t|<\delta} \left(\frac{f(x-t) - f(x)}{2i\sin(\pi t)} e^{i\pi t}\right) e^{-2\pi i(-N)t}\,dt - \int_{|t|<\delta} \left(\frac{f(x-t) - f(x)}{2i\sin(\pi t)} e^{-i\pi t}\right) e^{-2\pi i(N)t}\,dt \\
&= (p \cdot e^{i\pi})\widehat{\ }(-N) - (p \cdot e^{-i\pi})\widehat{\ }(N)
\end{aligned}$$

where

$$p(t) = \frac{f(x-t) - f(x)}{2i\sin(\pi t)} \chi_{\{|t|<\delta\}}(t).$$

The important thing to notice here is that we are integrating on $|t| < \delta$. From here we now show that we can approximate $\sin(\pi t)$ to be πt for small t. We start by investigating the Maclaurin series of $\sin(x)$.

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots$$

From here we can look at the error of an approximation by x

$$R(x) = |\sin(x) - x| = \left| -\frac{x^3}{3!} + \frac{x^5}{5!} - \dots \right|.$$

Then we look at the relative error which is given by the error term divided by x

$$\text{Relative Error } = \frac{\left| -\frac{x^3}{3!} + \frac{x^5}{5!} - \dots \right|}{|x|} = \frac{x^2}{3!} + \frac{x^4}{4!} + O(x^6)$$

As we take the limit of $|x| \to 0$ we see that the relative error also $\to 0$. Therefore for $x = \pi t$ we see that $\sin(\pi t)$ and πt are equivalent when $|t| < \delta$ for an appropriately small δ. Now we can look at $1/\sin(\pi t)$ and $1/\pi t$

$$\frac{1}{\sin(\pi t)} = \frac{1}{\pi t - \frac{(\pi t)^3}{3!} + \frac{(\pi t)^5}{5!} - \dots} = \frac{1}{\pi t \left(1 - \frac{(\pi t)^2}{3!} + \frac{(\pi t)^4}{5!} - \dots \right)} = \frac{1}{\pi t} \cdot \frac{1}{1 - \frac{(\pi t)^2}{3!} + \frac{(\pi t)^4}{5!} - \dots}$$

From here we investigate the error and the relative error

$$R(x) = \left| \frac{1}{\sin(\pi t)} - \frac{1}{\pi t} \right| = \left| \frac{1}{\pi t} \cdot \frac{1}{1 - \frac{(\pi t)^2}{3!} + \frac{(\pi t)^4}{5!} - \dots} - \frac{1}{\pi t} \right|$$

$$\text{Relative Error } = \frac{\left| \frac{1}{\pi t} \left(\frac{1}{1 - \frac{(\pi t)^2}{3!} + \frac{(\pi t)^4}{5!} - \dots} - 1 \right) \right|}{\left| \frac{1}{\pi t} \right|} = \left| \frac{1}{1 - \frac{(\pi t)^2}{3!} + \frac{(\pi t)^4}{5!} - \dots} - 1 \right|$$

From here we see that as we take the limit of $|t| \to 0$ we see the relative error $\to 0$. Hence we have proven that $1/\sin(\pi t)$ and $1/\pi t$ are equivalent when $|t| < \delta$ for an appropriately small δ.

Now that we have proven that $1/\sin(\pi t)$ and $1/\pi t$ are equivalent when $|t| < \delta$ for an appropriately small δ we can say that the following two expressions are equivalent for the specified δ (we show both $(p \cdot e^{i\pi})\widehat{\ }(-N)$ and $(p \cdot e^{-i\pi})\widehat{\ }(N)$)

$$(p \cdot e^{i\pi})\widehat{\ }(-N){:}\ \int_{|t|<\delta} \left(\frac{f(x-t) - f(x)}{2i\sin(\pi t)} e^{i\pi t} \right) e^{-2\pi i(-N)t}\, dt = \int_{|t|<\delta} \left(\frac{f(x-t) - f(x)}{2i\pi t} e^{i\pi t} \right) e^{-2\pi i(-N)t}\, dt$$

and

$$(p \cdot e^{-i\pi})\widehat{\ }(N){:}\ \int_{|t|<\delta} \left(\frac{f(x-t) - f(x)}{2i\sin(\pi t)} e^{-i\pi t} \right) e^{-2\pi i(N)t}\, dt = \int_{|t|<\delta} \left(\frac{f(x-t) - f(x)}{2i\pi t} e^{-i\pi t} \right) e^{-2\pi i(N)t}\, dt$$

From this we can pull out the $1/2i\pi$ to get

$$\frac{1}{2i\pi} \int_{|t|<\delta} \left(\frac{f(x-t) - f(x)}{t} e^{i\pi t} \right) e^{-2\pi i(-N)t}\, dt \text{ and } \frac{1}{2i\pi} \int_{|t|<\delta} \left(\frac{f(x-t) - f(x)}{t} e^{-i\pi t} \right) e^{-2\pi i(N)t}\, dt$$

To show that these pieces are integrable we look at

$$\left|\int_{|t|<\delta} \frac{f(x-t)-f(x)}{t} e^{\pm i\pi t}\, dt\right| \leq \int_{|t|<\delta} \left|\frac{f(x-t)-f(x)}{t} e^{\pm i\pi t}\right| dt$$

From here we see that $|e^{\pm i\pi t}| = \sqrt{(\cos(\pm\pi t))^2 + (\sin(\pm\pi t))^2} = \sqrt{1} = 1$ therefore

$$\int_{|t|<\delta} \left|\frac{f(x-t)-f(x)}{t}\right| \cdot \left|e^{\pm i\pi t}\right| dt = \int_{|t|<\delta} \left|\frac{f(x-t)-f(x)}{t}\right| dt$$

which by the hypothesis of Theorem 2.1

$$\text{Hypothesis: } \int_{|t|<\delta} \left|\frac{f(x+t)-f(x)}{t}\right| dt < \infty$$

converge. Therefore we have shown that these integrands are integrable so $(p \cdot e^{i\pi})\widehat{\ }(-N)$ and $(p \cdot e^{-i\pi})\widehat{\ }(N)$ are correct and their limits can be taken. Finally we see by the Riemann-Lebesgue Lemma (since $f \in L^1(\mathbb{T})$) that the first integral evaluates to 0 as we take the limit as $N \to \infty$.
Hence, since both integrals tend to 0 as we take $N \to \infty$ we have as desired

$$\lim_{N\to\infty} S_N f(x) - f(x) = 0$$
$$\lim_{N\to\infty} S_N f(x) = f(x).$$

□

We now prove Theorem 2.2

Proof of Theorem 2.2. We start by letting f be a function of bounded variation in a neighborhood of x. To understand this proof, we must unpack what it means to be a function of bounded variation. We start by supposing that f is a function defined on $[a, b]$ and $a = t_0 < t_1 < \cdots < t_N = b$ is a partition of this interval. The variation of f on this partition is defined by

$$\sum_{j=1}^{N} |f(t_j) - f(t_{j-1})|.$$

We also note that the total variation of f on $[a, x]$ (where $a \leq x \leq b$) is defined by

$$T_f(a, x) = \sup \sum_{j=1}^{N} |f(t_j) - f(t_{j-1})|.$$

where the supremum is over all partitions of $[a, x]$. The function f is said to be of bounded variation if the variations of f over all partitions are bounded, aka there exists $M < \infty$ so that

$$\sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| \leq M$$

for all partitions $a = t_0 < t_1 < \cdots < t_N = b$. This is the same as saying that the total variation of f, $T_f(a,b)$, is finite.
We will now show that a real-valued function f on $[a,b]$ is of bounded variation if and only if f is the difference of two increasing bounded functions.
To do this we need to prove a lemma which uses the following two definitions:
The positive variation of f on $[a,x]$ is

$$P_f(a,x) = \sup \sum_{(+)} f(t_j) - f(t_{j-1}),$$

where the sum is over all j such that $f(t_j) \geq f(t_{j-1})$, and the supremum is over all partitions of $[a,x]$.
The negative variation of f on $[a,x]$ is

$$N_f(a,x) = \sup_{(-)} - [f(t_j) - f(t_{j-1})],$$

where the sum is over all j such that $f(t_j) \leq f(t_{j-1})$, and the supremum is over all partitions of $[a,x]$.

Lemma 2.5 (Lemma 3.2 from Chapter 3 of Stein & Shakarchi Vol. 3)**.** *Suppose f is real-valued and of bounded variation on $[a,b]$. Then for all $a \leq x \leq b$ one has*

$$f(x) - f(a) = P_f(a,x) - N_f(a,x),$$

and

$$T_f(a,x) = P_f(a,x) + N_f(a,x).$$

Proof of Lemma 2.5. Given $\epsilon > 0$ there exists a partition $a = t_0^1 < t_1^1 < \cdots < t_N^1 = x$ of $[a,x]$ (we denote as $\{t_n^1\}$), such that

$$\left| P_f(a,x) - \sum_{(+)_{t_n^1}} f(t_j) - f(t_{j-1}) \right| < \epsilon/2$$

It is important to note that this sum is only summing when $t_j^1 \geq t_{j-1}^1$ for $\{t_n^1\}$ meaning that each term of the sum is non-negative.
There also exists another partition $a = t_0^2 < t_1^2 < \cdots < t_N^2 = x$ of $[a,x]$ (we denote as $\{t_n^2\}$), such that

$$\left| N_f(a,x) - \sum_{(-)_{t_n^2}} - [f(t_j) - f(t_{j-1})] \right| < \epsilon/2$$

where this sum is only summing when $t_j^2 \leq t_{j-1}^2$ for $\{t_n^2\}$ meaning that each term of the sum is non-negative.
From here, with the remembrance that $P_f(a,x)$ is always greater than $\sum_{(+)} f(t_j) - f(t_{j-1})$ and $N_f(a,x)$ is always greater than $\sum_{(-)} - [f(t_j) - f(t_{j-1})]$ no matter the partition, we state

that a refinement of a partition can only increase the amount of variation. This can be seen by the following:
Given the partition P of $[a,b]$ given by $a = t_0 < t_1 < \cdots < t_N = b$ (denoted as $\{t_n\}$) and a refinement of P which we will denote as $\tilde{P}$ of $[a,b]$ given by $a = \tilde{t}_0 < \tilde{t}_1 < \cdots < \tilde{t}_M = b$ (denoted as $\{\tilde{t}_n\}$) we see that we can expand out the variation of the partition of P until we attain the variation of partition $\tilde{P}$. Note that a refinement of P means that $\{\tilde{t}_n\}$ contains all elements of the sequence $\{t_n\}$ plus more elements; we name these extra elements $\{\tilde{t}_{n_k}\}$. If there exists any elements of $\{\tilde{t}_{n_k}\}$ between elements of $\{t_n\}$ in the sequence $\{\tilde{t}_n\}$ we do the following to obtain our desired inequality:

Variation of P: $\displaystyle\sum_{t_n} |f(t_j) - f(t_{j-1})| = |f(t_1) - f(t_0)| + |f(t_2) - f(t_1)| + \cdots + |f(t_N) - f(t_{N-1})|$

For sake of example let the first element of $\{\tilde{t}_{n_k}\}$ exist between t_1 and t_2.
This would mean that $\tilde{t}_0 = t_0$, $\tilde{t}_1 = t_1$, $\tilde{t}_2 = \tilde{t}_{n_0}$, and $\tilde{t}_3 = t_2$. Then we could do the following:
Variation of P: $= |f(t_1) - f(t_0)| + |f(t_2) - f(t_1)| + \cdots + |f(t_N) - f(t_{N-1})|$
$= |f(t_1) - f(t_0)| + \left|f(t_2) - f(\tilde{t}_{n_0}) + f(\tilde{t}_{n_0}) - f(t_1)\right| + \cdots + |f(t_N) - f(t_{N-1})|$
$\leq |f(t_1) - f(t_0)| + \left|f(\tilde{t}_{n_0}) - f(t_1)\right| + \left|f(t_2) - f(\tilde{t}_{n_0})\right| + \cdots + |f(t_N) - f(t_{N-1})|$

Doing this addition of the $\{\tilde{t}_{n_k}\}$ into the RHS creates the variation of $\tilde{P}$ meaning that

$$\text{Variation of } P\text{: } \sum_{t_n} |f(t_j) - f(t_{j-1})| \leq \sum_{\tilde{t}_n} |f(t_j) - f(t_{j-1})| \ : \text{Variation of } \tilde{P}$$

This endeavor allows us to refine our partitions for the positive variation and negative variation estimate, t_n^1 and t_n^2 respectively, together to create a single partition t_n^*. Since the refinement of partitions only increases the variation of our sums, $\sum_{(+)} f(t_j) - f(t_{j-1})$ and $\sum_{(-)} -[f(t_j) - f(t_{j-1})]$, and they are both less than $P_f(a,x)$ and $N_f(a,x)$ respectively, we see then that

$$\epsilon/2 > \left|P_f(a,x) - \sum_{(+)_{t_n^1}} f(t_j) - f(t_{j-1})\right| \geq \left|P_f(a,x) - \sum_{(+)_{t_n^*}} f(t_j) - f(t_{j-1})\right|$$

and

$$\epsilon/2 > \left|N_f(a,x) - \sum_{(-)_{t_n^2}} -[f(t_j) - f(t_{j-1})]\right| \geq \left|N_f(a,x) - \sum_{(-)_{t_n^*}} -[f(t_j) - f(t_{j-1})]\right|$$

So we have

$$\left|P_f(a,x) - \sum_{(+)_{t_n^*}} f(t_j) - f(t_{j-1})\right| < \epsilon/2 \text{ and } \left|N_f(a,x) - \sum_{(-)_{t_n^*}} -[f(t_j) - f(t_{j-1})]\right| < \epsilon/2$$

From here we see that

$$f(x) - f(a) = \sum_{(+)_{t_n^*}} f(t_j) - f(t_{j-1}) - \sum_{(-)_{t_n^*}} -[f(t_j) - f(t_{j-1})]$$

This is due to the fact that for our given partition, $\{t_n^*\}$, we have

$$f(x) - f(a) = \sum_{t_n^*} f(t_j) - f(t_{j-1})$$

Note that $t_0^* = a$ and the last element in t_n^* is x meaning that the sum is telescoping and leaves only the first and last elements of the sequence.
From here we see that we can separate the groups in the sum above into three parts: the positive values ($f(t_j) - f(t_{j-1}) > 0$), the negative values ($f(t_j) - f(t_{j-1}) < 0$), and the zero values ($f(t_j) - f(t_{j-1}) = 0$, also note that these terms in the sum have value 0). Giving us the following

$$\begin{aligned} f(x) - f(a) = \sum_{t_n^*} f(t_j) - f(t_{j-1}) &= \sum_{(+)_{t_n^*}} f(t_j) - f(t_{j-1}) + \sum_{(-)_{t_n^*}} f(t_j) - f(t_{j-1}) + \sum_{(0)_{t_n^*}} f(t_j) - f(t_{j-1}) \\ &= \sum_{(+)_{t_n^*}} f(t_j) - f(t_{j-1}) - \sum_{(-)_{t_n^*}} -[f(t_j) - f(t_{j-1})] \end{aligned}$$

From here

$$f(x) - f(a) - [P_f(a,x) - N_f(a,x)] = \sum_{(+)_{t_n^*}} f(t_j) - f(t_{j-1}) - \sum_{(-)_{t_n^*}} -[f(t_j) - f(t_{j-1})] - [P_f(a,x) - N_f(a,x)]$$

$$\begin{aligned} |f(x) - f(a) - [P_f(a,x) - N_f(a,x)]| &= \Big| \sum_{(+)_{t_n^*}} f(t_j) - f(t_{j-1}) - P_f(a,x) \\ &\quad + N_f(a,x) - \sum_{(-)_{t_n^*}} -[f(t_j) - f(t_{j-1})] \Big| \\ &\leq \left| P_f(a,x) - \sum_{(+)_{t_n^*}} f(t_j) - f(t_{j-1}) \right| \\ &\quad + \left| N_f(a,x) - \sum_{(-)_{t_n^*}} -[f(t_j) - f(t_{j-1})] \right| \\ &< \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

