5: INTEGRATION

STEVEN HEILMAN

1. Introduction

The integrals which we have obtained are not only general expressions which satisfy the differential equation, they represent in the most distinct manner the natural effect which is the object of the phenomenon... when this condition is fulfilled, the integral is, properly speaking, the equation of the phenomenon; it expresses clearly the character and progress of it, in the same manner as the finite equation of a line or curved surface makes known all the properties of those forms.

Jean-Baptiste-Joseph Fourier, Théorie Analytique de la Chaleur, 1822.

Along with the derivative, the integral is one of the two most fundamental concepts that we find in Calculus. Unfortunately, the formal definition of the integral is more complicated than that of the derivative. However, we should still try to understand these formal definitions, since the ideas that go into the construction of the derivative and the integral are pervasive throughout mathematics and the sciences. In the case of the integral, the quantity $\int_a^b f(x)dx$ intuitively represents the area under the curve $y = f(x)$ on the interval $[a, b]$. We will make more precise statements about the integral in Section 3.

Special attention should be given to the Fundamental Theorem of Calculus in Section 5. This Theorem truly is fundamental, as it provides a link between our two fundamental concepts, the derivative and the integral.

In this course we have seen an important though minute part of mathematics. For example, we are unable to discuss the integrals that Fourier has referenced above. His discovery is arguably one of the most influential mathematical findings since the creation of Calculus, yet Fourier's work is still around 200 years old. A great deal has happened since the time of Fourier. Many applications of mathematics to the sciences were discovered, and there are still many areas being investigated today in pure and applied mathematics. However, for now we discuss the antiderivative, which serves as a gentle introduction to the integral.

2. Antiderivatives

Recall the following proposition from the previous set of notes.

Proposition 2.1. Let $f, g\colon (a, b) \to \mathbb{R}$ be differentiable functions. If $f'(x) = g'(x)$ for all $x \in (a, b)$, then there is a constant $C \in \mathbb{R}$ such that $f(x) = g(x) + C$.

Definition 2.2. Let $f\colon [a, b] \to \mathbb{R}$. We say that F is an antiderivative of f if F is differentiable, and for all $x \in (a, b)$ we have $F'(x) = f(x)$.

Date: December 6, 2012.

Remark 2.3. Let F, G be antiderivatives of f, so that $F'(x) = G'(x) = f$. From Proposition 2.1, there must be a constant C such that $F(x) = G(x)+C$. So, if we have one antiderivative F of f, then the set of all antiderivatives of f is given by the set of all $F(x) + C$, $C \in \mathbb{R}$.

Example 2.4. Let $f(x) = x^2$. Then the set of all antiderivatives of f is given by $F(x) = (1/3)x^3 + C$, $C \in \mathbb{R}$.

Example 2.5. (**Idealized Trajectories**) Suppose I throw a ball straight up in the air at a velocity v_0 m/s, with initial vertical position s_0 meters. We ignore air friction. Suppose the ball has mass m kg. The only acceleration that acts on the ball is a constant acceleration due to gravity, of roughly $a(t) = -10$ m/s^2, where t is the time after the ball is thrown, measured in seconds. Taking the antiderivative and using Remark 2.3, the ball must have velocity $v(t) = -10t + C$. Since $v(0) = v_0$, we conclude that $v(t) = -10t + v_0$. Taking the antiderivative again and using Remark 2.3, the ball must have position $s(t) = -5t^2 + v_0 t + C$. My initial vertical position is s_0, so we conclude that the ball has position

$$s(t) = -5t^2 + v_0 t + s_0.$$

For now, antiderivatives may seem a bit strange. Also, if we have a function f, how can we know whether or not an antiderivative exists? It turns out that, if f is continuous, then you can create an antiderivative of f by measuring the areas under the curve f. So, calculating areas under a curve has the "opposite" effect of taking a derivative of a curve. This statement will be made more precise when we state the Fundamental Theorem of Calculus. For now, we give a precise description of how to find the area under a curve, via the Riemann integral.

