The Fundamental Theorem of Calculus

not the most rigorous proof

Sai Sivakumar

March 28, 2021

Outline

Riemann integrals

To begin...

We will briefly discuss what the conditions are for a function to be Riemann integrable and how the Riemann integral is constructed.

Riemann integrability. The details for Riemann integrability can be confusing, so I will omit the details. Basically, we want functions $f : [a, b] \to \mathbb{R}$ to be bounded and piecewise continuous almost everywhere, that is, we can have a few discontinuities so long as they're not *horrible*.

Some functions which are not Riemann integrable

- $f(x) = \frac{1}{x}$ on $[0, 1]$
- $g(x) = \begin{cases} 0 & \text{if } x \text{ is irrational} \\ 1 & \text{if } x \text{ is rational} \end{cases}$

Why? These discontinuities are just *really bad*, but there are more rigorous explanations that you would get in a real analysis course.

The Riemann integral

Here we will motivate the definition by constructing the Riemann integral of a Riemann integrable function $f : [a, b] \to \mathbb{R}$.

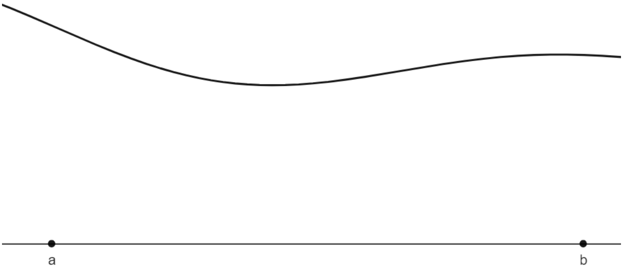
First begin by *partitioning* the interval $[a, b]$ into a set of n subsets (or subintervals) like so:

We call this set of subintervals P, and the "length" of the longest subinterval in P is called the *norm* or *mesh* of P, $\max\{|x_i - x_{i-1}| : 1 \leq i \leq n\}$ denoted as $|P|$.

Then consider

$$\sum_{i=1}^{n} f(t_i)(x_i - x_{x-1}),\ t_i \in [x_{i-1}, x_i]\,.$$

This summation can be interpreted graphically as the sum of rectangular areas, where $f(t_i)$ is the height and $(x_i - x_{i-1})$ is the length.

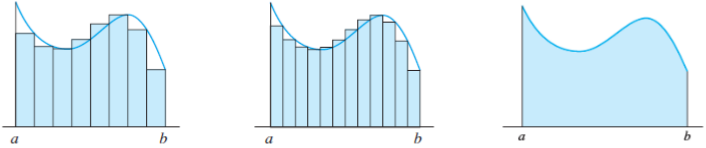


Then consider what happens when we let n become arbitrarily large. The partition of the interval $[a, b]$ would then be made of arbitrarily small subintervals, meaning their lengths approach zero.

Since $|P|$ is the length of the longest subinterval, it too will approach zero.

So we say that as n tends to infinity, or that $|P|$ tends to zero, that the sum is equivalent to the definite integral of f over $[a, b]$. So

$$\lim_{\substack{n\to\infty \\ (|P|\to 0)}} \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) = \int_a^b f(x)\,\mathrm{d}x\,.$$



Statement of the theorem

The statement

Theorem (Fundamental Theorem of Calculus)

If $f : [a, b] \to \mathbb{R}$ is Riemann integrable, then:

(1) F given by

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$

is continuous.

(2) Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x) = f(x)$$

whenever f is continuous at x.

Theorem (Antiderivative theorem)

Again, let $f : [a, b] \to \mathbb{R}$ *be Riemann integrable.*

(3) Let G *be any antiderivative of* f*, that is,* $\frac{\mathrm{d}G}{\mathrm{d}x} = f(x)$*. Then*

$$\int_a^b f(x)\,\mathrm{d}x = G(b) - G(a).$$

So (1) states that the integral as a function is always continuous. I will not prove this, but I will give an intuitive explanation. Then (2) states that we can differentiate F wherever f is continuous, and (3) is our familiar result where we use an antiderivative to evaluate definite integrals.