Therefore

$$|f(x) - f(a) - [P_f(a,x) - N_f(a,x)]| < \epsilon$$

proving the first identity.
Now for the second identity we start by taking any partition of $a = t_0 < t_1 < \cdots < t_N = x$ of $[a,x]$ we can write $\sum_{j=1}^{N} f(t_j) - f(t_{j-1})$ into its three pieces again as so (note again that

the third grouping, $(0)_{t_n}$, is when $f(t_j) = f(t_{j-1})$ making the summation of these 0)

$$\begin{aligned}\sum_{j=1}^{N} f(t_j) - f(t_{j-1}) &= \sum_{(+)_{t_n}} f(t_j) - f(t_{j-1}) + \sum_{(-)_{t_n}} f(t_j) - f(t_{j-1}) + \sum_{(0)_{t_n}} f(t_j) - f(t_{j-1}) \\ &= \sum_{(+)_{t_n}} f(t_j) - f(t_{j-1}) + \sum_{(-)_{t_n}} f(t_j) - f(t_{j-1})\end{aligned}$$

we see that the negative terms can be written as positive values since for them we have $f(t_{j-1}) \geq f(t_j)$ meaning that $-[f(t_j) - f(t_{j-1})]$ is positive for each term. This means that if we rewrite the second sum and then take absolute values on both sides after making all of the terms on the RHS positive, we are left with

$$\begin{aligned}\sum_{j=1}^{N} f(t_j) - f(t_{j-1}) &= \sum_{(+)_{t_n}} f(t_j) - f(t_{j-1}) - \sum_{(-)_{t_n}} -[f(t_j) - f(t_{j-1})] \\ \sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| &= \sum_{(+)_{t_n}} |f(t_j) - f(t_{j-1})| + \sum_{(-)_{t_n}} |-[f(t_j) - f(t_{j-1})]| \\ \sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| &= \sum_{(+)_{t_n}} f(t_j) - f(t_{j-1}) + \sum_{(-)_{t_n}} -[f(t_j) - f(t_{j-1})]\end{aligned}$$

Since this is true for any partition of $a = t_0 < t_1 < \cdots < t_N = x$ of $[a, x]$, and $P_f(a, x)$ and $N_f(a, x)$ are the supremum of the two terms on the RHS, we see that the LHS, $T_f(a, x)$, is less than or equal to $P_f(a, x) + N_f(a, x)$

$$\begin{gathered}\sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| = \sum_{(+)_{t_n}} f(t_j) - f(t_{j-1}) + \sum_{(-)_{t_n}} -[f(t_j) - f(t_{j-1})] \\ \Rightarrow \\ T_f(a, x) \leq P_f(a, x) + N_f(a, x).\end{gathered}$$

Now to show $T_f(a, x) \geq P_f(a, x) + N_f(a, x)$ we start by taking the partition of $a = t_{0_P} < t_{1_P} < \cdots < t_{N_P} = x$ of $[a, x]$, where $\{t_{n_P}\}$ is the partition for $P_f(a, x)$ and the partition of $a = t_{0_N} < t_{1_N} < \cdots < t_{N_N} = x$ of $[a, x]$, where $\{t_{n_N}\}$ is the partition for $N_f(a, x)$. From here we remember that taking a refinement of these partitions by combining them, calling it $\{t_{n_{PN}}\}$, can only increase the value of the sums since they are non-negative, therefore

$$P_{f_{t_{n_P}}}(a, x) + N_{f_{t_{n_N}}}(a, x) \leq P_{f_{t_{n_{PN}}}}(a, x) + N_{f_{t_{n_{PN}}}}(a, x)$$

From here we remember

$$\sum_{t_{n_{PN}}} f(t_j) - f(t_{j-1}) = \sum_{(+)_{t_{n_{PN}}}} f(t_j) - f(t_{j-1}) - \sum_{(-)_{t_{n_{PN}}}} -[f(t_j) - f(t_{j-1})]$$

and

$$\sum_{t_{n_{PN}}} |f(t_j) - f(t_{j-1})| = \sum_{(+)t_{n_{PN}}} |f(t_j) - f(t_{j-1})| + \sum_{(-)t_{n_{PN}}} |-[f(t_j) - f(t_{j-1})]|$$
$$= P_{f_{t_{n_{PN}}}}(a, x) + N_{f_{t_{n_{PN}}}}(a, x)$$

We now realize that $T_f(a, x)$ is defined to be the supremum over all possible partitions therefore $T_f(a, x) \geq \sum_{t_{n_{PN}}} |f(t_j) - f(t_{j-1})|$ meaning that we have our desired inequality

$$T_f(a, x) \geq \sum_{t_{n_{PN}}} |f(t_j) - f(t_{j-1})| = P_{f_{t_{n_{PN}}}}(a, x) + N_{f_{t_{n_{PN}}}}(a, x) \geq P_{f_{t_{n_P}}}(a, x) + N_{f_{t_{n_N}}}(a, x)$$

thus proving the second identity and concluding the proof of the lemma. □

Now that we have proved the lemma, we will prove the statement: a real-valued function f on $[a, b]$ is of bounded variation if and only if f is the difference of two increasing bounded functions (Theorem 3.3 from Chapter 3 of Stein & Shakarchi Vol. 3).

Proof of Theorem 3.3 from Chapter 3 of Stein & Shakarchi Vol. 3. We start by proving that the difference of two increasing bounded functions is of bounded variation. We see that each function, since they are increasing and bounded, have the following for any partition (let $|f_1(x)| \leq M_1$ and $|f_2(x)| \leq M_2$)

$$\sum_{j=1}^{N} |f_1(t_j) - f_1(t_{j-1})| = \sum_{j=1}^{N} f_1(t_j) - f_1(t_{j-1})$$
$$= f_1(b) - f_1(a) \leq 2M_1$$

and

$$\sum_{j=1}^{N} |f_2(t_j) - f_2(t_{j-1})| = \sum_{j=1}^{N} f_2(t_j) - f_2(t_{j-1})$$
$$= f_2(b) - f_2(a) \leq 2M_2$$

This means that for $f(x) = f_1(x) - f_2(x)$, for any partition, we have

$$\begin{aligned}
\sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| &= \sum_{j=1}^{N} \big| [f_1(t_j) - f_2(t_j)] - [f_1(t_{j-1}) - f_2(t_{j-1})] \big| \\
&= \sum_{j=1}^{N} \big| f_1(t_j) - f_1(t_{j-1}) + (-1) \cdot [f_2(t_j) - f_2(t_{j-1})] \big| \\
&\leq \sum_{j=1}^{N} |f_1(t_j) - f_1(t_{j-1})| + |f_2(t_j) - f_2(t_{j-1})| \\
&\leq 2M_1 + 2M_2 < \infty
\end{aligned}$$

meaning that $f(x)$ is a function of bounded variation.
Now we wish to show if $f(x)$ is of bounded variation, then it is a difference of two increasing bounded functions. We start by having $f_1(x) = P_f(a, x) + f(a)$ and $f_2(x) = N_f(a, x)$. Since $f(x)$ is of bounded variation $f(a)$ is finite. Also, $P_f(a, x)$ and $N_f(a, x)$ are both increasing functions by definition, this is due to the fact that as we take a larger interval (increase of x) we see that the two functions can only increase in output due to a new term added to the summation, for both functions, can only be greater than or equal to 0. This means that both $f_1(x)$ and $f_2(x)$ are increasing. To see that these functions are bounded we see that since $f(x)$ is of bounded variation we have $T_f(a, x) < \infty$ for any x. From Lemma 2.5 we then have

$$0 \leq P_f(a, x) + N_f(a, x) = T_f(a, x) < \infty \text{ for all } x$$

meaning that both $P_f(a, x)$ and $N_f(a, x)$ are bounded making $f_1(x)$ and $f_2(x)$ bounded.
Lastly, we see from Lemma 2.5

$$\begin{aligned} f(x) - f(a) &= P_f(a, x) - N_f(a, x) \\ f(x) &= (P_f(a, x) + f(a)) - N_f(a, x) \\ f(x) &= f_1(x) - f_2(x) \end{aligned}$$

Proving that the function $f(x)$ of bounded variation is the difference of $f_1(x)$ and $f_2(x)$, two increasing bounded functions. □

Now that we have proven that every function of bounded variation can be written as the difference of two increasing bounded functions, we may assume that f is increasing and bounded in a neighborhood of x. We do this so that we can prove the theorem for a monotonic (specifically increasing) and bounded function and then show that it holds for the difference between two monotonic and bounded functions implying that it holds for functions of bounded variation.
First we see

$$\begin{aligned} S_N f(x) &= \int_{-1/2}^{1/2} f(x-t) D_N(t)\, dt \\ &= \int_{-1/2}^{0} f(x-t) D_N(t)\, dt + \int_{0}^{1/2} f(x-t) D_N(t)\, dt \end{aligned}$$

We now do a change of variables for the first integral, $t = -v$ and $dt = -dv$ and note that $D_N(t)$ is an even function. This is due to the fact that $D_N(t)$ is equal to a sine function multiplied by a 1/sine function making it equal to an odd function times an odd function making $D_N(t)$ an even function. This means $D_N(-t) = D_N(t)$.

$$\begin{aligned} S_N f(x) &= \int_{1/2}^{0} f(x-(-v)) D_N(-v) \cdot (-1)\, dv + \int_{0}^{1/2} f(x-t) D_N(t)\, dt \\ &= (-1)(-1) \int_{0}^{1/2} f(x+v) D_N(v)\, dv + \int_{0}^{1/2} f(x-t) D_N(t)\, dt \end{aligned}$$

From here we relabel $v = t$ to achieve

$$S_N f(x) = \int_0^{1/2} [f(x-t) + f(x+t)] D_N(t)\, dt.$$

Now let $g(t) = f(x+t)$ and $h(t) = f(x-t)$. This means that if we take the limit as $t \to 0^+$ we have $g(0+) = f(x+)$ and $h(0+) = f(x-)$. Note that we are assuming that f is increasing and bounded and therefore g and h are increasing, monotonic, and bounded. Our goal now is to show

$$\lim_{N\to\infty} \int_0^{1/2} f(x+t) D_N(t)\, dt = \frac{1}{2} g(0+) \text{ and } \lim_{N\to\infty} \int_0^{1/2} f(x-t) D_N(t)\, dt = \frac{1}{2} h(0+)$$
$$\lim_{N\to\infty} \int_0^{1/2} g(t) D_N(t)\, dt = \frac{1}{2} g(0+) \text{ and } \lim_{N\to\infty} \int_0^{1/2} h(t) D_N(t)\, dt = \frac{1}{2} h(0+).$$

Proving this would complete the proof as we would have

$$\lim_{N\to\infty} S_N f(x) = \lim_{N\to\infty} \int_0^{1/2} [f(x-t) + f(x+t)] D_N(t)\, dt = \frac{1}{2} g(0+) + \frac{1}{2} h(0+) = \frac{1}{2}\left[f(x+) + f(x-)\right].$$

We start by proving this for $g(t)$. From here we use the proven facts that $\int_0^1 D_N(t)\, dt = 1$, the Dirichlet kernel is of period 1, and that the Dirichlet kernel is even($\int_{-a}^{a} even = 2\int_0^a even$).

$$\begin{aligned}
1 &= \int_0^1 D_N(t)\, dt = \int_{-1/2}^{1/2} D_N(t)\, dt \\
1 &= 2\int_0^{1/2} D_N(t)\, dt \\
\frac{1}{2} &= \int_0^{1/2} D_N(t)\, dt
\end{aligned}$$

So now we have, by substituting in,

$$\begin{aligned}
\lim_{N\to\infty} \int_0^{1/2} g(t) D_N(t)\, dt &= \int_0^{1/2} g(0+) D_N(t)\, dt \\
\lim_{N\to\infty} \int_0^{1/2} [g(t) - g(0+)] D_N(t)\, dt &= 0
\end{aligned}$$

Our new goal is to validate this, so we start by splitting up the integral. For some $\delta > 0$ we have

$$\int_0^{1/2} [g(t) - g(0+)] D_N(t)\, dt = \int_0^{\delta} [g(t) - g(0+)] D_N(t)\, dt + \int_{\delta}^{1/2} [g(t) - g(0+)] D_N(t)\, dt.$$

Just as before we see that this second integral goes to 0 by the Riemann-Lebesgue Lemma

$$\begin{aligned}
&\int_{\delta}^{1/2} [g(t) - g(0+)] D_N(t)\, dt = \int_{\delta}^{1/2} \frac{g(t) - g(0+)}{2i\sin(\pi t)} \cdot [e^{i\pi 2Nt} \cdot e^{i\pi t} - e^{-i\pi 2Nt} \cdot e^{-i\pi t}]\, dt \\
&= \int_{\delta}^{1/2} \left(\frac{g(t) - g(0+)}{2i\sin(\pi t)} e^{i\pi t}\right) e^{-2\pi i(-N)t}\, dt - \int_{\delta}^{1/2} \left(\frac{g(t) - g(0+)}{2i\sin(\pi t)} e^{-i\pi t}\right) e^{-2\pi i(N)t}\, dt \\
&= (h \cdot e^{i\pi})\widehat{\ }(-N) - (h \cdot e^{-i\pi})\widehat{\ }(N)
\end{aligned}$$

where

$$h(t) = \frac{g(t) - g(0+)}{2i\sin(\pi t)} \chi_{\{\delta \le t < 1/2\}}(t).$$

Therefore by the Riemann-Lebesgue Lemma (since $g \in L^1(\mathbb{T})$) and by the fact that our denominator is bounded by δ avoiding the problem of dividing by 0 we see that this integral evaluates to 0 as we take the limit as $N \to \infty$.
We now turn our attention to the first integral $\int_0^\delta [g(t) - g(0+)] D_N(t)\, dt$. First let us create a new function

$$\tilde{g}(t) = \begin{cases} 0, & \text{if } t = 0 \\ g(t) - g(0+), & \text{for } t \in (0, \delta] \end{cases}$$

We do this to make sure that our new function $\tilde{g}(t)$ is non-negative and monotonically increasing (and bounded since $g(t)$ is bounded). This change makes sure of this since for $t \in (0, \delta]$ we know $g(t)$ is increasing so $g(0+)$ is the infimum on this interval making $g(t) - g(0+) \ge 0$ for our interval.
Next, given $\epsilon > 0$, choose $\delta > 0$ such that $\tilde{g}(t) < \epsilon$ if $0 < t < \delta$. This is completely true as $\tilde{g}(t)$ is an increasing (monotonic) function. This also means that $|g(t) - g(0+)| < \epsilon$ if $0 < t < \delta$.
We then prove a version of the second mean value theorem for integrals that states:

Let $\phi(x)$ be continuous on $[a, b]$ and $h(x)$ monotonic on $[a, b]$, and let $A, B \in \mathbb{R}$ such that $A \le h(a+) \le h(b-) \le B$. Then there exists a $c \in [a, b]$ such that

$$\int_a^b h(x)\phi(x)\, dx = B \int_c^b \phi(x)\, dx + A \int_a^c \phi(x)\, dx$$

We will prove this then prove a corollary from this theorem.

Proof of a version of the Second Mean Value Theorem for Integrals. Before we start this very complicated proof we will prove the following theorems in order which we will use and refer to later in this proof and later in this paper.
These theorems are (in order of proving) Bolzano Theorem, sign-preserving property for continuous functions, Intermediate Value Theorem, Boundedness Theorem, Nested Intervals Theorem, Monotone Convergence Theorem, and Extreme Value Theorem. Note that all of these theorems have something to do with $\phi(x)$ being continuous on a closed interval $[a, b]$.
First:

Bolzano Theorem:
If ϕ is continuous on $[a, b]$ and $\phi(a) < 0 < \phi(b)$, then there is some $c \in [a, b]$ such that $\phi(c) = 0$ (we also prove this for $\phi(a) > 0 > \phi(b)$ after).

Proof of Bolzano Theorem. We start by assuming $\phi(x)$ is continuous on $[a, b]$ and $\phi(a) < 0 < \phi(b)$. From here we note that there may exists multiple $x \in [a, b]$ such that $\phi(x) = 0$, but we will find the largest x in the interval.
Let $S = \{x \in [a, b] : \phi(x) \le 0\}$. We see that $S \ne \emptyset$ since $\phi(a) < 0$. Note that S is bounded above since $S \subseteq [a, b]$ therefore S has a least upper bound c. We will show $\phi(c) = 0$.

To do this we will use and prove the following lemma:

Sign-preserving property for continuous functions:
Let ϕ be continuous at a and suppose $\phi(a) \neq 0$. Then there exists an interval $(a-\delta, a+\delta)$ about a which we call $B(a, \delta)$ in which ϕ has the same sign as $\phi(a)$.