3. The Riemann Integral

Our presentation of the construction of the Riemann integral is slightly different than the textbook. Though the presentations differ, they are equivalent. Unfortunately, any construction of the Riemann integral requires a lot of notation. Fortunately, there is a nice geometric picture of the construction of the integral. We recommend seeing this picture in action with the help of the MIT JAVA applet, Riemann sums.

Ultimately, we want to find the area under a given curve. The strategy is similar in spirit to our construction of the derivative. We would like to perform some process that requires infinitely many steps, and as noted by Zeno, doing so does not make any sense. To resolve this issue, we approximate some infinite thing by a finite number of steps. And we hope that, as our approximation gets "finer," some number will approach some limit.

Using this paradigm, we first approximate the area under a given curve by a finite number of rectangles. We know the area of a rectangle, so we therefore know the area of several non-overlapping rectangles. We then want to make our approximation of rectangles finer and finer, and then take some limit. If we complete this process in the right way, and if our curve is nice enough, then this limit will exist. Unfortunately, the limit of the sum of the areas of these rectangles may not always exist, so we have to be careful in our construction of the integral. So, although the notation below and the details may appear pedantic or unnecessary, these things really are necessary in order to get a sensible answer in the end.

Definition 3.1. Let $a < b$. A **partition** P of the interval $[a, b]$ is a set of points $x_0, x_1, \ldots, x_n$ such that

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

A partition P defines the rectangles that approximate the area under a curve. See Remark 3.3 below.

Let P, P' be partitions, so that $P = \{x_i\}_{i=0}^n$ and $P' = \{y_i\}_{i=0}^m$. We say that P is a **refinement** of the partition P' if the set of points P contains the set of points P'.

Definition 3.2. Let $a < b$. Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function, and let P be a partition of $[a,b]$. By assumption, f is continuous on each closed interval $[x_i, x_{i-1}]$ of the partition, for $i = 1,\ldots,n$. So, the Extreme Value Theorem says that f assumes its minimum and maximum values on each such interval. We denote the minimum and maximum values of f on each such interval by $\min_{y\in[x_i,x_{i-1}]} f(y)$ and $\max_{y\in[x_i,x_{i-1}]} f(y)$, respectively. Define the **lower Riemann sum** $L(f,P)$ by the following formula

$$L(f,P) = \sum_{i=1}^{n} (x_i - x_{i-1}) \left(\min_{y\in[x_i,x_{i-1}]} f(y) \right).$$

Define the **upper Riemann sum** $U(f,P)$ by the following formula

$$U(f,P) = \sum_{i=1}^{n} (x_i - x_{i-1}) \left(\max_{y\in[x_i,x_{i-1}]} f(y) \right).$$

Remark 3.3. Draw a picture representing the the upper and lower Riemann sums. For $i = 1,\ldots,n$, the number $x_i - x_{i-1}$ denotes the width of a rectangle that sits on the interval $[x_i, x_{i-1}]$. In the case of $L(f,P)$, $\min_{y\in[x_i,x_{i-1}]} f(y)$ represents the height of the rectangle that sits on $[x_i, x_{i-1}]$. And we see that $L(f,P)$ should be less than the area under the curve of f on $[a,b]$. Similarly, in the case of $U(f,P)$, $\max_{y\in[x_i,x_{i-1}]} f(y)$ represents the height of the rectangle that sits on $[x_i, x_{i-1}]$. And we see that $U(f,P)$ should be more than the area under the curve of f on $[a,b]$. Note that $L(f,P) \leq U(f,P)$. The upper and lower Riemann sums will serve as an approximation of the integral.