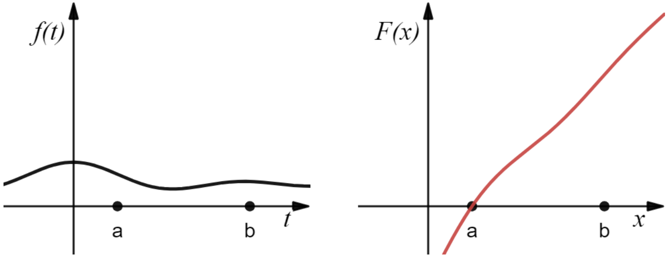
The proof

The continuity of $F(x)$, intuitively

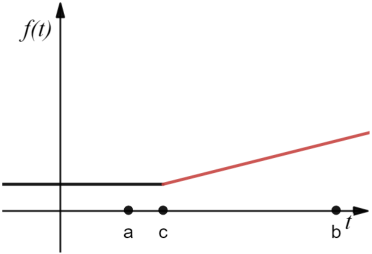
Here we will discuss how part (1) of the theorem works.

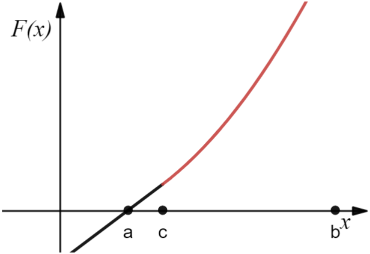
Visually, the quantity $F(x)$ represents the *accumulation* of the signed area under the graph of f from a to x.

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$



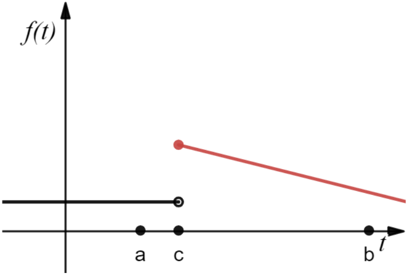
Piecewise continuous example:

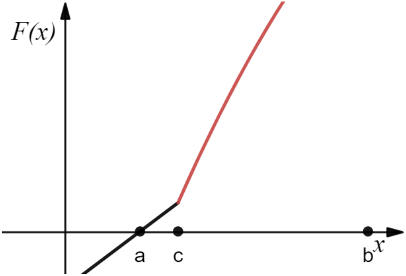




(Remember that $F(x) = \int_a^x f(t)\,\mathrm{d}t$.)

We can imagine how F is continuous whenever f is, but how can we think about what happens when f is discontinuous? An example:





Thus concludes our discussion of part (1) of the theorem, which states that F is continuous.

Differentiating F whenever f is continuous

Now we move into part (2) of the theorem.

By definition,

$$\frac{\mathrm{d}F}{\mathrm{d}x} = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{\int_a^{x+h} f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t}{h}$$

$$= \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)\,\mathrm{d}t\,.$$

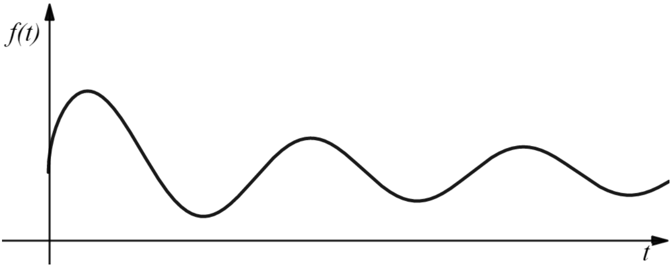
We must show that this limit exists and is equal to $f(x)$.

Consider the following quantities

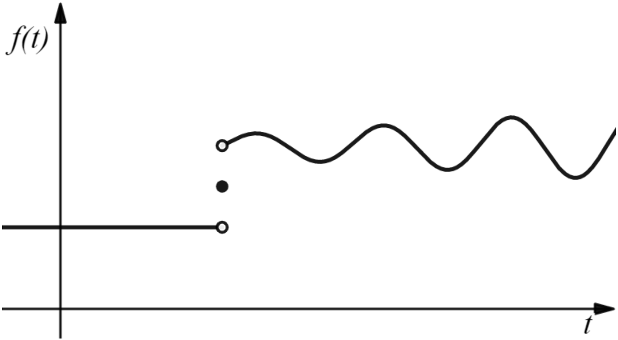
$$m(x,h) = \inf\{f(s) : |s-x| \le |h|\} \tag{1}$$

$$M(x,h) = \sup\{f(s) : |s-x| \le |h|\}, \tag{2}$$

which can be graphically thought of as values that bound* f in the neighborhood $[x-h, x+h]$.



When f is continuous at x, $m(x,h)$ and $M(x,h)$ converge to $f(x)$ as h tends to 0. We can imagine this working out with a convenient example.