Proof of sign-preserving property for continuous functions. Suppose $\phi(a) > 0$. From continuity at a we have for every $\epsilon > 0$, there exists a $\delta > 0$ such that for $|x-a| < \delta \Rightarrow |\phi(x) - \phi(a)| < \epsilon$. Aka for $a - \delta < x < a + \delta$ we have $\phi(a) - \epsilon < \phi(x) < \phi(a) + \epsilon$.
Take $\epsilon = \frac{\phi(a)}{2}$, then there exists $\delta > 0$ such that $|x-a| < \delta \Rightarrow |\phi(x) - \phi(a)| < \frac{\phi(a)}{2}$. Aka for $a - \delta < x < a + \delta$ we have $\phi(a) - \frac{\phi(a)}{2} < \phi(x) < \phi(a) + \frac{\phi(a)}{2}$, aka (since we supposed $\phi(a) > 0$) $0 < \frac{\phi(a)}{2} < \phi(x) < \frac{3\phi(a)}{2}$.
Therefore we have shown $\phi(x) > 0$ on $(a-\delta, a+\delta)$ and $\phi(a)$ and $\phi(x)$ have the same sign.
Now suppose $\phi(a) < 0$. From continuity at a we have for every $\epsilon > 0$, there exists a $\delta > 0$ such that $(a-\delta, a+\delta) \Rightarrow \phi(a) - \epsilon < \phi(x) < \phi(a) + \epsilon$.
Now take $\epsilon = -\frac{\phi(a)}{2}$ which is positive since $\phi(a) < 0$. Therefore for the δ connected to this ϵ we have for $x \in (a-\delta, a+\delta) \Rightarrow \phi(a) - \left(-\frac{\phi(a)}{2}\right) < \phi(x) < \phi(a) + \left(-\frac{\phi(a)}{2}\right)$. Aka $\frac{3\phi(a)}{2} < \phi(x) < \frac{\phi(a)}{2} < 0$.
Therefore we have shown $\phi(x) < 0$ on $(a-\delta, a+\delta)$ and $\phi(a)$ and $\phi(x)$ have the same sign.
This concludes the proof of the sign-preserving property for continuous functions. $\square$

Now we wish to show that $\phi(c) = 0$. We know that there are only three options:

$$\phi(c) > 0, \phi(c) < 0, \text{ or } \phi(c) = 0.$$

If $\phi(c) > 0$, then, by the sign-preserving property for continuous functions, there exists an interval $(c-\delta, c+\delta)$ or $(c-\delta, c]$ if $c = b$, in which $\phi > 0$ for the interval. Note that this means nothing to the right of $c - \delta$ can be in S making $c - \delta$ an upper bound for S, but we chose c to be the least upper bound for S and since $c - \delta < c$ we have arrived at a contradiction and $\phi(c) \not> 0$.
If $\phi(c) < 0$ then, by the sign-preserving property for continuous functions, there exists an interval $(c-\delta, c+\delta)$ or $[c, c+\delta)$ if $c = a$, in which $\phi < 0$ for the interval. This means that $\phi(x) < 0$ for some $x > c$ which gives us a contradiction since c is the least upper bound meaning that $\phi(c) \not< 0$.
This means that we must have $\phi(c) = 0$ and $\phi(a) < \phi(c) < \phi(b)$ since $\phi(a) < 0$ and $\phi(b) > 0$.
Also note that $c \in [a, b]$ since $c \in S$ and $S \subseteq [a, b]$.
Also note that this holds for $\phi(a) > 0 > \phi(b)$. We prove this as follows:
Given $\phi(x)$ continuous on $[a, b]$ and $\phi(a) > 0 > \phi(b)$ we define a function $g(x) = -\phi(x)$. From here we see that $g(a) < 0 < g(b)$ and we will prove that $g(x)$ is continuous on $[a, b]$. To show this, we need to prove that for any $x, y \in [a, b]$ we have for every $\epsilon > 0$, there exists a $\delta > 0$ such that for $|x-y| < \delta$ we have $|g(x) - g(y)| < \epsilon$. Now notice that $\phi(x)$ is continuous on $[a, b]$ so we have for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|x-y| < \delta \Rightarrow |\phi(x) - \phi(y)| < \epsilon$. Notice that if we factor out a -1 from the inside of the absolute values, which will then be turned into 1 by the absolute values and not affect the end value of the expression of the absolute values, we see that it is the same as $|\phi(y) - \phi(x)| < \epsilon$. From here we see that $|g(x) - g(y)| = |-\phi(x) - (-\phi(y))| = |\phi(y) - \phi(x)| < \epsilon$. This means that from the continuity

of $\phi(x)$ we have proven that $g(x)$ is continuous on $[a, b]$.
Now since $g(x)$ is continuous on $[a, b]$ and $g(a) < 0 < g(b)$ we see from above that there exists a $c \in [a, b]$ such that $g(c) = 0$. Hence, since $g(x) = -\phi(x) \Rightarrow g(c) = -\phi(c) = 0 = \phi(c)$, we have produced a $c \in [a, b]$ where $\phi(c) = 0$ proving the theorem for a $\phi(x)$ where $\phi(x)$ is continuous on $[a, b]$ and $\phi(a) > 0 > \phi(b)$.
This concludes the proof of Bolzano Theorem. □

We now move on to proving the Intermediate Value Theorem:

Intermediate Value Theorem:
If $\phi(x)$ is continuous on $[a, b]$ and k is any number between $\phi(a)$ and $\phi(b)$ then there is at least one number c between a and b such that $\phi(c) = k$.

Proof of Intermediate Value Theorem. If $\phi(a) = k$, then $c = a$, and if $\phi(b) = k$, then $c = b$. If neither of these are the case then we have either $\phi(a) < k < \phi(b)$ or $\phi(b) < k < \phi(a)$.
In either case we subtract k from all expressions giving

$$\phi(a) - k < 0 < \phi(b) - k \text{ and } \phi(b) - k < 0 < \phi(a) - k$$

Now define $g(x) = \phi(x) - k$. We will prove that $g(x)$ is continuous on $[a, b]$. To prove that $g(x)$ is continuous on $[a, b]$ we must have for any $x, y \in [a, b]$ the case that for every $\epsilon > 0$, there exists a $\delta > 0$ such that for $|x - y| < \delta \Rightarrow |g(x) - g(y)| < \epsilon$. From here realize that $|g(x) - g(y)| = |\phi(x) - k - (\phi(y) - k)| = |\phi(x) - \phi(y)|$. Since $\phi(x)$ is continuous on $[a, b]$ we have that for any $x, y \in [a, b]$ the following holds: for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|x - y| < \delta \Rightarrow |\phi(x) - \phi(y)| < \epsilon$. This means that we have $|g(x) - g(y)| < \epsilon$ for $|x - y| < \delta$ for any $x, y \in [a, b]$ proving that $g(x)$ is continuous on $[a, b]$.
Now that we've proven that $g(x)$ is continuous on $[a, b]$ we see that both cases above allow us to use Bolzano Theorem which means that there exists a $c \in [a, b]$ such that $g(c) = 0$. This means that for both

$$g(a) < 0 < g(b) \text{ and } g(a) > 0 > g(b),$$

we have found them each a c in their own respective intervals $[a, b]$. This means that $0 = g(c) = \phi(c) - k \Rightarrow \phi(c) = k$. This proves that for $\phi(x)$ continuous on $[a, b]$ and k being a number between $\phi(a)$ and $\phi(b)$, we have found a number $c \in [a, b]$ such that $\phi(c) = k$.
This concludes the proof of the Intermediate Value Theorem. □

From here we prove the Boundedness Theorem:

Boundedness Theorem:
If $\phi(x)$ is continuous on $[a, b]$ then $\phi(x)$ is bounded on $[a, b]$, i.e. there exists an $M < \infty$ such that $|\phi(x)| \leq M$ for all $x \in [a, b]$.

Proof of Boundedness Theorem. Note that inside this proof we will prove the Nested Interval Theorem and the Monotone Convergence Theorem.
To prove the Boundedness Theorem we will do a proof by contradiction. This means that we will assume that $\phi(x)$ is continuous on $[a, b]$ and unbounded on $[a, b]$.
We start by splitting [a,b] into two closed halves, since ϕ is unbounded on $[a, b]$ we have that ϕ must be unbounded on one of these two halves as well. Whichever half this is, label it $[a_1, b_1]$

here note that $a_1 \in \{a, a + \frac{b-a}{2}\}$ and we have the length of this new interval $|b_1 - a_1| = \frac{b-a}{2}$. From here we split $[a_1, b_1]$ into two halves and since ϕ is unbounded on $[a_1, b_1]$ we see that it must be unbounded on one of these two new halves. Label the half where ϕ is unbounded as $[a_2, b_2]$. Note that $a_2 \in \{a_1, a_1 + \frac{b_1-a_1}{2}\}$ and the length of $[a_2, b_2]$ is $|b_2 - a_2| = \frac{b-a}{2^2}$. We repeat these steps constructing $[a_{n+1}, b_{n+1}]$ from $[a_n, b_n]$ noting that $a_{n+1} \in \{a_n, a_n + \frac{b_n-a_n}{2}\}$, ϕ is unbounded on each interval, $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$ for all $n \in \mathbb{N}$, and the length of $[a_{n+1}, b_{n+1}]$ is $|b_{n+1} - a_{n+1}| = \frac{b-a}{2^{n+1}}$. Also note that the length of $[a_n, b_n]$ goes to 0 as $n \to \infty$ because $\frac{b-a}{2^n} \to 0$ as $n \to \infty$.
From here we use and prove the Nested Interval Theorem:

Nested Interval Theorem:
Given a sequence of nested closed intervals $[a_n, b_n] \in \mathbb{R}$ where $a_n \leq a_{n+1}$ and $b_n \geq b_{n+1}$ for all $n \in \mathbb{N}$, and $\lim_{n\to\infty} (b_n - a_n) = 0$, there exists a unique point $x \in \mathbb{R}$ such that $x \in [a_n, b_n]$ for all $n \in \mathbb{N}$.

Proof. We note that we will prove and use the Monotone Convergence Theorem in this proof. We start by noticing that $a_n \leq b_1$ for all $n \in \mathbb{N}$ making the sequence $\{a_n\}$ bounded above by b_1. Also we have that $b_n \geq a_1$ for all $n \in \mathbb{N}$ making the sequence $\{b_n\}$ bounded below by a_1. From the premise we have that $\{a_n\}$ is a non-decreasing, $a_n \leq a_{n+1}$ for all $n \in \mathbb{N}$, and bounded above sequence and $\{b_n\}$ is a non-increasing, $b_n \geq b_{n+1}$ for all $n \in \mathbb{N}$, and bounded below sequence.
We now wish to use and prove the Monotone Convergence Theorem:

Monotone Convergence Theorem:
If $\{a_n\}$ is a monotonically increasing sequence that is bounded above, then $\{a_n\}$ converges to a value, specifically its supremum/least upper bound. If $\{b_n\}$ is a monotonically decreasing sequence that is bounded below, then $\{b_n\}$ converges to a value, specifically its infimum/greatest lower bound.

Proof of Monotone Convergence Theorem. We start by proving if $\{a_n\}$ is a monotonically increasing sequence that is bounded above, then $\{a_n\}$ converges to a value, specifically its supremum/least upper bound.
Given that $\{a_n\}$ is a monotone increasing sequence, it means

$$a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n \leq \ldots$$

Also, since the sequence is bounded above, this means that there exists an $M < \infty$ such that $a_n < M$ for all $n \in \mathbb{N}$. Since the sequence is bounded above we have, by the completeness property of the real numbers, that the set $\{a_n\}$ has a least upper bound which we will call L that is in $\mathbb{R}$. This means $L = \sup\{a_n : n \in \mathbb{N}\}$. We now need to show that $\{a_n\}$ converges to L. To do this we will prove that for every $\epsilon > 0$, there exists an integer N such that for all $n \geq N$ we have $|a_n - L| < \epsilon$.
Since L is the least upper bound of $\{a_n\}$ we have that for any $\epsilon > 0$, $L - \epsilon$ is not an upper bound of $\{a_n\}$. This means that there exists an integer N such that $a_N > L - \epsilon$ and since the sequence is increasing we have

$$a_n \geq a_N > L - \epsilon$$

for all $n \geq N$. Again, since L is an upper bound of $\{a_n\}$, for all n, meaning $a_n \leq L$, we achieve the following by combining the two inequalities

$$L - \epsilon < a_n \leq L$$

Meaning that we have (since $L \geq a_n$ for all $n \in \mathbb{N} \Rightarrow \; |a_n - L| = L - a_n$)

$$\begin{aligned} L - \epsilon < a_n &\Rightarrow \; L - a_n < \epsilon \\ &\Rightarrow \; |a_n - L| < \epsilon. \end{aligned}$$

Hence we have proven that for every $\epsilon > 0$, there exists an integer N such that for all $n \geq N$ we have $|a_n - L| < \epsilon$ meaning that $\{a_n\}$ converges to L. Proving that given a monotone increasing sequence, $\{a_n\}$, that is bounded above we see that $\{a_n\}$ converges to a value, specifically its supremum/least upper bound.
Now we will prove if $\{b_n\}$ is a monotonically decreasing sequence that is bounded below, then $\{b_n\}$ converges to a value, specifically its infimum/greatest lower bound.
Given that $\{b_n\}$ is a monotonically decreasing sequence we have

$$b_1 \geq b_2 \geq b_3 \geq \cdots \geq b_n \geq \ldots$$

Since the sequence is bounded below there exists an $-\infty < M < \infty$ such that $b_n > M$ for all $n \in \mathbb{N}$. Since the sequence is bounded below we have, by the completeness property of the real numbers, that the set $\{b_n\}$ has a greatest lower bound which we will denote as L that is in $\mathbb{R}$. This means that $L = \inf\{b_n : n \in \mathbb{N}\}$. We now need to show that $\{b_n\}$ converges to L. To do this we will prove that for every $\epsilon > 0$, there exists an integer N such that for all $n \geq N$ we have $|b_n - L| < \epsilon$.
Since L is the greatest lower bound of $\{b_n\}$ we have that for any $\epsilon > 0$, $L + \epsilon$ is not a lower bound of $\{b_n\}$. This means that there exists an integer N such that $b_N < L + \epsilon$ and since the sequence is decreasing we have

$$b_n \leq b_N < L + \epsilon$$

for all $n \geq N$. Again, since L is a lower bounded of $\{b_n\}$, for all $n \in \mathbb{N}$, meaning that $b_n \geq L$ for all $n \in \mathbb{N}$, we achieve the following combining the two inequalities

$$L + \epsilon > b_n \geq L$$

Meaning that we have (since $L \leq b_n$ for all $n \in \mathbb{N} \Rightarrow \; |b_n - L| = b_n - L$)

$$\begin{aligned} L + \epsilon > b_n &\Rightarrow \; b_n - L < \epsilon \\ &\Rightarrow \; |b_n - L| < \epsilon. \end{aligned}$$

Hence we have proven that for every $\epsilon > 0$, there exists an integer N such that for all $n \geq N$ we have $|b_n - L| < \epsilon$ meaning that $\{b_n\}$ converges to L. Proving that given a monotone decreasing sequence, $\{b_n\}$, that is bounded below we see that $\{b_n\}$ converges to a value, specifically its infimum/greatest lower bound.
This concludes the proof of the Monotone Convergence Theorem. $\square$

Now that we have proven the Monotone Convergence Theorem we see that $\{a_n\}$, being a monotonically increasing sequence that is bounded above, converges to a value, specifically its supremum/least upper bound, let it be α. Also we that $\{b_n\}$, being a monotonically decreasing sequence that is bounded below, converges to a value specifically its infimum/greatest lower bound, let it be β.
Again, from the premise we have that $\lim_{n\to\infty}(b_n - a_n) = 0$. Note that $\lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \beta - \alpha$. Therefore $\beta - \alpha = 0 \Rightarrow\ \beta = \alpha$.
Let $x = \beta = \alpha$. We then have $a_n \leq x \leq b_n$ for all $n \in \mathbb{N}$ from the limits above. From this we see that $x \in [a_n, b_n]$ for all $n \in \mathbb{N}$.
Now all we have to do is prove that x is unique. To do this we assume that there exists a $y \in [a_n, b_n]$ for all $n \in \mathbb{N}$. This means that we would have $a_n \leq y \leq b_n$ for all $n \in \mathbb{N}$. As $n \to \infty$ we see that $a_n \to \alpha = x$ and $b_n \to \beta = x$. There we see that by taking the limit as $n \to \infty$ we get $x \leq \lim_{n\to\infty} y \leq x$. Therefore by the Squeeze Theorem we have that $y = x$ must be the case validating the uniqueness of x.
This concludes the proof of the Nested Interval Theorem. $\square$

Now that we have proven the Nested Interval Theorem we see that our constructed intervals $[a_n, b_n]$ align with the necessary properties to use the Nested Interval Theorem. We see that by constructing $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$, we have $a_n \leq a_{n+1}$ and $b_n \geq b_{n+1}$ for all $n \in \mathbb{N}$ and we have that the intervals, aka $b_n - a_n$, go to 0 as $n \to \infty$ (this is because $b_n - a_n = \frac{b-a}{2^n}$ and $\frac{b-a}{2^n} \to 0$ as we take the limit of $n \to \infty$). It is also very important to remember that ϕ is unbounded on $[a_n, b_n]$ for every $n \in \mathbb{N}$.
This means that all of these intervals intersect at a single unique point which we will call α. From here we remember that $\alpha \in [a, b]$, therefore by the continuity of $\phi(x)$ at α (which we get from $\phi(x)$ being continuous on $[a, b]$) we have that $\phi(\alpha)$ is defined and is in $\mathbb{R}$ meaning that $\phi(\alpha) \neq \pm\infty$. Furthermore, since $\phi(x)$ is continuous at α we have that for any $\epsilon > 0$, there exists a $\delta > 0$ such that for an interval $(\alpha - \delta, \alpha + \delta)$ we get $|\phi(x) - \phi(\alpha)| < \epsilon$ for x in this interval. This means that we can let $\epsilon = 1$ and there exists a $\delta > 0$ that makes this true. Since $\epsilon = 1$ we see $|\phi(x) - \phi(\alpha)| < 1$.
Now by the triangle inequality we have

$$\begin{aligned}
|\phi(x)| &= |\phi(x) - \phi(\alpha) + \phi(\alpha)| \\
&\leq |\phi(x) - \phi(\alpha)| + |\phi(\alpha)| \\
&\leq \epsilon + |\phi(\alpha)| \\
&= 1 + |\phi(\alpha)|
\end{aligned}$$

Since we know from earlier that $\phi(\alpha)$ is some finite value in $\mathbb{R}$ we see that $|\phi(x)| \leq 1 + |\phi(\alpha)| < \infty$ for the set interval $(\alpha - \delta, \alpha + \delta)$. This means that $\phi(x)$ is bounded on $(\alpha - \delta, \alpha + \delta)$, but since $\delta > 0$ is fixed and $[a_n, b_n] \to \alpha$ as $n \to \infty$ and $|b_n - a_n| = \frac{b-a}{2^n} \to 0$ as $n \to \infty$ we see that there exists a finite $n' \in \mathbb{N}$ such that $[a_{n'}, b_{n'}] \subseteq (\alpha - \delta, \alpha + \delta)$. But this is gives us a contradiction, since $\phi(x)$ is bounded on the interval $(\alpha - \delta, \alpha + \delta)$ but $\phi(x)$ is unbounded on $[a_{n'}, b_{n'}]$ since $\phi(x)$ is unbounded on $[a_n, b_n]$ for all $n \in \mathbb{N}$. This means that $\phi(x)$ must be bounded on $[a, b]$ given that $\phi(x)$ is continuous on $[a, b]$.
This concludes the proof of the Boundedness Theorem. $\square$

Now that we have proven the Boundedness Theorem we will prove the Extreme Value Theorem.