Definition 3.4. Let $a < b$, and let P be a partition of $[a,b]$. The maximum spacing of the partition is denoted by

$$\max_{i=1,\ldots,n} (x_i - x_{i-1}).$$

This number is the maximum of the numbers $(x_1 - x_0), (x_2 - x_1), (x_3 - x_2), \ldots, (x_n - x_{n-1})$. If $\max_{i=1,\ldots,n}(x_i - x_{i-1})$ is small, then the partition is very fine. More specifically, all of our approximating rectangles will have small width. In order to construct the integral, we will let $\max_{i=1,\ldots,n}(x_i - x_{i-1})$ approach zero.

We can finally define the Riemann Integral.

Definition 3.5. (**The Riemann Integral**) Let $f\colon [a,b] \to \mathbb{R}$. If the following two limits exist and are equal, we say that f is integrable on $[a,b]$.

$$\lim_{\left(\max_{i=1,\ldots,n} (x_i - x_{i-1})\right) \to 0} L(f,P),$$

$$\lim_{\left(\max_{i=1,\ldots,n} (x_i - x_{i-1})\right) \to 0} U(f,P).$$

If these limits exist and are equal, we denote both limits with the notation $\int_a^b f(x)dx$. That is, we have

$$\int_a^b f(x)dx = \lim_{\left(\max_{i=1,\ldots,n}(x_i - x_{i-1})\right) \to 0} L(f,P) = \lim_{\left(\max_{i=1,\ldots,n}(x_i - x_{i-1})\right) \to 0} U(f,P).$$

Remark 3.6. We now describe the limit appearing in the Riemann Integral more explicitly. For $\lim_{[\max_{i=1,\ldots,n}(x_i - x_{i-1})] \to 0} L(f,P)$ to exist and be equal to I, we mean the following. For every $\varepsilon > 0$, there exists a $\delta(\varepsilon) > 0$ such that: for all partitions P with $\max_{i=1,\ldots,n}(x_i - x_{i-1}) < \delta(\varepsilon)$, we have

$$|L(f,P) - I| < \varepsilon.$$

That is, for the limit $\lim_{[\max_{i=1,\ldots,n}(x_i - x_{i-1})] \to 0} L(f,P)$ to exist, we require that any sufficiently fine partition has a lower Riemann sum that is close to the value I.

Similarly, for $\lim_{[\max_{i=1,\ldots,n}(x_i - x_{i-1})] \to 0} U(f,P)$ to exist and be equal to U, we mean the following. For every $\varepsilon > 0$, there exists a $\delta(\varepsilon) > 0$ such that: for all partitions P with $\max_{i=1,\ldots,n}(x_i - x_{i-1}) < \delta(\varepsilon)$, we have

$$|U(f,P) - U| < \varepsilon.$$

Remark 3.7. Strictly speaking, we should replace the min in Definition 3.2 with an infimum, and we should replace the max in Definition 3.2 with a supremum. We have made a similar omission in the previous notes, though we will not really need to use these concepts, which is why we have not mentioned them. For example, for a continuous function, the maximum and supremum are equivalent, and the minimum and infimum are equivalent.

We will now explore some consequences of Definition 3.5. We first show that, if f is Riemann integrable, then any Riemann sum will approximate the integral. Then, we will show an extremely important fact, that tells us when the Riemann integral actually exists. If $f\colon [a,b] \to \mathbb{R}$ is continuous, then $\int_a^b f(x)dx$ exists.

Definition 3.8. Let $a < b$. Let $f\colon [a,b] \to \mathbb{R}$, and let P be a partition of $[a,b]$. For $i = 1, \ldots, n$, let $y_i \in [x_i, x_{i-1}]$. Then a **Riemann sum** is an expression of the following form

$$\sum_{i=1}^{n} (x_i - x_{i-1}) f(y_i).$$

Proposition 3.9. Let $f\colon [a,b] \to \mathbb{R}$ be integrable. Then any Riemann sum converges to $\int_a^b f(x)dx$, as the partition becomes finer. That is, for $y_i \in [x_i, x_{i-1}]$, $i = 1, \ldots, n$,

$$\int_a^b f(x)dx = \lim_{\left(\max_{i=1,\ldots,n}(x_i - x_{i-1})\right) \to 0} \sum_{i=1}^{n} (x_i - x_{i-1}) f(y_i) \qquad (*)$$