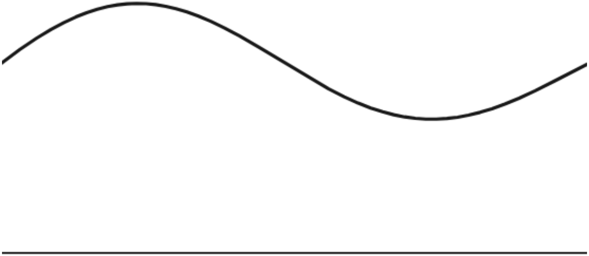


So more explicitly, $\lim_{h \to 0} m(x,h) = f(x)$ and $\lim_{h \to 0} M(x,h) = f(x)$.

We have that

$$\int_x^{x+h} m(x,h)\,\mathrm{d}t \leq \int_x^{x+h} f(t)\,\mathrm{d}t \leq \int_x^{x+h} M(x,h)\,\mathrm{d}t$$

since $m(x,h) \leq f(x) \leq M(x,h)$.



Thus

$$\frac{1}{h}\int_x^{x+h} m(x,h)\,\mathrm{d}t \le \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t \le \frac{1}{h}\int_x^{x+h} M(x,h)\,\mathrm{d}t\,,$$

and notice that since $m(x,h)$ is constant with respect to t, we have that

$$\frac{1}{h}\int_x^{x+h} m(x,h) = \frac{m(x,h)}{h}\,(x+h-x) = m(x,h).$$

Similarly $\frac{1}{h}\int_x^{x+h} M(x,h)\,\mathrm{d}t = M(x,h)$.

So then

$$m(x,h) \le \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t \le M(x,h),$$

and *when f is continuous at x* we can use the fact that $m(x,h)$ and $M(x,h)$ converge to $f(x)$ as h tends to zero to find that

$$\lim_{h\to 0} m(x,h) \le \lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t \le \lim_{h\to 0} M(x,h)$$

becomes

$$f(x) \le \frac{\mathrm{d}F}{\mathrm{d}x} \le f(x) \iff \frac{\mathrm{d}}{\mathrm{d}x}F = f(x),$$

which is part (2) of the theorem.

Using antiderivatives to compute definite integrals

We are (finally) beginning discussion of part (3) of the theorem. This part is going to be a little more algebraic.

We need to define what an antiderivative is. An antiderivative G of $f : [a, b] \to \mathbb{R}$ satisfies

$$G'(x) = f(x)$$

for all $x \in [a, b]$.

A nice thing to think about is *when* functions have antiderivatives.

We also have that antiderivatives $G(x)$ differ from the indefinite integral $\int_a^x f(t)\,\mathrm{d}t$ by a constant, that is,

$$G(x) = \int_a^x f(t)\,\mathrm{d}t + C$$

for some constant C.

Similarly as before, partition $[a, x]$, which is only part of $[a, b]$, into n subintervals, where $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = x$.

Consider the following quantity

$$\frac{G(x_k) - G(x_{k-1})}{x_k - x_{k-1}},$$

and note that due to the mean value theorem, the above quantity is equal to $G'(t_k) = f(t_k)$, where $t_k \in [x_{k-1}, x_k]$. This is true *no matter how small the subinterval* $[x_{k-1}, x_k]$ *is.*

So then we have

$$G(x_k) - G(x_{k-1}) = f(t_k)(x_k - x_{k-1})$$

Then we can sum over all of the subintervals, so

$$\sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}) = \sum_{k=1}^{n} G(x_k) - G(x_{k-1})$$

$$= (G(x_1) - G(x_0)) + (G(x_2) - G(x_1)) + (G(x_3) - G(x_2)) + \cdots + (G(x_{n-1}) - G(x_{n-2})) + (G(x_n) - G(x_{n-1}))$$

$$= G(x_n) - G(x_0) = G(x) - G(a)$$

Since $G(x_k) - G(x_{k-1}) = f(t_k)(x_k - x_{k-1})$ holds regardless of the size of the subinterval $[x_{k-1}, x_k]$, we can consider taking the following limit:

$$\lim_{n\to\infty} \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}) = \lim_{n\to\infty} G(x) - G(a)$$

But the left hand side is the definition of the integral from a to x of $f(t)$, and the right hand side remains unchanged. So then the previous equation becomes

$$\int_a^x f(t)\,\mathrm{d}t = G(x) - G(a).$$

So

$$G(x) = \int_a^x f(t)\,\mathrm{d}t + G(a),$$

where $G(a)$ is a constant. This proves part (3) of the theorem.

Notice that if we go back one step and evaluate both sides of the equation at $x = b$, we get the familiar formula we learned to evaluate integrals.

$$\int_a^b f(t)\,\mathrm{d}t = G(b) - G(a)$$

Closing remarks and/or questions

Extra slide in case I wanted to write something down