Extreme Value Theorem:
If $\phi(x)$ is continuous on $[a,b]$, then there are numbers $u, v \in [a,b]$ such that $\phi(u) \leq \phi(x) \leq \phi(v)$ for all $x \in [a,b]$. Moreover

$$\phi(v) = \max\{\phi(x) : x \in [a,b]\} = \sup\{\phi(x) : x \in [a,b]\}$$
$$\phi(u) = \min\{\phi(x) : x \in [a,b]\} = \inf\{\phi(x) : x \in [a,b]\}$$

Proof of Extreme Value Theorem. We start by using the Boundedness Theorem. Since $\phi(x)$ is continuous on $[a,b]$ we have that $\phi(x)$ is bounded on $[a,b]$. This means that the range of ϕ is a bounded set of numbers and therefore has a supremum/least upper bound which we will call p and an infimum/great lower bound which we will call q.
We start by proving that for some $v \in [a,b]$ we have $\phi(v) = p$. Assume $\phi(v) \neq p$ for any $v \in [a,b]$, then define $g(x) = p - \phi(x)$. Since p is the supremum of $\phi(x)$ on $[a,b]$ we have that $g(x) > 0$ for all $x \in [a,b]$. We start by proving that $g(x)$ is continuous on $[a,b]$ and then we prove $\frac{1}{g(x)}$ is continuous on $[a,b]$.
To see that $g(x)$ is continuous on $[a,b]$ we must prove that for every $\epsilon > 0$, there exists a $\delta > 0$ such that for any $x, y \in [a,b]$ where $|x-y| < \delta$ we get $|g(x) - g(y)| < \epsilon$. Note that $|g(x) - g(y)| = |p - \phi(x) - (p - \phi(y))| = |\phi(y) - \phi(x)|$. Now we bring in the fact that $\phi(x)$ is continuous on $[a,b]$ which means that we have for every $\epsilon > 0$, there exists a $\delta > 0$ such that for any $x, y \in [a,b]$ where $|x-y| < \delta$ we get $|\phi(x) - \phi(y)| < \epsilon$. Remember from an earlier proof that we can bring out a -1 inside the absolute values and not change the value of the expression, therefore $|\phi(y) - \phi(x)| = |\phi(x) - \phi(y)| < \epsilon$ which means that $|g(x) - g(y)| < \epsilon$ whenever $|x-y| < \delta$ for $x, y \in [a,b]$ that satisfy the δ inequality. Thus we have proven $g(x)$ is continuous.
Now we wish to prove that $\frac{1}{g(x)}$ is continuous when $g(x)$ is continuous and $g(x) \neq 0$. We do this by proving that for any $\xi \in [a,b]$ we have: if g is continuous at ξ and $g(\xi) \neq 0$ then $\frac{1}{g}$ is continuous at ξ.
We start this proof by taking $\epsilon = \frac{1}{2}|g(\xi)|$. Since $g(x)$ is continuous at ξ there exists a $\delta_1 > 0$ such that $|g(x) - g(\xi)| < \epsilon$ if $|x - \xi| < \delta_1$ then we have

$$|g(x)| = |g(x) - g(\xi) + g(\xi)| \leq |g(x) - g(\xi)| + |g(\xi)| < \epsilon + |g(\xi)| = \frac{3}{2}|g(\xi)| \text{ and}$$
$$|g(\xi)| = |g(\xi) - g(x) + g(x)| \leq |g(\xi) - g(x)| + |g(x)| < \epsilon + |g(x)| = \frac{1}{2}|g(\xi)| + |g(x)|$$
$$\Rightarrow$$
$$\frac{1}{2}|g(\xi)| \leq |g(x)|$$

This means we have $\frac{1}{2}|g(\xi)| \leq |g(x)| \leq \frac{3}{2}|g(\xi)|$. Inherently this means $0 \leq \frac{1}{2}|g(\xi)| \leq |g(x)| \leq$

$\frac{3}{2}|g(\xi)|$ which we use to deduce

$$
\begin{aligned}
&0 \leq \frac{1}{2}|g(\xi)| \leq |g(x)| \\
&0 \leq 1 \leq \frac{|g(x)|}{\frac{1}{2}|g(\xi)|} \\
&0 \leq \frac{1}{|g(x)|} \leq \frac{1}{\frac{1}{2}|g(\xi)|} \\
&\text{and} \\
&0 \leq |g(x)| \leq \frac{3}{2}|g(\xi)| \\
&0 \leq 1 \leq \frac{\frac{3}{2}|g(\xi)|}{|g(x)|} \\
&0 \leq \frac{1}{\frac{3}{2}|g(\xi)|} \leq \frac{1}{|g(x)|} \\
&\text{which gives} \\
&0 \leq \frac{1}{\frac{3}{2}|g(\xi)|} \leq \frac{1}{|g(x)|} \leq \frac{1}{\frac{1}{2}|g(\xi)|}
\end{aligned}
$$

Now with $\frac{1}{\frac{3}{2}|g(\xi)|} \leq \frac{1}{|g(x)|} \leq \frac{1}{\frac{1}{2}|g(\xi)|}$ we let ϵ be arbitrary. Since $g(x)$ is continuous at ξ there exists a $\delta > 0$ (which we can choose to be $\leq \delta_1$) such that $|g(x) - g(\xi)| < \epsilon_1 = \epsilon\frac{1}{2}|g(\xi)|^2$ when $|x - \xi| < \delta$. Now we prove that $\frac{1}{g(x)}$ is continuous at ξ.
To do this we need to prove that for every $\epsilon > 0$, there exists a $\delta > 0$, such that for $x \in [a, b]$ where $|x - \xi| < \delta$ we have $\left|\frac{1}{g(x)} - \frac{1}{g(\xi)}\right| < \epsilon$. From here we investigate $\left|\frac{1}{g(x)} - \frac{1}{g(\xi)}\right|$

$$\left|\frac{1}{g(x)} - \frac{1}{g(\xi)}\right| = \left|\frac{g(\xi) - g(x)}{g(x)g(\xi)}\right| = \left|\frac{g(x) - g(\xi)}{g(x)g(\xi)}\right|$$

Since $\frac{1}{|g(x)|} \leq \frac{1}{\frac{1}{2}|g(\xi)|}$ we have

$$\left|\frac{1}{g(x)} - \frac{1}{g(\xi)}\right| = \left|\frac{g(x) - g(\xi)}{g(x)g(\xi)}\right| \leq \frac{|g(x) - g(\xi)|}{\frac{1}{2}|g(\xi)|^2}$$

and $|g(x) - g(\xi)| < \epsilon_1$ so

$$\left|\frac{1}{g(x)} - \frac{1}{g(\xi)}\right| \leq \frac{|g(x) - g(\xi)|}{\frac{1}{2}|g(\xi)|^2} < \frac{\epsilon_1}{\frac{1}{2}|g(\xi)|^2} = \frac{\epsilon\frac{1}{2}|g(\xi)|^2}{\frac{1}{2}|g(\xi)|^2} = \epsilon$$

proving that $\frac{1}{g(x)}$ is continuous at ξ whenever $g(\xi) \neq 0$ and $g(x)$ continuous at ξ. Since ξ was arbitrarily in $[a, b]$ we have that $\frac{1}{g(x)}$ is continuous on $[a, b]$ whenever $g(x) \neq 0$ for $x \in [a, b]$ and $g(x)$ continuous on $[a, b]$.
Now back to the proof of the Extreme Value Theorem.
Since $g(x)$ is continuous on $[a, b]$ and $g(x) > 0$ for all $x \in [a, b]$ we see that $\frac{1}{g(x)}$ is continuous

on $[a,b]$. Since $\frac{1}{g(x)}$ is continuous on $[a,b]$ we have by the Boundedness Theorem that $\frac{1}{g(x)}$ is bounded on $[a,b]$. This means that, since $\frac{1}{g(x)} > 0$ on $[a,b]$ because both 1 and $g(x)$ are greater than 0 on $[a,b]$ and division of two positive terms is a positive output, we have $0 < \frac{1}{g(x)} < C < \infty$ for all $x \in [a,b]$ for our bound on $\frac{1}{g(x)}$ that we call C.
This means $0 < \frac{1}{g(x)} < C \Rightarrow\ 0 < \frac{1}{C} < g(x)$. Plugging in $g(x) = p - \phi(x)$ we get

$$0 < \frac{1}{C} < p - \phi(x) \Rightarrow\ 0 < \phi(x) < p - \frac{1}{C}.$$

But p is the supremum/least upper bound which means this is a contradiction. Hence we must have that $p = \phi(v)$ for some $v \in [a,b]$.
We now want to prove that for some $u \in [a,b]$ we have $\phi(u) = q$. Assume $\phi(u) \neq q$ for all $u \in [a,b]$ and define $g(x) = \phi(x) - q > 0$. Notice that $g(x)$ is continuous on $[a,b]$ because $\phi(x)$ is continuous on $[a,b]$. To prove that this $g(x)$ is continuous on $[a,b]$ we must have that for every $\epsilon > 0$, there exists a $\delta > 0$, such that for $x, y \in [a,b]$ for which $|x-y| < \delta$ we have $|g(x) - g(y)| < \epsilon$. Note that $|g(x) - g(y)| = |\phi(x) - q - (\phi(y) - q)| = |\phi(x) - \phi(y)|$. Since $\phi(x)$ is continuous on $[a,b]$ we see that for every $\epsilon > 0$, there exists a $\delta > 0$, such that for $x, y \in [a,b]$ for which $|x-y| < \delta$ we have $|\phi(x) - \phi(y)| < \epsilon$ therefore proving that $g(x)$ is continuous on $[a,b]$.
Now that we see that $g(x)$ is continuous on $[a,b]$ and $g(x) > 0$ for all $x \in [a,b]$ we have from earlier, which we proved, that $\frac{1}{g(x)}$ is also continuous on $[a,b]$. Since $\frac{1}{g(x)}$ is continuous on $[a,b]$ we have by the Boundedness Theorem that $\frac{1}{g(x)}$ is bounded on $[a,b]$. This means that there exists a $C < \infty$ such that $0 < \frac{1}{g(x)} < C$ for all $x \in [a,b]$. Note similarly as before we have that $\frac{1}{g(x)} > 0$ for all $x \in [a,b]$ since $g(x) > 0$ for all $x \in [a,b]$. This means that

$$0 < \frac{1}{g(x)} < C \Rightarrow\ 0 < \frac{1}{C} < g$$

Plugging back in $g(x) = \phi(x) - q$ gets us

$$0 < \frac{1}{C} < \phi(x) - q \Rightarrow\ 0 < \frac{1}{C} + q < \phi(x).$$

But this is a contradiction since q is our greatest lower bound and we have just proven that $\phi(x) > \frac{1}{C} + q$ for all $x \in [a,b]$. This means we must have $q = \phi(u)$ for some $u \in [a,b]$.
This concludes the Extreme Value Theorem. $\square$

We have now finally proved all of the prerequisite little proofs that we will use in this proof.

Now that we have the preliminaries out of the way:
We start by defining the function $\alpha(x)$

$$\alpha(x) = \int_a^x \phi(t)\, dt.$$

We see that $\alpha(x)$ is continuous because ϕ is continuous. This is due to the Fundamental Theorem of Calculus 1 which we will prove.

Mini proof: To prove that the function $\alpha(x) = \int_a^x \phi(t)\,dt$ is continuous, when $\phi(t)$ is continuous, we will show that for a given $\epsilon > 0$, we can find a $\delta > 0$ such that $|x - y| < \delta \Rightarrow |\alpha(x) - \alpha(y)| < \epsilon$.
Without loss of generality, let $x > y$. Given $|x - y| < \delta$ we look at

$$\begin{aligned} |\alpha(x) - \alpha(y)| &= \left| \int_a^x \phi(t)\,dt - \int_a^y \phi(t)\,dt \right| \\ &= \left| \int_y^x \phi(t)\,dt \right| \\ &\leq \int_y^x |\phi(t)|\,dt \end{aligned}$$

Now, since $\phi(t)$ is continuous, we have that $\phi(t)$ attains an absolute maximum value $\phi(c_M) = M \geq \phi(t)$ for any t in the interval $[x, y]$ and an absolute minimum value $\phi(c_m) = m \leq \phi(t)$ for any t in the interval $[x, y]$ by the Extreme Value Theorem. We proved this earlier at the start of the big proof. From here we see that $|\phi(t)| \leq max(|M|, |m|) = \tilde{M}$. This means that

$$\int_y^x |\phi(t)|\,dt \leq \int_x^y \tilde{M}\,dt = \tilde{M}\,|x - y| < \tilde{M}\delta$$

Therefore, letting $\delta = \epsilon/\tilde{M}$ we conclude

$$|\alpha(x) - \alpha(y)| < \tilde{M}\delta = \tilde{M} \cdot \frac{\epsilon}{\tilde{M}} = \epsilon$$

proving that $\alpha(x)$ is continuous.

We also note that $\alpha'(x) = \phi(x)$ which is due to the Fundamental Theorem of Calculus 1.

Mini proof: If x and $x + h$ are in (a, b), then

$$\alpha(x+h) - \alpha(x) = \int_a^{x+h} \phi(t)\,dt - \int_a^x \phi(x)\,dt = \int_x^{x+h} \phi(t)\,dt$$

Therefore, for $h \neq 0$ we have

$$\frac{\alpha(x+h) - \alpha(x)}{h} = \frac{1}{h} \int_x^{x+h} \phi(t)\,dt$$

First, assume $h > 0$ (we will prove this for $h < 0$ next). Since ϕ is continuous on $[x, x+h]$, the Extreme Value Theorem, which we proved earlier, says that there are numbers $u, v \in [x, x+h]$ such that $\phi(u) = m$ and $\phi(v) = M$, where m and M are the absolute minimum and maximum values of ϕ on $[x, x+h]$ therefore we have

$$\begin{aligned} mh &\leq \int_x^{x+h} \phi(t)\,dt \leq Mh \\ \phi(u)h &\leq \int_x^{x+h} \phi(t)\,dt \leq \phi(v)h \end{aligned}$$

Since $h > 0$ we can divide by h and replace the middle of this inequality with what we have from earlier

$$\phi(u) \leq \frac{1}{h}\int_x^{x+h} \phi(t)\,dt \leq \phi(v)$$
$$\phi(u) \leq \frac{\alpha(x+h) - \alpha(x)}{h} \leq \phi(v)$$

Now letting $h \to 0$ we get $u \to x$ and $v \to x$ since u, v are between x and $x+h$. Therefore taking the limit as $h \to 0$:

$$\lim_{h\to 0} \phi(u) \leq \lim_{h\to 0} \frac{\alpha(x+h) - \alpha(x)}{h} \leq \lim_{h\to 0} \phi(v)$$
$$\phi(x) \leq \lim_{h\to 0} \frac{\alpha(x+h) - \alpha(x)}{h} \leq \phi(x)$$

Therefore, by the Squeeze Theorem we have

$$\alpha'(x) = \lim_{h\to 0} \frac{\alpha(x+h) - \alpha(x)}{h} = \phi(x).$$

For the case that $h < 0$ we have again that since ϕ is continuous on $[x+h, x]$ by the Extreme Value Theorem, which we proved earlier, there exist $u, v \in [x+h, x]$ such that $\phi(u) = m$ and $\phi(v) = M$, where m and M are the absolute minimum and maximum values of ϕ on $[x+h, x]$ therefore we have

$$mh \leq \int_x^{x+h} \phi(t)\,dt \leq Mh$$
$$\phi(u)h \leq \int_x^{x+h} \phi(t)\,dt \leq \phi(v)h$$

Since $h < 0$ we can divide by h which flips the inequality signs. We also replace the middle of this inequality with what we have from earlier

$$\phi(u) \geq \frac{1}{h}\int_x^{x+h} \phi(t)\,dt \geq \phi(v)$$
$$\phi(u) \geq \frac{\alpha(x+h) - \alpha(x)}{h} \geq \phi(v)$$

Following the same steps as earlier for taking the limit as $h \to 0$ we have

$$\lim_{h\to 0} \phi(u) \geq \lim_{h\to 0} \frac{\alpha(x+h) - \alpha(x)}{h} \geq \lim_{h\to 0} \phi(v)$$
$$\phi(x) \geq \lim_{h\to 0} \frac{\alpha(x+h) - \alpha(x)}{h} \geq \phi(x)$$

Therefore, by the Squeeze Theorem we have

$$\alpha'(x) = \lim_{h\to 0} \frac{\alpha(x+h) - \alpha(x)}{h} = \phi(x).$$

Which concludes the mini proof.