Proof. Let $P = \{x_i\}_{i=0}^{n}$ be a partition. Then, by Definition 3.5,

$$L(f,P) \leq \sum_{i=1}^{n} (x_i - x_{i-1}) f(y_i) \leq U(f,P). \qquad (**)$$

Now, let $\max_{i=1,\ldots,n}(x_i - x_{i-1}) \to 0$. Since f is integrable on $[a,b]$, the left and right side of $(**)$ approach $\int_a^b f(x)dx$. So, by the Squeeze Theorem,

$$\int_a^b f(x)dx = \lim_{\left(\max_{i=1,\ldots,n}(x_i - x_{i-1})\right) \to 0} L(f,P)$$
$$\leq \lim_{\left(\max_{i=1,\ldots,n}(x_i - x_{i-1})\right) \to 0} \sum_{i=1}^{n}(x_i - x_{i-1})f(y_i)$$
$$\leq \lim_{\left(\max_{i=1,\ldots,n}(x_i - x_{i-1})\right) \to 0} U(f,P) = \int_a^b f(x)dx.$$

We conclude that $(*)$ holds, as desired. □

For a function $f\colon [a,b] \to \mathbb{R}$, we do not yet have a way to determine whether or not $\int_a^b f(x)dx$ exists. Thankfully, if $f\colon [a,b] \to \mathbb{R}$ is continuous, then $\int_a^b f(x)dx$ exists, as the following very important theorem shows. However, there are situations where the integral of a function does not exist. We will investigate these situations more in the Exercises and Problems below. If we understand when integrals do not exist, then our understanding of the integral is improved, just as an understanding of nonexistence of derivatives improves our understanding of derivatives. We should also mention that the closed interval condition is crucial in Theorem 3.10. See Exercise 8.9(6).

Theorem 3.10. ***(Continuous Functions on Closed Intervals are Integrable)*** *Let $a < b$. Let $f\colon [a,b] \to \mathbb{R}$ be continuous. Then $\int_a^b f(x)dx$ exists.*

Since a composition of continuous functions is continuous, and a product of continuous functions is continuous, we have the following Corollary of Theorem 3.10.

Corollary 3.11. *Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be continuous. Then $\int_a^b f(g(x))dx$ exists, and $\int_a^b f(x)g(x)dx$ exists.*

Proof of Theorem 3.10. The function $f\colon [a,b] \to \mathbb{R}$ is continuous. So, for every $\varepsilon > 0$ and for every $x \in [a,b]$, there is a $\delta(x,\varepsilon) > 0$ such that, if $y \in [a,b]$ satisfies $|y - x| < \delta(x,\varepsilon)$, then $|f(x) - f(y)| < \varepsilon$. For convenience, we label the set of $y \in [a,b]$ such that $|y - x| < \delta(x,\varepsilon)$ by $B(x,\delta) = B(x,\delta(x,\varepsilon))$. That is, $B(x,\delta)$ is the set of points in $[a,b]$ within distance δ of x. We claim: there exists a number $\alpha > 0$ such that, for any $y \in [a,b]$, there exists an $x \in [a,b]$ such that $B(x,\delta(x,\varepsilon))$ contains $B(y,\alpha)$. To prove this claim, we argue by contradiction.

Suppose our claim is false. By negating the claim, we have: for any $n \in \mathbb{Z}$, $n > 0$, there is a $y_n \in [a,b]$ such that $B(y_n, 1/n)$ is not contained in $B(x,\delta(x,\varepsilon))$ for all $x \in [a,b]$. By repeating the proof of the Extreme Value Theorem from the third set of notes, there is a $y \in [a,b]$ such that $y_n \to y$ as $n \to \infty$. In particular, there is an $N \in \mathbb{Z}$ such that, for all $n \geq N$, we have $|y_n - y| < \delta(y,\varepsilon)/2$. Let $M \in \mathbb{Z}$ such that $1/M < \delta(y,\varepsilon)/2$. Then for $n \geq N$ and $n \geq M$, we have $|y_n - y| < \delta(y,\varepsilon)/2$, and $1/n < \delta(y,\varepsilon)/2$. Therefore, $B(y_n, 1/n)$ is contained in $B(y,\delta(y,\varepsilon))$. But this containment contradicts our definition of the points y_n. Since we have achieved a contradiction, we conclude that our claim is true. There is some $\alpha > 0$ such that, for any $y \in [a,b]$, there is some $B(x,\delta(x,\varepsilon))$ that contains $B(y,\alpha)$.

With this claim proven, we can conclude the proof of the theorem. Let $\alpha > 0$ as given by the Claim. Let $J \in \mathbb{Z}$, $n > 0$ such that $(b-a)/J < \alpha/2$. For $i = 0, 1, \ldots, J$, let $y_i = (b-a)(i/J) + a$. Then every $y \in [a, b]$ is contained in some $B(y_i, \alpha)$, $i = 0, 1, \ldots, J$, since $y_i - y_{i-1} = (b-a)/J < \alpha/2$. By the Claim, every $B(y_i, \alpha)$ is contained in some $B(x, \delta(x, \varepsilon))$. So, applying the definition of the $\delta(x, \varepsilon)$, we have the following statement. For every $\varepsilon > 0$ there is an $\alpha(\varepsilon) > 0$ such that, for every $x \in [a, b]$, if $y \in [a, b]$ satisfies $|x - y| < \alpha(\varepsilon)$ then $|f(x) - f(y)| < \varepsilon$. Note that this statement is similar to, but stronger than, the usual definition of continuity, since the $\alpha(\varepsilon)$ term does not depend on x.

Now, for $j = 1, 2, 3, \ldots$, let P_j be the partition of $[a, b]$ such that $P_j = \{x_{i,j}\}_{i=1}^{2^j} = \{(b - a)(i2^{-j}/n) + a\}_{i=0}^{2^j}$. Note that P_{j+1} is a refinement of P_j. Let $\varepsilon > 0$ and let $\alpha(\varepsilon)$ as guaranteed by our strengthened form of continuity. Then let $j \in \mathbb{Z}$ such that $2^{-j}(b-a) < \alpha(\varepsilon)$. As a result, note that $|x_{i,j} - x_{i-1,j}| = 2^{-j}(b-a) < \alpha(\varepsilon)$. So, using Definition 3.2 and our strengthened form of continuity, we have

$$\begin{aligned} L(f, P_j) &= \sum_{i=1}^{n} (x_{i,j} - x_{i-1,j}) \left(\min_{y \in [x_{i,j}, x_{i-1,j}]} f(y) \right) \leq \sum_{i=1}^{n} (x_{i,j} - x_{i-1,j}) \left(\max_{y \in [x_{i,j}, x_{i-1,j}]} f(y) - \varepsilon \right) \\ &= U(f, P_j) - (b-a)\varepsilon \end{aligned}$$

Since P_{j+1} is a refinement of P_j, Definition 3.2 shows that $L(f, P_{j+1}) \geq L(f, P_j)$. Similarly, $U(f, P_{j+1}) \leq U(f, P_j)$. Also, Definition 3.2 shows that $L(f, P_j) \leq U(f, P_j)$. So, the sequence of numbers $L(f, P_j)$ increases for $j = 1, 2, 3, \ldots$, and the sequence of numbers $U(f, P_j)$ decreases for $j = 1, 2, 3, \ldots$. And for any $\varepsilon > 0$, if j is sufficiently large, then $0 \leq U(f, P_j) - L(f, P_j) < (b-a)\varepsilon$. So, letting $\varepsilon \to 0$, we see that, as $j \to \infty$, $U(f, P_j)$ and $L(f, P_j)$ approach the same limit I. To conclude the theorem, it therefore remains to show that any sufficiently fine partition has upper and lower Riemann sums that approach I.