From here, since $h(x)$ is increasing on $[a, b]$ and $\alpha(x)$ is continuous on $[a, b]$, we enact the Second Mean Value Theorem for Riemann-Stieltjes Integrals:

> Given that $h(x)$ is increasing on $[a, b]$ and $\alpha(x)$ is continuous on $[a, b]$, there exists a $c \in [a, b]$ such that:
> $$\int_a^b h(x)\, d\alpha(x) = h(a) \int_a^c d\alpha(x) + h(b) \int_c^b d\alpha(x)$$

Which we will now prove:

Proof of Second Mean Value Theorem for Riemann-Stieltjes Integrals. To understand this proof we need to breakdown what it means for two functions to be Riemann-Stieltjes integrable to each other, which we will now prove and explain.
Note: since $h(x)$ is increasing on $[a, b]$ and bounded, then $h(x)$ is of bounded variation on $[a, b]$. This we proved, that an increasing and bounded function on an interval $[a, b]$ is of bounded variation on $[a, b]$, in the proof of Theorem 3.3 from Chapter 3 of Stein & Shakarchi Vol. 3.
We start this proof by showing that α is Riemann-Stieltjes integrable with respect to h. We now prove this:

> Given that α is continuous on $[a, b]$ and h is increasing and bounded on $[a, b]$ (meaning h is of bounded variation) then α is Riemann-Stieltjes integrable with respect to h.

Proof. For α to be Riemann-Stieltjes integrable with respect to h we need to show that the upper and lower Riemann-Stieltjes sums are equal. We start by defining a normal Riemann-Stieltjes sum and these two sums.
Take a partition $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$ of the interval $[a, b]$. Since $h(x)$ is monotonically increasing we have that $h(x_i) - h(x_{i-1}) \geq 0$. For the normal Riemann-Stieltjes sum we take a point $\xi_i \in [x_{i-1}, x_i]$ for each $i = 1, ..., n$ and define the sum as follows

$$\mathcal{S}(P, \alpha, h) = \sum_{i=1}^{n} \alpha(\xi_i) \left[h(x_i) - h(x_{i-1})\right]$$

For the upper and lower sums we see that since $\alpha(x)$ is continuous on $[a, b]$ we have, by the Extreme Value Theorem, which we proved earlier, that for each interval $[x_{i-1}, x_i]$ there exists a $u_i, v_i \in [x_{i-1}, x_i]$ such that $\alpha(u_i) = m_i$ and $\alpha(v_i) = M_i$ are the absolute minimum and maximum respectively for $\alpha(x)$ on their specific interval $[x_{i-1}, x_i]$. Note that $\alpha(u_i) \leq \alpha(\xi_i) \leq \alpha(v_i)$ for all $i = 1, ..., n$. From here we define the upper Riemann-Stieltjes sum of α with respect to h and the partition P as

$$\mathcal{U}(P, \alpha, h) = \sum_{i=1}^{n} M_i \left[h(x_i) - h(x_{i-1})\right]$$

and the lower Riemann-Stieltjes sum of α with respect to h and the partition P as

$$\mathcal{L}(P, \alpha, h) = \sum_{i=1}^{n} m_i \left[h(x_i) - h(x_{i-1})\right].$$

From here we note that the norm or mesh of a partition $||P||$ is the length of the largest sub interval and denote

$$\overline{\int_a^b} \alpha(x)\, dh(x) = \inf\{\mathcal{U}(P, \alpha, h) : P \text{ is a partition of } [a, b]\},$$

and

$$\underline{\int_a^b} \alpha(x)\, dh(x) = \sup\{\mathcal{L}(P, \alpha, h) : P \text{ is a partition of } [a, b]\}.$$

From here we note that taking a refinement of P which means adding in more points creating more sub intervals of $[a, b]$ only increases $\mathcal{L}(P, \alpha, h)$. Aka for a refinement P^* ($P \subseteq P^*$) we have $\mathcal{L}(P^*, \alpha, h) \geq \mathcal{L}(P, \alpha, h)$. Similarly taking such a refinement only decreases $\mathcal{U}(P, \alpha, h)$. Aka for a refinement P^* ($P \subseteq P^*$) we have $\mathcal{U}(P^*, \alpha, h) \leq \mathcal{U}(P, \alpha, h)$. We shall now prove this

Mini proof: To see that $\mathcal{L}(P^*, \alpha, h) \geq \mathcal{L}(P, \alpha, h)$ we will prove $\mathcal{L}(P^*, \alpha, h) - \mathcal{L}(P, \alpha, h) \geq 0$. To see this we, for simplicity, assume that the refinement P^* adds in one point $x_{j'}$ between x_{j-1} and x_j of the partition P for some $j = i$ for $i = 1, ..., n$ and is equal to partition P for all other sub intervals. From here we take $m =$ absolute minimum/infimum of $\alpha(x)$ on $[x_{j-1}, x_j]$ and take $m' =$ absolute minimum/infimum of $\alpha(x)$ on $[x_{j-1}, x_{j'}]$ and take $m'' =$ absolute minimum/infimum of $\alpha(x)$ on $[x_{j'}, x_j]$. Since m is the absolute minimum/infimum on the largest interval, which contains the other two intervals, we have $m \leq m'$ and $m \leq m''$. We then see

$$\begin{aligned}
\mathcal{L}(P^*, \alpha, h) - \mathcal{L}(P, \alpha, h) &= \sum_{P^*} m_i[h(x_i) - h(x_{i-1})] - \sum_{P} m_i[h(x_i) - h(x_{i-1})] \\
&= \left[m'(h(x_{j'}) - h(x_{j-1})) + m''(h(x_j) - h(x_{j'}))\right] - m(h(x_j) - h(x_{j-1}))
\end{aligned}$$

From here we remember that $h(x)$ is monotonically increasing so $h(x_{j-1}) \leq h(x_{j'}) \leq h(x_j)$ so m, m', and m'' are all multiplied by non-negative factors. Therefore we see

$$\begin{aligned}
\mathcal{L}(P^*, \alpha, h) - \mathcal{L}(P, \alpha, h) &= \left[m'(h(x_{j'}) - h(x_{j-1})) + m''(h(x_j) - h(x_{j'}))\right] \\
&\quad - m(h(x_j) - h(x_{j'}) + h(x_{j'}) - h(x_{j-1})) \\
&= (m' - m)(h(x_{j'}) - h(x_{j-1})) + (m'' - m)(h(x_j) - h(x_{j'})),
\end{aligned}$$

since $m \leq m'$ and $m \leq m''$ and $h(x)$ being monotonically increasing we see that both terms are non-negative meaning

$$\mathcal{L}(P^*, \alpha, h) - \mathcal{L}(P, \alpha, h) = (m' - m)(h(x_{j'}) - h(x_{j-1})) + (m'' - m)(h(x_j) - h(x_{j'})) \geq 0.$$

Doing this procedure for every new point in a refined partition proves that refining the partition only increases the lower Riemann-Stieltjes sum, $\mathcal{L}(P, \alpha, h)$.

Now to prove that refining the partition only decreases $\mathcal{U}(P, \alpha, h)$ we prove $\mathcal{U}(P, \alpha, h) - \mathcal{U}(P^*, \alpha, h) \geq 0$. Following the same procedure as above we assume that the refinement P^* adds in one point $x_{j'}$ between x_{j-1} and x_j of the partition P for some $j = i$ for $i = 1, ..., n$ and is equal to partition P for all other sub intervals. From here we take $M =$ absolute maximum/supremum of $\alpha(x)$ on $[x_{j-1}, x_j]$ and take M' = absolute maximum/supremum of $\alpha(x)$ on $[x_{j-1}, x_{j'}]$ and take M'' = absolute maximum/supremum of $\alpha(x)$ on $[x_{j'}, x_j]$. Since M is the absolute maximum/supremum on the largest interval, which contains the other two intervals, we have $M \geq M'$ and $M \geq M''$. We then see

$$\begin{aligned}\mathcal{U}(P,\alpha,h) - \mathcal{U}(P^*,\alpha,h) &= \sum_P M_i[h(x_i) - h(x_{i-1})] - \sum_{P^*} M_i[h(x_i) - h(x_{i-1})] \\ &= M(h(x_j) - h(x_{j-1})) - \left[M'(h(x_{j'}) - h(x_{j-1})) + M''(h(x_j) - h(x_{j'}))\right]\end{aligned}$$

Again, we remember $h(x)$ is monotonically increasing so $h(x_{j-1}) \leq h(x_{j'}) \leq h(x_j)$ and $M \geq M'$ and $M \geq M''$ therefore

$$\begin{aligned}\mathcal{U}(P,\alpha,h) - \mathcal{U}(P^*,\alpha,h) &= M(h(x_j) - h(x_{j'}) + h(x_{j'}) - h(x_{j-1})) \\ &\quad - \left[M'(h(x_{j'}) - h(x_{j-1})) + M''(h(x_j) - h(x_{j'}))\right] \\ &= (M - M')(h(x_{j'}) - h(x_{j-1})) + (M - M'')(h(x_j) - h(x_{j'})) \\ &\Rightarrow \text{ Since both terms are non-negative:} \\ \mathcal{U}(P,\alpha,h) - \mathcal{U}(P^*,\alpha,h) &= (M - M')(h(x_{j'}) - h(x_{j-1})) + (M - M'')(h(x_j) - h(x_{j'})) \geq 0.\end{aligned}$$

Doing this procedure for every new point in a refined partition proves that refining the partition only decreases the upper Riemann-Stieltjes sum, $\mathcal{U}(P,\alpha,h)$. Concluding the mini proof.

Since the infimum is the greatest lower bound and the supremum is the least upper bound we see that taking $||P|| \to 0$ is equivalent to the taking the infimum and supremum of the two integrals denoted earlier. This means

$$\overline{\int_a^b} \alpha(x)\, dh(x) = \mathcal{U}(P,\alpha,h) \text{ as } ||P|| \to 0,$$

and

$$\underline{\int_a^b} \alpha(x)\, dh(x) = \mathcal{L}(P,\alpha,h) \text{ as } ||P|| \to 0,$$

as well. Note that since $\mathcal{L}(P,\alpha,h) \leq \mathcal{S}(P,\alpha,h) \leq \mathcal{U}(P,\alpha,h)$ for any partition P, we see that taking a refinement of a partition only decreases the error of $\mathcal{S}(P,\alpha,h)$ to $\int_a^b \alpha(x)\, dh(x)$. The Riemann-Stieltjes integral for α with respect to h is properly defined if

$$\overline{\int_a^b} \alpha(x)\, dh(x) = \underline{\int_a^b} \alpha(x)\, dh(x).$$

From here we want to show that as $||P|| \to 0$ we obtain this equality. To do this we will show that for any $\epsilon > 0$, there exists a $\delta > 0$ such that for $||P|| < \delta$ we have

$$|\mathcal{U}(P,\alpha,h) - \mathcal{L}(P,\alpha,h)| < \epsilon.$$

To prove this we first prove that since $\alpha(x)$ is continuous on $[a, b]$, which is a closed interval of $\mathbb{R}$, $\alpha(x)$ is uniformly continuous on $[a, b]$. Recall that the definition of uniform continuity on $[a, b]$ is: α is uniformly continuous on $[a, b] \iff \forall \epsilon > 0$, $\exists \delta > 0$ such that $\forall x, y \in [a, b]$ we have $|x - y| < \delta \Rightarrow\ |\alpha(x) - \alpha(y)| < \epsilon$.

Mini proof: To prove that $\alpha(x)$ being continuous on $[a, b] \Rightarrow\ \alpha(x)$ is uniformly continuous on $[a, b]$ we proceed by contradiction:
This means we suppose that $\alpha(x)$ is continuous on $[a, b]$ but $\exists \epsilon > 0$ such that $\forall \delta > 0$ we have that $\exists x, y \in [a, b]$ such that $|x - y| < \delta$ and $|\alpha(x) - \alpha(y)| \geq \epsilon$. This means that for $\delta = 1/n$ for $n \in \mathbb{N}$, there exists x_n and y_n such that

$$|x_n - y_n| < \frac{1}{n} \text{ but } |\alpha(x_n) - \alpha(y_n)| \geq \epsilon > 0$$

From here we can extract sequences $\{x_n\}_{n\in\mathbb{N}} \in [a, b]$ and $\{y_n\}_{n\in\mathbb{N}} \in [a, b]$. From here we can use the Bolzano-Weierstrass Theorem:
Every bounded sequence of real numbers has a convergent subsequence. (Which we will now prove)

Mini mini proof: To prove the Bolzano-Weierstrass Theorem we start by taking a sequence $\{s_n\}_{n\in\mathbb{N}}$ such that $|s_n| < L$ for all s_n in this sequence. Setting $a_0 = -L$ and $b_0 = L$ we see that $\{s_n\}_{n\in\mathbb{N}} \in [a_0, b_0]$ and $|b_0 - a_0| = 2L$. We then split this interval into two halves. Since our sequence is infinite one of these two halves must contain an infinite number of elements from $\{s_n\}$. Take this half and name it $[a_1, b_1]$ where $a_1 \in \{a_0, a_0 + \frac{b_0 - a_0}{2}\}$ meaning a_1 is either a_0 or $a_0 + \frac{b_0 - a_0}{2} = a_0 + L$ meaning $a_1 \geq a_0$. Note that this also means that $|b_1 - a_1| = \frac{1}{2}|b_0 - a_0| = L$. Since $[a_1, b_1]$ contains infinitely elements in $\{s_n\}$ pick one and call it s_{i_1}. Again, we divide our new interval $[a_1, b_1]$ into two halves, and since $\{s_n\}$ has infinitely elements one of these two halves must contain an infinite amount of elements of $\{s_n\}$. From here we take this new half and name it $[a_2, b_2]$ where $a_2 \in \{a_1, a_1 + \frac{b_1 - a_1}{2}\}$ so $a_2 \geq a_1$ and $|b_2 - a_2| = \frac{1}{2}|b_1 - a_1| = \frac{1}{2}L$. Alas, there are infinitely many elements of $\{s_n\}$ in $[a_2, b_2]$ therefore we can select one with $n > i_1$ and label it s_{i_2}.
Repeating this process we create a set of intervals that contain each other with shortening length:

$$[a_0, b_0] \supseteq [a_1,b_1] \supseteq [a_2, b_2] \supseteq \dots$$
$$\text{length } 2L \mid \text{length } L \mid \text{length } \frac{1}{2}\text{L} \mid \dots$$

It is important to realize that $|b_n - a_n| = \frac{1}{2}|b_{n-1} - a_{n-1}|$ for all $n \in \mathbb{N}$ so then $|b_n - a_n| = \frac{1}{2^n}|b_0 - a_0| = \frac{L}{2^{n-1}}$ for all $n \in \mathbb{N}$. Note that we also have $a_0 \leq a_1 \leq a_2 \leq \dots$.
The sequence $a_1, a_2, a_3, \dots$ is monotone increasing and bounded above by $b_0 = L$. This means that the sequence converges to a limit and we will call this limit s. This is due to the Monotone Convergence Theorem which states:
If $\{a_n\}$ is a monotonically increasing sequence that is bounded above, then $\{a_n\}$ converges to its supremum/least upper bound. Also if $\{b_n\}$ is a monotonically decreasing sequence that is bounded below, then $\{b_n\}$ converges to its infimum/greatest lower bound.
We proved this theorem at the beginning of the proof that we are technically in (version of the second mean value theorem for integrals). Note that we are only using the first half of

the theorem.
From this we see truly that the sequence, $a_1, a_2, a_3, \ldots$ converges to a limit which we will call s. Our goal to prove the Bolzano-Weierstrass Theorem is to now show that $\lim_{n\to\infty} s_{i_n} = s$. Let $\epsilon > 0$. Since $\{a_n\}_{n\in\mathbb{N}}$ converges to s there exists an $N_1 \in \mathbb{N}$ such that $|a_n - s| < \frac{\epsilon}{2}$ whenever $n \geq N_1$. Since s_{i_n} lies in the interval $[a_n, b_n]$ and the length of the interval $[a_n, b_n]$ is $\frac{L}{2^{n-1}}$ we see that the distance between s_{i_n} and a_n is at most $\frac{L}{2^{n-1}}$ which tends to 0 as $n \to \infty$. This means that there exists an $N_2 \in \mathbb{N}$ such that $|s_{i_n} - a_n| < \frac{\epsilon}{2}$ whenever $n \geq N_2$. From here we choose $N' = max(N_1, N_2)$ and we see

$$n \geq N' \Rightarrow \ |s_{i_n} - s| = |s_{i_n} - a_n + a_n - s| \leq |s_{i_n} - a_n| + |a_n - s| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Concluding the mini mini proof of the Bolzano-Weierstrass Theorem.

Now that we have proven the Bolzano-Weierstrass Theorem we see that the two sequence $\{x_n\}$ and $\{y_n\}$ that are both contained in $[a, b]$ contain subsequences, $\{x_{n_k}\}$ and $\{y_{n_k}\}$, that converge as $n_k \to \infty$. Let $\{x_{n_k}\} \to x_0$ and $\{y_{n_k}\} \to y_0$ as $n_k \to \infty$. From here we have

$$\begin{aligned}\lim_{n_k\to\infty} |x_0 - y_0| &= \lim_{n_k\to\infty} |x_0 - x_{n_k} + x_{n_k} - y_{n_k} + y_{n_k} - y_0| \\ &\leq \lim_{n_k\to\infty} |x_0 - x_{n_k}| + |x_{n_k} - y_{n_k}| + |y_{n_k} - y_0|\end{aligned}$$

Since $\{x_{n_k}\} \to x_0$ and $\{y_{n_k}\} \to y_0$ as $n_k \to \infty$ and $|x_n - y_n| < \frac{1}{n}$ we have

$$\begin{aligned}\lim_{n_k\to\infty} |x_0 - y_0| &\leq \lim_{n_k\to\infty} 0 + |x_{n_k} - y_{n_k}| + 0 \\ &< \lim_{n_k\to\infty} \frac{1}{n_k} = 0\end{aligned}$$

which means that $x_0 = y_0$.
But since $\alpha(x)$ is continuous this means that $x_{n_k} \to x_0 \Rightarrow \ \alpha(x_{n_k}) \to \alpha(x_0)$ and $y_{n_k} \to y_0 \Rightarrow \alpha(y_{n_k}) \to \alpha(y_0)$. But we have from the beginning of our assumption that $|\alpha(x_{n_k}) - \alpha(y_{n_k})| \geq \epsilon > 0$ which we see results in a contradiction:

$$\begin{aligned}0 < \epsilon \leq |\alpha(x_{n_k}) - \alpha(y_{n_k})| &= |\alpha(x_{n_k}) - \alpha(x_0) + \alpha(x_0) - \alpha(y_0) + \alpha(y_0) - \alpha(y_{n_k})| \\ &\leq |\alpha(x_{n_k}) - \alpha(x_0)| + |\alpha(x_0) - \alpha(y_0)| + |\alpha(y_0) - \alpha(y_{n_k})|\end{aligned}$$

Now taking the limit as $n_k \to \infty$, we see $\alpha(x_{n_k}) \to \alpha(x_0)$ and $\alpha(y_{n_k}) \to \alpha(y_0)$ and note that $x_0 = y_0$

$$\begin{aligned}\lim_{n_k\to\infty} 0 < \lim_{n_k\to\infty} \epsilon &\leq \lim_{n_k\to\infty} |\alpha(x_{n_k}) - \alpha(x_0)| + |\alpha(x_0) - \alpha(y_0)| + |\alpha(y_0) - \alpha(y_{n_k})| \\ 0 < \epsilon &\leq 0 + 0 + 0 = 0\end{aligned}$$

Giving us a contradiction which means that $\alpha(x)$ must be uniformly continuous on $[a, b]$. This concludes the mini proof.

Now that we have proven that $\alpha(x)$ being continuous on $[a, b]$ implies that $\alpha(x)$ is uniformly

continuous on $[a,b]$ we turn out attention back to proving:
For any $\epsilon > 0$, there exists a $\delta > 0$ such that for $||P|| < \delta$ we have

$$|\mathcal{U}(P,\alpha,h) - \mathcal{L}(P,\alpha,h)| < \epsilon.$$

Since $\alpha(x)$ is uniformly convergent on $[a,b]$ we have that for every $\epsilon > 0$, $\exists\delta > 0$ such that for any $x, y \in [a,b]$ with $|x-y| < \delta$ we have $|\alpha(x) - \alpha(y)| < \frac{\epsilon}{M'}$ where M' = total variation of h over $[a,b]$. Note that since $h(x)$ is a monotonically increasing and bounded function, the total variation of $h(x)$ on $[a,b]$ is

$$T_h(a,b) = \sum_{i=1}^{n} |h(x_i) - h(x_{i-1})| = \sum_{i=1}^{n} h(x_i) - h(x_{i-1}) = h(x_n) - h(x_0) = M' < \infty$$

Now we take a look at $|\mathcal{U}(P,\alpha,h) - \mathcal{L}(P,\alpha,h)|$

$$\begin{aligned} |\mathcal{U}(P,\alpha,h) - \mathcal{L}(P,\alpha,h)| &= \left| \sum_{i=1}^{n} M_i \left[h(x_i) - h(x_{i-1})\right] - \sum_{i=1}^{n} m_i \left[h(x_i) - h(x_{i-1})\right] \right| \\ &= \left| \sum_{i=1}^{n} (M_i - m_i) \left[h(x_i) - h(x_{i-1})\right] \right| \end{aligned}$$

From here we remember that $||P||$ is the largest sub interval of the partition P. If we have $||P|| < \delta$ we see that $[x_{i-1}, x_i] < \delta$ for all $i = 1, ..., n$. From here we remember that each pair of u_i and v_i are contained in one of the $[x_{i-1}, x_i]$ sub intervals. This means that $|v_i - u_i| < \delta$ for all $i = 1, ..., n$. Note that $\alpha(v_i) = M_i$ and $\alpha(u_i) = m_i$ for all $i = 1, ..., n$. Since $\alpha(x)$ is uniformly continuous we have

$$\begin{aligned} |v_i - u_i| < \delta \Rightarrow \ & |f(v_i) - f(u_i)| < \frac{\epsilon}{M'} \text{ for all } i = 1, ..., n. \\ \Rightarrow \ & |M_i - m_i| < \frac{\epsilon}{M'} \text{ for all } i = 1, ..., n. \end{aligned}$$

Combining this with the earlier equality

$$\begin{aligned} |\mathcal{U}(P,\alpha,h) - \mathcal{L}(P,\alpha,h)| &= \left| \sum_{i=1}^{n} (M_i - m_i) \left[h(x_i) - h(x_{i-1})\right] \right| \\ &< \left| \sum_{i=1}^{n} \frac{\epsilon}{M'} \cdot \left[h(x_i) - h(x_{i-1})\right] \right| \\ &\le \frac{\epsilon}{M'} \sum_{i=1}^{n} |[h(x_i) - h(x_{i-1})]| \\ &= \frac{\epsilon}{M'} \cdot M' \\ &= \epsilon \end{aligned}$$

Hence proving that $|\mathcal{U}(P,\alpha,h) - \mathcal{L}(P,\alpha,h)| < \epsilon$. This means that if we take $\delta \to 0$ we have $||P|| \to 0$ and we get

$$\overline{\int_a^b} \alpha(x)\, dh(x) = \underline{\int_a^b} \alpha(x)\, dh(x)$$

for α continuous on $[a, b]$ and h increasing and bounded on $[a, b]$. Thus proving that α is Riemann-Stieltjes integrable with respect to h. $\square$

We have now proven that α is Riemann-Stieltjes integrable with respect to h. From this we will prove that h is Riemann-Stieltjes integrable with respect to α.

Mini proof: We will prove that h is Riemann-Stieltjes integrable with respect to α given that α is Riemann-Stieltjes integrable with respect to h.
To prove that h is Riemann-Stieltjes integrable with respect to α we must show

$$|\mathcal{U}(P, h, \alpha) - \mathcal{L}(P, h, \alpha)| < \epsilon \text{ as } ||P|| \to 0.$$

To do this we investigate the LHS.

$$|\mathcal{U}(P, h, \alpha) - \mathcal{L}(P, h, \alpha)| = \left| \sum_{i=1}^{n} M_i \left[\alpha(x_i) - \alpha(x_{i-1})\right] - \sum_{i=1}^{n} m_i \left[\alpha(x_i) - \alpha(x_{i-1})\right] \right|$$

where M_i is the absolute maximum of $h(x)$ on $[x_{i-1}, x_i]$ and m_i is the absolute minimum of $h(x)$ on $[x_{i-1}, x_i]$. Since $h(x)$ is a monotonically increasing and bounded function, meaning $h(x) \geq h(y)$ for all $x > y \in [a, b]$, we see that the absolute minimum and maximum of each sub interval $[x_{i-1}, x_i]$ is $h(x_{i-1}) = m_i$ and $h(x_i) = M_i$ respectively. This means that we have

$$\begin{aligned}
|\mathcal{U}(P, h, \alpha) - \mathcal{L}(P, h, \alpha)| &= \left| \sum_{i=1}^{n} h(x_i) \left[\alpha(x_i) - \alpha(x_{i-1})\right] - \sum_{i=1}^{n} h(x_{i-1}) \left[\alpha(x_i) - \alpha(x_{i-1})\right] \right| \\
&= \left| \sum_{i=1}^{n} \left(h(x_i) - h(x_{i-1})\right) \left[\alpha(x_i) - \alpha(x_{i-1})\right] \right|
\end{aligned}$$

Now, since $\alpha(x)$ is uniformly continuous on $[a, b]$ we see that if we take $||P|| < \delta$ this means that $|x_i - x_{i-1}| < \delta$ for all $i = 1, ..., n$. This means, because of the uniform continuity, that $|\alpha(x_i) - \alpha(x_{i-1})| < \frac{\epsilon}{M'}$ for all $i = 1, ..., n$ where again $M' = T_h(a, b)$ which is finite because $h(x)$ is a monotonically increasing and bounded function and therefore a function of bounded variation. This means that we have

$$\begin{aligned}
|\mathcal{U}(P, h, \alpha) - \mathcal{L}(P, h, \alpha)| &= \left| \sum_{i=1}^{n} \left(h(x_i) - h(x_{i-1})\right) \left[\alpha(x_i) - \alpha(x_{i-1})\right] \right| \\
&\leq \left| \sum_{i=1}^{n} \left(h(x_i) - h(x_{i-1})\right) \cdot \frac{\epsilon}{M'} \right| \\
&\leq \frac{\epsilon}{M'} \sum_{i=1}^{n} |h(x_i) - h(x_{i-1})| \\
&= \frac{\epsilon}{M'} \cdot M' \\
&= \epsilon
\end{aligned}$$

Therefore, since we have proven $|\mathcal{U}(P,h,\alpha)-\mathcal{L}(P,h,\alpha)|<\epsilon$, we see that as $\delta \to 0$ we have $||P|| \to 0$ which makes

$$\overline{\int_a^b} h(x)\,d\alpha(x) = \underline{\int_a^b} h(x)\,d\alpha(x)$$

proving that h is Riemann-Stieltjes integrable with respect to α.

Now that we have proven that α is Riemann-Stieltjes integrable with respect to h and h is Riemann-Stieltjes integrable with respect to α we wish to use the formula for integration by parts of Riemann-Stieltjes integrals which we will prove:

Let h be a Riemann-Stieltjes integrable function with respect to α on the interval $[a,b]$. Then α is a Riemann-Stieltjes integrable function with respect to h on the interval $[a,b]$ and

$$\int_a^b h(x)\,d\alpha(x) + \int_a^b \alpha(x)\,dh(x) = h(b)\alpha(b) - h(a)\alpha(a).$$

Proof of Riemann-Stieltjes Integration by Parts. To show that $\int_a^b h(x)\,d\alpha(x)+\int_a^b \alpha(x)\,dh(x) = h(b)\alpha(b)-h(a)\alpha(a)$ we will show that for any $\epsilon>0$ we have

$$\left|\int_a^b h(x)\,d\alpha(x) + \int_a^b \alpha(x)\,dh(x) - [h(b)\alpha(b) - h(a)\alpha(a)]\right| < \epsilon$$

To do this, we take an arbitrary $\epsilon > 0$. Now, since $h(x)$ is Riemann-Stieltjes integrable with respect to $\alpha(x)$ on the interval $[a,b]$ we see that for some $A \in \mathbb{R}$ we have $\int_a^b h(x)\,d\alpha(x) = A$. We also have for $\epsilon_1 = \frac{\epsilon}{2} > 0$ there exists a partition P_{ϵ_1} of $[a,b]$ such that for any finer partition of P_{ϵ_1} ($P_{\epsilon_1} \subseteq P$) and for any $\xi_{i_1} \in [x_{i-1},x_i]$ for each sub interval, we have

$$|\mathcal{S}(P,h,\alpha) - A| < \epsilon_1 = \frac{\epsilon}{2}$$

This is due to the fact that taking a refinement only decreases the error of $\mathcal{S}(P,h,\alpha)$ to $\int_a^b h(x)\,d\alpha(x)$.
Similarly, since $\alpha(x)$ is Riemann-Stieltjes integrable with respect to $h(x)$ on the interval $[a,b]$ we see that for some $B \in \mathbb{R}$ we have $\int_a^b \alpha(x)\,dh(x) = B$. Again we can say that we have for $\epsilon_2 = \frac{\epsilon}{2} > 0$ there exists a partition P_{ϵ_2} of $[a,b]$ such that for any finer partition of P_{ϵ_2} ($P_{\epsilon_2} \subseteq P$) and for any $\xi_{i_2} \in [x_{i-1},x_i]$ for each sub interval, we have

$$|\mathcal{S}(P,\alpha,h) - B| < \epsilon_2 = \frac{\epsilon}{2}$$

Next, define $P_\epsilon = P_{\epsilon_1} \cup P_{\epsilon_2}$ (which is finer than the two partitions on their own). We now have for all partitions that are finer than P_ϵ ($P_\epsilon \subseteq P$) and any $\xi_i \in [x_{i-1},x_i]$ for each sub interval, which satisfies the two inequalities above, gives

$$|\mathcal{S}(P,h,\alpha) + \mathcal{S}(P,\alpha,h) - (A+B)| \le |\mathcal{S}(P,h,\alpha) - A| + |\mathcal{S}(P,\alpha,h) - B| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Now, we let $\xi_{i_1} = x_i$ and $\xi_{i_2} = x_{i-1}$ for each sub interval for $|\mathcal{S}(P, h, \alpha)$ and $\mathcal{S}(P, \alpha, h)$ respectively.

$$|\mathcal{S}(P, h, \alpha) + \mathcal{S}(P, \alpha, h) - (A + B)| < \epsilon$$
$$\left| \sum_{i=1}^{n} h(\xi_{i_1})[\alpha(x_i) - \alpha(x_{i-1})] + \sum_{i=1}^{n} \alpha(\xi_{i_2})[h(x_i) - h(x_{i-1})] - (A + B) \right| < \epsilon$$
$$\left| \sum_{i=1}^{n} h(x_i)\alpha(x_i) - h(x_i)\alpha(x_{i-1}) + \alpha(x_{i-1})h(x_i) - \alpha(x_{i-1})h(x_{i-1}) - (A + B) \right| < \epsilon$$
$$\left| \sum_{i=1}^{n} h(x_i)\alpha(x_i) - \alpha(x_{i-1})h(x_{i-1}) - (A + B) \right| < \epsilon$$

This first term in the absolute values is a telescoping sum where $\sum_{i=1}^{n} h(x_i)\alpha(x_i) - \alpha(x_{i-1})h(x_{i-1}) = h(b)\alpha(b) - h(a)\alpha(a)$. Hence we have

$$|h(b)\alpha(b) - h(a)\alpha(a) - (A + B)| < \epsilon$$

Multiplying by -1 inside the absolute values (changing nothing of the output of the absolute values) gives us our desired result

$$\left| \int_a^b h(x)\, d\alpha(x) + \int_a^b \alpha(x)\, dh(x) - [h(b)\alpha(b) - h(a)\alpha(a)] \right| < \epsilon. \tag{17}$$

By taking $\epsilon \to 0$ we have

$$\int_a^b h(x)\, d\alpha(x) + \int_a^b \alpha(x)\, dh(x) = h(b)\alpha(b) - h(a)\alpha(a),$$

proving Riemann-Stieltjes Integration by Parts. □

Now that we have proven Riemann-Stieltjes Integration by Parts we can progress on proving the Second Mean Value Theorem for Riemann-Stieltjes integrals.
From Riemann-Stieltjes Integration by Parts, since h is a Riemann-Stieltjes integrable function with respect to α on the interval $[a, b]$ and α is a Riemann-Stieltjes integrable function with respect to h on the interval $[a, b]$, we have

$$\int_a^b h(x)\, d\alpha(x) + \int_a^b \alpha(x)\, dh(x) = h(b)\alpha(b) - h(a)\alpha(a)$$
$$\int_a^b h(x)\, d\alpha(x) = h(b)\alpha(b) - h(a)\alpha(a) - \int_a^b \alpha(x)\, dh(x)$$

Next we see that since $\alpha(x)$ is continuous on the closed and bounded interval $[a, b]$ we have that $\alpha(x)$ is bounded on $[a, b]$. This is due to the Boundedness Theorem which we proved at the beginning of the big proof. We will use this to prove the First Mean Value Theorem for Riemann-Stieltjes integrals and a corollary to the First Mean Value Theorem for Riemann-Stieltjes integrals.