Let $\varepsilon > 0$ and let $P = \{x_i\}_{i=0}^{n}$ be any partition of $[a, b]$ with $\max_{i=1,\ldots,n}(x_i - x_{i-1}) < \varepsilon$. Consider the partition P' that consists of the points in P and of the points in P_j. Then P' is a refinement of P and P' is a refinement of P_j. By Definition 3.2, we therefore have

$$L(f, P') \geq L(f, P),\ L(f, P') \geq L(f, P_j),\ U(f, P') \leq U(f, P),\ U(f, P') \leq U(f, P_j).$$

For j sufficiently, large, $0 \leq U(f, P_j) - L(f, P_j) \leq (b-a)\varepsilon$. Also, for j sufficiently large, $|U(f, P_j) - I| < \varepsilon$ and $|L(f, P_j) - I| < \varepsilon$. Combining these inequalities,

$$L(f, P_j) - I \leq L(f, P') - I \leq U(f, P') - I \leq U(f, P_j) - I.$$

We therefore conclude that $|L(f, P') - I| < \varepsilon$ and $|U(f, P') - I| < \varepsilon$.

Since $\max_{i=1,\ldots,n}(x_i - x_{i-1}) < \varepsilon$, we have $0 \leq U(f, P) - L(f, P) \leq (b-a)\varepsilon$, using strengthened continuity as above. Combining this inequality with our inequalities above,

$$L(f, P) \geq U(f, P) - (b-a)\varepsilon \geq U(f, P') - (b-a)\varepsilon \geq I - (b - a + 1)\varepsilon.$$

Similarly,

$$U(f, P) \leq L(f, P) + (b-a)\varepsilon \leq L(f, P') + (b-a)\varepsilon \leq I + (b - a + 1)\varepsilon.$$

In summary,

$$I - (b - a + 1)\varepsilon \leq L(f, P) \leq U(f, P) \leq I + (b - a + 1)\varepsilon$$

So, P is an arbitrary partition, and if it is sufficiently fine, its upper and lower Riemann sums are close to the number I. Therefore, by Definition 3.5 and Remark 3.6, we conclude that the integral $\int_a^b f(x)dx$ exists, as desired. □

4. Properties of the Riemann Integral

Let $a, b, c \in \mathbb{R}$, $a < b$. Let $f, g\colon [a, b] \to \mathbb{R}$ be integrable

(1) $\int_a^b f(x)dx = -\int_b^a f(x)dx$.
(2) $\int_a^a f(x)dx = 0$.
(3) $\int_a^b c\,dx = c(b-a)$.
(4) $\int_a^b [f(x)+g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx$.
(5) $\int_a^b cf(x)dx = c\int_a^b f(x)dx$.
(6) $\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$.
(7) If $f \geq 0$, then $\int_a^b f(x)dx \geq 0$.
(8) If $f \geq g$, then $\int_a^b f(x)dx \geq \int_a^b g(x)dx$.
(9) If $m \leq f \leq M$, then $m(b-a) \leq \int_a^b f(x)dx \leq M(b-a)$.
(10) $|\int_a^b f(x)dx| \leq \int_a^b |f(x)|\,dx$.
(11) $\int_0^1 f(x)dx = \lim_{n\to\infty} \sum_{i=1}^n \frac{1}{n} f\left(\frac{i}{n}\right)$, if $a = 0, b = 1$.
(12) $\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^n \left(\frac{b-a}{n}\right) f\left(a + \frac{i(b-a)}{n}\right)$.