First Mean Value Theorem for Riemann-Stieltjes integrals:
Let $\alpha(x)$ be bounded on $[a, b]$ and let $h(x)$ be an increasing function on $[a, b]$. Furthermore, let $\alpha(x)$ be Riemann-Stieltjes integrable with respect to $h(x)$ on $[a, b]$. If $M = \sup\{\alpha(x) : x \in [a, b]\}$ and $m = \inf\{\alpha(x) : x \in [a, b]\}$ then there exists a $c \in [m, M]$ such that

$$\int_a^b \alpha(x)\, dh(x) = c[h(b) - h(a)].$$

Proof of the First Mean Value Theorem for Riemann-Stieltjes Integrals. Suppose that $h(a) = h(b)$. Then since $h(x)$ is increasing on $[a, b]$ we must have that $h(x)$ is constant on $[a, b]$ and so $\int_a^b \alpha(x)\, dh(x) = 0$. Furthermore, $h(b) - h(a) = 0$ so any $c \in [m, M]$ satisfies the equality. Now suppose that $h(a) < h(b)$ and let P be a partition of $[a, b]$ where $P = \{a = x_0, x_1, \ldots, x_n = b\}$ and consider the following Riemann-Stieltjes sum for some $\xi_i \in [x_{i-1}, x_i]$ for all $i = 1, ..., n$:

$$\mathcal{S}(P, \alpha, h) = \sum_{i=1}^{n} \alpha(\xi_i)[h(x_i) - h(x_{i-1})]$$

Since $m = \inf\{\alpha(x) : x \in [a, b]\}$ and $M = \sup\{\alpha(x) : x \in [a, b]\}$ we see that $m \leq \alpha(\xi_i) \leq M$ for all $\xi_i \in [x_{i-1}, x_i]$ for all sub intervals. Therefore we see

$$\begin{aligned}\sum_{i=1}^{n} m[h(x_i) - h(x_{i-1})] &\leq \sum_{i=1}^{n} \alpha(\xi_i)[h(x_i) - h(x_{i-1})] \leq \sum_{i=1}^{n} M[h(x_i) - h(x_{i-1})] \\ m\sum_{i=1}^{n}[h(x_i) - h(x_{i-1})] &\leq \sum_{i=1}^{n} \alpha(\xi_i)[h(x_i) - h(x_{i-1})] \leq M\sum_{i=1}^{n}[h(x_i) - h(x_{i-1})]\end{aligned}$$

Since $h(x)$ is an increasing sequence we see that $\sum_{i=1}^{n}[h(x_i) - h(x_{i-1})]$ is a telescoping series and equals $[h(b) - h(a)]$ so

$$m[h(b) - h(a)] \leq \sum_{i=1}^{n} \alpha(\xi_i)[h(x_i) - h(x_{i-1})] \leq M[h(b) - h(a)]$$

This inequality holds for all partitions of $[a, b]$, and since $\alpha(x)$ is Riemann-Stieltjes integrable with respect to $h(x)$ on $[a, b]$ we have

$$\begin{aligned}m[h(b) - h(a)] \leq \int_a^b \alpha(x)\, dh(x) \leq M[h(b) - h(a)] \\ m \leq \frac{1}{h(b) - h(a)} \int_a^b \alpha(x)\, dh(x) \leq M\end{aligned}$$

From here we take $c = \frac{1}{h(b)-h(a)} \int_a^b \alpha(x)\, dh(x)$ as we can see that $c \in [m, M]$.
Also

$$c = \frac{1}{h(b) - h(a)} \int_a^b \alpha(x)\, dh(x) \Rightarrow\ c[h(b) - h(a)] = \int_a^b \alpha(x)\, dh(x).$$

This concludes the proof of the First Mean Value Theorem for Riemann-Stieltjes integrals. □

Now that we have proven the First Mean Value Theorem for Riemann-Stieltjes integrals we will prove a corollary to this:

Corollary to First Mean Value Theorem for Riemann-Stieltjes integrals:
Let $\alpha(x)$ be a continuous and bounded function on $[a, b]$ and let $h(x)$ be an increasing function on $[a, b]$. Furthermore, let $\alpha(x)$ be Riemann-Stieltjes integrable with respect to $h(x)$ on $[a, b]$. If $M = \sup\{\alpha(x) : x \in [a, b]\}$ and $m = \inf\{\alpha(x) : x \in [a, b]\}$ then there exists an $c' \in [a, b]$ such that

$$\int_a^b \alpha(x)\, dh(x) = \alpha(c')[h(b) - h(a)].$$

Proof of Corollary. Since $\alpha(x)$ is continuous and bounded on $[a, b]$ and $h(x)$ be increasing on $[a, b]$ as well as α being Riemann-Stieltjes integrable with respect to h we can enact the First Mean Value Theorem for Riemann-Stieltjes integrals meaning that we have for some $c \in [m, M]$

$$\int_a^b \alpha(x)\, dh(x) = c[h(b) - h(a)].$$

Since $\alpha(x)$ is continuous on the closed and bounded interval $[a, b]$ we have that $M = \sup\{\alpha(x) : x \in [a, b]\} = \alpha(v)$ for some $v \in [a, b]$ and similarly $m = \inf\{\alpha(x) : x \in [a, b]\} = \alpha(u)$ for some $u \in [a, b]$. This means that $c \in [\alpha(u), \alpha(v)]$ aka $\alpha(u) \leq c \leq \alpha(v)$.
We then have by the Intermediate Value Theorem, which we proved earlier in the big proof, that since $\alpha(x)$ is continuous on $[a, b]$ there exists a $c' \in [a, b]$ such that $\alpha(c') = c$. Hence

$$\int_a^b \alpha(x)\, dh(x) = \alpha(c')[h(b) - h(a)].$$

This concludes the proof of the corollary. □

Now we can get back to proving the Second Mean Value Theorem for Riemann-Stieltjes integrals:
We left off with

$$\int_a^b h(x)\, d\alpha(x) = h(b)\alpha(b) - h(a)\alpha(a) - \int_a^b \alpha(x)\, dh(x)$$

From here we note that since $\alpha(x)$ is continuous on $[a, b]$ and therefore bounded on $[a, b]$ by the Boundedness Theorem, which we proved earlier, and since $h(x)$ is increasing on $[a, b]$ and $\alpha(x)$ is Riemann-Stieltjes integrable with respect to $h(x)$ on $[a, b]$ we are able to use the corollary to First Mean Value Theorem for Riemann-Stieltjes integrals which states that there exists a $c' \in [a, b]$ such that

$$\int_a^b \alpha(x)\, dh(x) = \alpha(c')[h(b) - h(a)].$$

For cleanliness we will relabel $c' = c$.
Substituting this into the recalled formula gives us

$$\int_a^b h(x)\, d\alpha(x) = h(b)\alpha(b) - h(a)\alpha(a) - \int_a^b \alpha(x)\, dh(x)$$
$$\int_a^b h(x)\, d\alpha(x) = h(b)\alpha(b) - h(a)\alpha(a) - \alpha(c)[h(b) - h(a)]$$
$$\int_a^b h(x)\, d\alpha(x) = h(b)\alpha(b) - h(a)\alpha(a) - h(b)\alpha(c) + h(a)\alpha(c)$$
$$\int_a^b h(x)\, d\alpha(x) = h(a)[\alpha(c) - \alpha(a)] + h(b)[\alpha(b) - \alpha(c)]$$

From here we remark that, due to the sums being telescopic,

$$\int_a^c d\alpha(x) = \sum_{i=1}^{n}[\alpha(x_i) - \alpha(x_{i-1})] = \alpha(c) - \alpha(a) \text{ for any partition } P \text{ of } [a, b].$$

and

$$\int_c^b d\alpha(x) = \sum_{i=1}^{n}[\alpha(x_i) - \alpha(x_{i-1})] = \alpha(b) - \alpha(c) \text{ for any partition } P \text{ of } [a, b].$$

Substituting these in gives us

$$\int_a^b h(x)\, d\alpha(x) = h(a)[\alpha(c) - \alpha(a)] + h(b)[\alpha(b) - \alpha(c)]$$
$$\int_a^b h(x)\, d\alpha(x) = h(a)\int_a^c d\alpha(x) + h(b)\int_c^b d\alpha(x)$$

as desired.
This, definitively, proves the Second Mean Value Theorem for Riemann-Stieltjes Integrals.
□

Now that we have finished proving the Second Mean Value Theorem for Riemann-Stieltjes integrals we continue on with our goal of proving a version of the Second Mean Value Theorem for integrals.
If we can remember, we had defined $\alpha(x) = \int_a^x \phi(t)\, dt$ which we proved to be continuous on $[a, b]$ and proved that $\alpha'(x) = \phi(x)$. From here we see that since $h(x)$ is increasing on $[a, b]$ and $\alpha(x)$ is continuous on $[a, b]$ we can use the Second Mean Value Theorem for Riemann-Stieltjes integrals which tells us that there exists a $c \in [a, b]$ such that

$$\int_a^b h(x)\, d\alpha(x) = h(a)\int_a^c d\alpha(x) + h(b)\int_c^b d\alpha(x)$$

Now we wish to use the Reducing Riemann-Stieltjes Integrals to Riemann Integrals Lemma. To use this we must prove it!

Reducing Riemann-Stieltjes Integrals to Riemann Integrals Lemma:

Let h be a Riemann-Stieltjes integrable function with respect to α on the interval $[a, b]$, Furthermore, let $h(x)$ be bounded on [a,b], let $\alpha'(x)$ exist and be continuous on $[a, b]$. Then

$$\int_a^b h(x)\, d\alpha(x) = \int_a^b h(x)\alpha'(x)\, dx.$$

Proof of Lemma. From the assumptions we have that h is a Riemann-Stieltjes integrable function with respect to α on the interval $[a, b]$ meaning that for some $A \in \mathbb{R}$ we have $\int_a^b h(x)\, d\alpha(x) = A$.
Our goal is to show that for all $\epsilon > 0$, we have that there exists a partition of $[a, b]$ which we call P_ϵ such that for all P finer than P_ϵ ($P_\epsilon \subseteq P$) we have

$$|\mathcal{S}(P, h\alpha', x) - A| < \epsilon$$

This would prove that $\int_a^b h(x)\alpha'(x)\, dx = A$.
We start by letting $\epsilon > 0$, since h is a Riemann-Stieltjes integrable function with respect to α on the interval $[a, b]$ we then have for $\epsilon_1 = \frac{\epsilon}{2} > 0$, there exists a partition P_{ϵ_1} such that for all partitions finer than P_{ϵ_1} ($P_{\epsilon_1} \subseteq P$) we have

$$|\mathcal{S}(P, h, \alpha) - A| < \epsilon_1 = \frac{\epsilon}{2}.$$

From the triangle inequality we get

$$\begin{aligned}
|\mathcal{S}(P, h\alpha', x) - A| &= |\mathcal{S}(P, h\alpha', x) - \mathcal{S}(p, h, \alpha) + \mathcal{S}(P, h, \alpha) - A| \\
&\leq |\mathcal{S}(P, h\alpha', x) - \mathcal{S}(p, h, \alpha)| + |\mathcal{S}(P, h, \alpha) - A| \\
&\leq \left| \sum_{i=1}^{n} h(\xi_i)\alpha'(\xi_i)[x_i - x_{i-1}] - \sum_{i=1}^{n} h(\xi_i)[\alpha(x_i) - \alpha(x_{i-1})] \right| + |\mathcal{S}(P, h, \alpha) - A|
\end{aligned}$$

From here we wish to use the Mean Value Theorem:

Mean Value Theorem:
Given $\alpha(x)$ continuous on $[a, b]$ and $\alpha'(x)$ exists on (a, b), then there exists a $c \in (a, b)$ where

$$\alpha'(c) = \frac{\alpha(b) - \alpha(a)}{b - a}$$

Proof of Mean Value Theorem. This proof may use Rolle's Theorem which I need to prove □

Since we have now proven the Mean Value Theorem we see that since $\alpha(x)$ is continuous on $[a, b]$ and $\alpha'(x)$ exists on (a, b) we have

$$\alpha'(c) = \frac{\alpha(b) - \alpha(a)}{b - a} \Rightarrow \ \alpha(b) - \alpha(a) = \alpha'(c)[b - a].$$

Since $[x_{i-1}, x_i] \subseteq [a, b]$ for all $i = 1, ..., n$ we see that we have

$$\alpha(x_i) - \alpha(x_{i-1}) = \alpha'(\xi_i')[x_i - x_{i-1}] \text{ for } \xi_i' \in [x_{i-1}, x_i] \text{ for all } i = 1, ..., n.$$

Subbing this in we have

$$\begin{aligned}
&\left|\sum_{i=1}^{n} h(\xi_i)\alpha'(\xi_i)[x_i - x_{i-1}] - \sum_{i=1}^{n} h(\xi_i)[\alpha(x_i) - \alpha(x_{i-1})]\right| \\
&= \left|\sum_{i=1}^{n} h(\xi_i)\alpha(\xi_i)[x_i - x_{i-1}] - \sum_{i=1}^{n} h(\xi_i)\alpha'(\xi_i')[x_i - x_{i-1}]\right| \\
&= \left|\sum_{i=1}^{n} h(\xi_i)(\alpha'(\xi_i) - \alpha'(\xi_i'))[x_i - x_{i-1}]\right|.
\end{aligned}$$

Since $h(x)$ is bounded on $[a, b]$ there exists an $M < \infty$ such that $|h(x)| \leq M$ for all $x \in [a, b]$. This means that $|h(\xi_i)| \leq M$ for all $i = 1, ..., n$. Therefore we have

$$\left|\sum_{i=1}^{n} h(\xi_i)(\alpha'(\xi_i) - \alpha'(\xi_i'))[x_i - x_{i-1}]\right| \leq M \left|\sum_{i=1}^{n} (\alpha'(\xi_i) - \alpha'(\xi_i'))[x_i - x_{i-1}]\right|.$$

Next we recall that since $\alpha'(x)$ exists and is continuous on $[a, b]$ we have $\alpha'(x)$ is uniformly continuous on $[a, b]$, which we proved earlier. This means that for $\epsilon_2 = \frac{\epsilon}{2M(b-a)} > 0$ there exists a $\delta > 0$ such that for any $x, y \in [a, b]$ that satisfy $|x - y| < \delta$ then $|\alpha'(x) - \alpha'(y)| < \frac{\epsilon}{2M(b-a)}$. Now we take a partition of $[a, b]$, call it P_{ϵ_2}, such that $||P_{\epsilon_2}|| < \delta$. We then have for all partitions P that are finer than P_{ϵ_2} ($P_{\epsilon_2} \subseteq P$) we have $||P|| \leq ||P_{\epsilon_2}|| < \delta$, so we have $|\xi_i - \xi_i'| < \delta$ for each $i = 1, ..., n$ since both $\xi_i, \xi_i' \in [x_{i-1}, x_i]$ for all sub intervals. Thus, by uniform continuity of $\alpha'(x)$ on $[a, b]$ we have $|\alpha'(x_i) - \alpha'(\xi_i')| < \frac{\epsilon}{2M(b-a)}$.
This means

$$\begin{aligned}
M\left|\sum_{i=1}^{n} (\alpha'(\xi_i) - \alpha'(\xi_i'))[x_i - x_{i-1}]\right| &\leq M \sum_{i=1}^{n} |\alpha'(\xi_i) - \alpha'(\xi_i')| \cdot |[x_i - x_{i-1}]| \\
&< M \frac{\epsilon}{2M(b-a)} \sum_{i=1}^{n} [x_i - x_{i-1}]
\end{aligned}$$

Since $\sum_{i=1}^{n} [x_i - x_{i-1}]$ is a telescoping sum it is equal to $b - a$.

$$M\left|\sum_{i=1}^{n} (\alpha'(\xi_i) - \alpha'(\xi_i'))[x_i - x_{i-1}]\right| < M \frac{\epsilon}{2M(b-a)}(b - a) = \frac{\epsilon}{2}.$$