Remark 4.1. Property (11) can be used to evaluate certain infinite sums, as we show in Problem 9.2 below.

Proof of (3). Let $P = \{x_i\}_{i=0}^n$ be a partition of $[a, b]$. Let $y_i \in [x_i, x_{i-1}]$, $i = 1, \ldots, n$. Then any Riemann sum of the constant function $f(x) = c$ is of the form

$$\sum_{i=1}^n (x_i - x_{i-1}) f(y_i) = \sum_{i=1}^n (x_i - x_{i-1}) c = c \sum_{i=1}^n (x_i - x_{i-1}) = c(x_n - x_0) = c(b-a). \qquad (*)$$

Now, let $\max_{i=1,\ldots,n}(x_i - x_{i-1}) \to 0$. Applying Proposition 3.9, the left side of $(*)$ approaches $\int_a^b c\,dx$, and the right side is $c(b-a)$. ☐

Proof of (4). Let $\varepsilon, \delta > 0$, let $a < b < c$. Let $P_1 = \{x_i\}_{i=0}^n$ be a partition of $[a, b]$. Let $P_2 = \{y_i\}_{i=0}^m$ be a partition of $[b, c]$. Assume that $\max_{i=1,\ldots,n}(x_i - x_{i-1}) < \delta$ and $\max_{i=1,\ldots,n}(y_i - y_{i-1}) < \delta$. Define a partition $P_3 = \{z_i\}_{i=0}^{n+m+1}$ of $[a, c]$ so that $P_3 = \{x_0, x_1, \ldots, x_n, y_0, y_1, \ldots, y_m\}$. Then $\max_{i=1,\ldots,n}(z_i - z_{i-1}) < \delta$. Let $x_i' \in [x_i, x_{i-1}]$, $i = 1, \ldots, n$, and let $y_j' \in [y_j, y_{j-1}]$, $j = 1, \ldots, m$. Assume that $\delta = \delta(\varepsilon) > 0$ is so small such that

$$\left|\sum_{i=1}^n (x_i - x_{i-1}) f(x_i') - \int_a^b f(x)dx\right| < \varepsilon, \quad \text{and}$$

$$\left|\sum_{j=1}^m (y_j - y_{j-1}) f(y_j') - \int_b^c f(x)dx\right| < \varepsilon.$$

Note that

$$S = \sum_{i=1}^n (x_i - x_{i-1}) f(x_i') + \sum_{j=1}^m (y_j - y_{j-1}) f(y_j')$$

is a Riemann sum for $\int_a^c f(x)dx$ with $\max_{i=1,\ldots,n+m+1}(z_i - z_{i-1}) < \delta$. So, choosing $\delta(\varepsilon)$ smaller if necessary,

$$\left|\sum_{i=1}^{n}(x_i - x_{i-1})f(x_i') + \sum_{j=1}^{m}(y_j - y_{j-1})f(y_j') - \int_a^c f(x)dx\right| < \varepsilon.$$

Then, by the triangle inequality,

$$\begin{aligned}
&\left|\int_a^c f(x)dx - \int_a^b f(x)dx - \int_b^c f(x)dx\right| \\
&\quad= \left|\int_a^c f(x)dx - S + S - \int_a^b f(x)dx - \int_b^c f(x)dx\right| \\
&\quad\leq \left|\int_a^c f(x)dx - S\right| + \left|S - \int_a^b f(x)dx - \int_b^c f(x)dx\right| \\
&\quad\leq \varepsilon + \left|\sum_{i=1}^{n}(x_i - x_{i-1})f(x_i') - \int_a^b f(x)dx\right| + \left|\sum_{j=1}^{m}(y_j - y_{j-1})f(y_j') - \int_b^c f(x)dx\right| \\
&\quad< 3\varepsilon
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we conclude that (4) holds. □