Thus proving that

$$\left|\sum_{i=1}^{n} h(\xi_i)\alpha'(\xi_i)[x_i - x_{i-1}] - \sum_{i=1}^{n} h(\xi_i)[\alpha(x_i) - \alpha(x_{i-1})]\right| \leq \frac{\epsilon}{2} \tag{18}$$

for partitions finer than P_{ϵ_2}.
Therefore, taking $P_\epsilon = P_{\epsilon_1} \cup P_{\epsilon_2}$ we have that P_ϵ is finer than both P_{ϵ_1} and P_{ϵ_2}. From here we see that for all partitions finer than P_ϵ ($P_\epsilon \subseteq P$) we have

$$\begin{aligned}
|\mathcal{S}(P, h\alpha', x) - A| &\leq \left|\sum_{i=1}^{n} h(\xi_i)\alpha'(\xi_i)[x_i - x_{i-1}] - \sum_{i=1}^{n} h(\xi_i)[\alpha(x_i) - \alpha(x_{i-1})]\right| + |\mathcal{S}(P, h, \alpha) - A| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Since ϵ was arbitrarily chosen we see that as we take $\epsilon \to 0$ we get

$$\mathcal{S}(P, h\alpha', x) = A$$

Hence

$$\int_a^b h(x)\alpha'(x)\,dx = \int_a^b h(x)\,d\alpha(x).$$

Thus, concluding the proof of the Reducing Riemann-Stieltjes Integrals to Riemann Integrals lemma. □

Now we continue on the proof for the version of the second mean value theorem for integrals. From the second mean value theorem for Riemann-Stieltjes integrals we have

$$\int_a^b h(x)\,d\alpha(x) = h(a)\int_a^c d\alpha(x) + h(b)\int_c^b d\alpha(x)$$

for $c \in [a,b]$. Since $h(x)$ is a Riemann-Stieltjes integrable function with respect to α on the interval $[a,b]$ and bounded (since we will be taking our function to be in $L^1(\mathbb{T})$) and $\alpha'(x)$ exists and is continuous on $[a,b]$, which we proved earlier, we can enact the most recent lemma, therefore

$$\int_a^b h(x)\alpha'(x)\,dx = \int_a^b h(x)\,d\alpha(x)$$

Note that if $h(x) = 1$ for the reduction to Riemann integration we have $\int_a^b 1\,d\alpha(x) = \int_a^b 1 \cdot \alpha'(x)\,dx = \int_a^b \alpha'(x)\,dx$.
So then we have

$$\int_a^b h(x)\alpha'(x)\,dx = h(a)\int_a^c \alpha'(x)\,dx + h(b)\int_c^b \alpha'(x)\,dx$$

Remembering that $\alpha'(x) = \phi(x)$ gets us to

$$\int_a^b h(x)\phi(x)\,dx = h(a)\int_a^c \phi(x)\,dx + h(b)\int_c^b \phi(x)\,dx.$$

Lastly, we note that if $A = h(a)$ and $B = h(b)$ then the theorem has been proven. If not, we redefine $h(x)$ at the endpoints of a and b. Then the integral of $h(x)$ keeps its value since changing the value of $h(x)$ at a finite number of points does not change the value of the integral. However, if we ensure that $A, B \in \mathbb{R}$ are such that $A \le h(a+) \le h(b-) \le B$ then $h(x)$ is still increasing function so we have

$$\int_a^b h(x)\phi(x)\,dx = A\int_a^c \phi(x)\,dx + B\int_c^b \phi(x)\,dx$$

Proving the version of the second mean value theorem for integrals. □

Lastly, we will prove a corollary to this version of the second mean value theorem for integrals called Bonnet's Theorem:

Bonnet's Theorem:
Let $h(x)$ be an increasing function on $[a, b]$, let $\phi(x)$ be continuous on $[a, b]$, have $B \in \mathbb{R}$ be such that $h(b-) \leq B$ and $h(x) \geq 0$ for all $x \in [a, b]$. Then there exists a $c \in [a, b]$ such that

$$\int_a^b h(x)\phi(x)\,dx = B\int_c^b \phi(x)\,dx.$$

Proof of Bonnet's Theorem. Since $h(x) \geq 0$ for all $x \in [a, b]$ we have that $h(a+) = \lim_{n\to\infty} h(x) \geq 0$. So letting $A = 0$ and $B \in \mathbb{R}$ be such that $h(b-) \leq B$. From here by the above version of the second mean value theorem for integrals there exists a $c \in [a, b]$ such that

$$\begin{aligned}\int_a^b h(x)\phi(x)\,dx &= 0 \cdot \int_a^c \phi(x)\,dx + B\int_c^b \phi(x)\,dx \\ &= B\int_c^b \phi(x)\,dx.\end{aligned}$$

This concludes Bonnet's Theorem. $\square$

Now that we have proven Bonnet's Theorem we want to make sure that our integral fits the requirements.
Note that we can make $D_N(t)$ continuous on this interval. We see this by the following

$$\begin{aligned}\lim_{t\to 0}\frac{\sin(\pi(2N+1)t)}{\sin(\pi t)} &= \frac{0}{0} \text{ apply L'Hôpital's rule} \\ &= \lim_{t\to 0}\frac{\pi(2N+1)\cos(\pi(2N+1)t)}{\pi\cos(\pi t)} = \frac{\pi(2N+1)}{\pi} = 2N+1\end{aligned}$$

Since $D_N(t)$ is continuous everywhere except at 0 and the limit approaches $2N + 1$ we can rewrite $D_N(t)$ as $\tilde{D}_N(t)$ where

$$\tilde{D}_N(t) = \begin{cases} 2N+1, & \text{if } t = 0 \\ D_N(t), & \text{for } t \in (0, \delta] \end{cases}$$

which changes nothing of the integration, allowing us to use Bonnet's Theorem.
Therefore taking $\tilde{g}(t)$ to be our monotonic function and $\tilde{g}(t) \geq 0$ for $t \in [0, \delta]$ and $\tilde{D}_N(t)$ to be our continuous function we have for some ν, $0 < \nu < \delta$,

$$\int_0^\delta \tilde{g}(t)\tilde{D}_N(t)\,dt = \tilde{g}(\delta-)\int_\nu^\delta \tilde{D}_N(t)\,dt$$

In this case we had $B = \tilde{g}(\delta-)$ for Bonnet's Theorem.
This means that we have

$$\int_0^\delta \tilde{g}(t)\tilde{D}_N(t)\,dt = \tilde{g}(\delta-)\int_\nu^\delta \tilde{D}_N(t)\,dt$$

Furthermore,

$$\begin{aligned}
\left|\int_\nu^\delta \tilde{D}_N(t)\,dt\right| &= \left|\int_\nu^\delta \sin(\pi(2N+1)t)\left(\frac{1}{\sin(\pi t)} - \frac{1}{\pi t}\right) + \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&\leq \left|\int_\nu^\delta \sin(\pi(2N+1)t)\left(\frac{1}{\sin(\pi t)} - \frac{1}{\pi t}\right)\,dt\right| + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&\leq \int_\nu^\delta |\sin(\pi(2N+1)t)|\left|\left(\frac{1}{\sin(\pi t)} - \frac{1}{\pi t}\right)\right|\,dt + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&\leq \int_\nu^\delta 1\cdot\left|\left(\frac{1}{\sin(\pi t)} - \frac{1}{\pi t}\right)\right|\,dt + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&\leq \int_\nu^\delta \left|\frac{1}{\sin(\pi t)}\right| + \left|\frac{1}{\pi t}\right|\,dt + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right|
\end{aligned}$$

Now note that since $0 < \nu < \delta < 1/2$ we see that $t > 0$ for the second term of the first integral. We also remember from earlier that we proved that $\sin(\pi t)$ is strictly increasing on $\left(-\frac{1}{2}, \frac{1}{2}\right)$ and since $\sin(\pi t) = 0$ when $t = 0$ we see that $\sin(\pi t) > 0$ on the interval $[\nu, \delta]$. Which means we have

$$\begin{aligned}
\left|\int_\nu^\delta \tilde{D}_N(t)\,dt\right| &\leq \int_\nu^\delta \frac{1}{\sin(\pi t)}\,dt + \int_\nu^\delta \frac{1}{\pi t}\,dt + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&= \int_\nu^\delta \csc(\pi t)\,dt + \frac{1}{\pi}(\ln(\delta) - \ln(\nu)) + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right|
\end{aligned}$$

We will now integrate the first term using u-substitution letting $u = \pi t$ and $du = \pi dt$. We will also multiply by $\frac{(\csc(u)+\cot(u))}{(\csc(u)+\cot(u))} = 1$

$$\begin{aligned}
\left|\int_\nu^\delta \tilde{D}_N(t)\,dt\right| &\leq \frac{1}{\pi}\int_{\nu\pi}^{\delta\pi} \csc(u)\,du + \frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\right) + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&= \frac{1}{\pi}\int_{\nu\pi}^{\delta\pi} \frac{\csc(u)\cdot(\csc(u)+\cot(u))}{(\csc(u)+\cot(u))}\,du + \frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\right) + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right|
\end{aligned}$$

We then do one more change of variables to $v = \csc(u) + \cot(u)$ and $dv = -\cot(u)\csc(u) -$

$\csc(u)^2du = -\csc(u)(\csc(u)+\cot(u))du$ which gives us

$$\begin{aligned}
\left|\int_\nu^\delta \tilde{D}_N(t)\,dt\right| &\leq -\frac{1}{\pi}\int_{\csc(\nu\pi)+\cot(\nu\pi)}^{\csc(\delta\pi)+\cot(\delta\pi)} \frac{1}{v}\,dv + \frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\right) + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&= -\frac{1}{\pi}\left[\ln(\csc(\delta\pi)+\cot(\delta\pi)) - \ln(\csc(\nu\pi)+\cot(\nu\pi))\right] + \frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\right) + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&= \frac{1}{\pi}\ln\left(\frac{\csc(\nu\pi)+\cot(\nu\pi)}{\csc(\delta\pi)+\cot(\delta\pi)}\right) + \frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\right) + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| \\
&= \frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\cdot\frac{\csc(\nu\pi)+\cot(\nu\pi)}{\csc(\delta\pi)+\cot(\delta\pi)}\right) + \left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right|
\end{aligned}$$

Here we make a note that since $0 < \nu < \delta < 1/2$ we have the inside of each ln being positive therefore we have foregone the absolute values. Furthermore, we notice that $\csc(\pi t)+\cot(\pi t)$ is decreasing on the interval $[0,1]$ because its derivative, $-\pi\csc(\pi t)(\csc(\pi t)+\cot(\pi t))$, is negative for $t \in [0,1]$. Also, since it approaches $+\infty$ as $t \to 0^+$ and is equal to 0 at $t = 1$ we see that $\frac{\csc(\nu\pi)+\cot(\nu\pi)}{\csc(\delta\pi)+\cot(\delta\pi)} > 1$. Because of this, and the fact that $\delta > \nu$, we see that the first term is positive.

Now we turn our attention to the second integral which we will prove is finite by a change of variables and multiple inequalities. The change of variables is $v = \pi(2N+1)t$ and $dv = \pi(2N+1)dt$

$$\begin{aligned}
\left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| &= \left|\int_{\nu\pi(2N+1)}^{\delta\pi(2N+1)} \frac{\sin(v)}{\frac{v}{2N+1}}\cdot\frac{1}{\pi(2N+1)}\,dv\right| \\
&= \frac{1}{\pi}\left|\int_{\nu\pi(2N+1)}^{\delta\pi(2N+1)} \frac{\sin(v)}{v}\,dv\right| \\
&\leq \frac{1}{\pi}\int_{\nu\pi(2N+1)}^{\delta\pi(2N+1)} \left|\frac{\sin(v)}{v}\right|\,dv
\end{aligned}$$

Now note that since $|\sin(v)| \leq 1$ for all $v \in \mathbb{R}$ we have

$$\begin{aligned}
\left|\int_\nu^\delta \frac{\sin(\pi(2N+1)t)}{\pi t}\,dt\right| &\leq \frac{1}{\pi}\int_{\nu\pi(2N+1)}^{\delta\pi(2N+1)} \left|\frac{\sin(v)}{v}\right|\,dv \\
&\leq \frac{1}{\pi}\int_{\nu\pi(2N+1)}^{\delta\pi(2N+1)} \frac{1}{v}\,dv \\
&= \frac{1}{\pi}\ln\left(\delta\pi(2N+1)\right) - \ln\left(\nu\pi(2N+1)\right) \\
&= \frac{1}{\pi}\ln\left(\frac{\delta\pi(2N+1)}{\nu\pi(2N+1)}\right) \\
&= \frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\right)
\end{aligned}$$

Therefore we have

$$\left|\int_{\nu}^{\delta} \tilde{D}_N(t)\,dt\right| \le \frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\cdot\frac{\csc(\nu\pi)+\cot(\nu\pi)}{\csc(\delta\pi)+\cot(\delta\pi)}\right)+\frac{1}{\pi}\ln\left(\frac{\delta}{\nu}\right)$$
$$= \frac{1}{\pi}\ln\left(\frac{\delta^2}{\nu^2}\cdot\frac{\csc(\nu\pi)+\cot(\nu\pi)}{\csc(\delta\pi)+\cot(\delta\pi)}\right)$$

Hence (since $0<t<\delta$ then $\tilde{g}(t)<\epsilon \Rightarrow\ \tilde{g}(\delta-)<\epsilon$)

$$\int_0^{\delta}\tilde{g}(t)\tilde{D}_N(t)\,dt = \tilde{g}(\delta-)\int_{\nu}^{\delta}\tilde{D}_N(t)\,dt \le \frac{1}{\pi}\ln\left(\frac{\delta^2}{\nu^2}\cdot\frac{\csc(\nu\pi)+\cot(\nu\pi)}{\csc(\delta\pi)+\cot(\delta\pi)}\right)\epsilon$$

Meaning

$$\lim_{N\to\infty}\int_0^{1/2}[g(t)-g(0+)]D_N(t)\,dt \le \frac{1}{\pi}\ln\left(\frac{\delta^2}{\nu^2}\cdot\frac{\csc(\nu\pi)+\cot(\nu\pi)}{\csc(\delta\pi)+\cot(\delta\pi)}\right)\epsilon$$

which goes to 0 as we take $\epsilon\to 0$.
From this we see for $h(t)=f(x-t)$ and $h(0+)=f(x-)$ the same procedure proves

$$\lim_{N\to\infty}\int_0^{1/2}\left[h(t)-h(0+)\right]D_N(t)\,dt = 0$$

with the use of the changed function

$$\tilde{h}(t)=\begin{cases}0, & \text{if } t=0\\ h(t)-h(0+), & \text{for } t\in(0,\delta].\end{cases}$$

This proves that

$$\lim_{N\to\infty}S_Nf(x)=\lim_{N\to\infty}\int_0^{1/2}[f(x-t)+f(x+t)]D_N(t)\,dt=\frac{1}{2}g(0+)+\frac{1}{2}h(0+)=\frac{1}{2}\left[f(x+)+f(x-)\right].$$

Which means we have proven the theorem for monotonically increasing functions. To prove it for functions of bounded variations we remember that a function of bounded variation, $f(x)$, can be written as the difference of two monotonically increasing functions: $f(x)=f_1(x)-f_2(x)$ where $f_1(x)$ and $f_2(x)$ are monotonically increasing. From here we prove the theorem

$$
\begin{aligned}
\lim_{N\to\infty} S_N f(x) &= \lim_{N\to\infty} \int_0^{1/2} [f(x-t) + f(x+t)] D_N(t)\, dt \\
&= \lim_{N\to\infty} \int_0^{1/2} \left[(f_1(x-t) - f_2(x-t)) + (f_1(x+t) - f_2(x+t)) \right] D_N(t)\, dt \\
&= \lim_{N\to\infty} \int_0^{1/2} [f_1(x-t) + f_1(x+t)] D_N(t)\, dt - \lim_{N\to\infty} \int_0^{1/2} [f_2(x-t) + f_2(x+t)] D_N(t)\, dt \\
&= \frac{1}{2} g_1(0+) + \frac{1}{2} h_1(0+) - \left(\frac{1}{2} g_2(0+) + \frac{1}{2} h_2(0+) \right) \\
&= \frac{1}{2} \left[g_1(0+) - g_2(0+) + h_1(0+) - h_2(0+) \right] \\
&= \frac{1}{2} \left[f_1(x+) - f_2(x+) + f_1(x-) - f_2(x-) \right] \\
&= \frac{1}{2} \left[f(x+) + f(x-) \right]
\end{aligned}
$$

Completely proving the theorem for functions of bounded variations. $\square$

3 References

References

[1] Duoandikoetxea, J., & Cruz-Uribe, D. (2000). Fourier analysis Javier Duoandikoetxea. transl. and rev. by David Cruz-Uribe. American Mathematical Society.