Proof of (7). Let $P = \{x_i\}_{i=0}^n$ be a partition of $[a, b]$. Let $y_i \in [x_i, x_{i-1}]$, $i = 1, \ldots, n$. By assumption, $f(y_i) \geq 0$. And $x_i - x_{i-1} \geq 0$. So, any Riemann sum of the function $f(x)$ satisfies

$$\sum_{i=1}^{n}(x_i - x_{i-1})f(y_i) \geq 0. \qquad (*)$$

Now, let $\max_{i=1,\ldots,n}(x_i - x_{i-1}) \to 0$. Applying Proposition 3.9, the left side of $(*)$ approaches $\int_a^b f(x)dx$, and the right is 0. Since non-strict inequalities are preserved by limits, we conclude that $\int_a^b f(x)dx \geq 0$. □

Proof of (10). Let $P = \{x_i\}_{i=0}^n$ be a partition of $[a, b]$. Let $y_i \in [x_i, x_{i-1}]$, $i = 1, \ldots, n$. By the triangle inequality, any Riemann sum of the function $f(x)$ satisfies

$$\left|\sum_{i=1}^{n}(x_i - x_{i-1})f(y_i)\right| \leq \sum_{i=1}^{n}|(x_i - x_{i-1})f(y_i)| = \sum_{i=1}^{n}(x_i - x_{i-1})\,|f(y_i)| \qquad (*)$$

Now, let $\max_{i=1,\ldots,n}(x_i - x_{i-1}) \to 0$. Applying Proposition 3.9 and Corollary 3.11, the right side of $(*)$ approaches $\int_a^b |f(x)|\,dx$. Using continuity of the absolute value function, the left side approaches $|\int_a^b f(x)dx|$ Since non-strict inequalities are preserved by limits, we conclude that $|\int_a^b f(x)dx| \leq \int_a^b |f(x)|\,dx$. □

Proof of (11). Apply Proposition 3.9. □

Exercise 4.2. Sketch the proofs of Properties 1,2,5,6,8,9, and 12.

5. The Fundamental Theorem of Calculus

We can finally describe the precise manner in which differentiation and integration are opposite.

Theorem 5.1. ***(Fundamental Theorem of Calculus)***

(i) *Let $f\colon [a,b]\to\mathbb{R}$ be continuous. For $x\in(a,b)$ define $g(x)=\int_a^x f(t)dt$. Then g is an antiderivative of f, i.e.*

$$g'(x)=\frac{d}{dx}\int_a^x f(t)dt=f(x).$$

(ii) *Let $f\colon [a,b]\to\mathbb{R}$ be differentiable. Assume also that $f'\colon[a,b]\to\mathbb{R}$ is continuous. Then*

$$\int_a^b f'(x)dx=f(b)-f(a).$$

Remark 5.2. Part (ii) of Theorem 5.1 can be used to evaluate many different integrals. For example, we have the following two corollaries.

Corollary 5.3. ***(Integrating Polynomials)*** *Let $\alpha\in\mathbb{R}$, $\alpha\neq-1$, $0<a<b$. Then*

$$\int_a^b x^\alpha dx=\left[\frac{1}{\alpha+1}x^{\alpha+1}\right]_{x=a}^{x=b}=\frac{1}{\alpha+1}(b^{\alpha+1}-a^{\alpha+1}).$$

Remark 5.4. What happens if we allow $a=-1$, $b=1$, $\alpha<0$?

Corollary 5.5. ***(An Alternate Definition of the Natural Logarithm)*** *Let $x>0$. Then*

$$\int_1^x \frac{1}{t}\,dt=\log(x).$$

Proof of Theorem 5.1(i). We treat the difference quotient directly. Let $h\in\mathbb{R}$. Then

$$\frac{g(x+h)-g(x)}{h}=\frac{1}{h}\int_x^{x+h} f(t)dt.$$

We now apply the definition of continuity. For any $\varepsilon>0$, there exists $\delta(\varepsilon)>0$ such that, if y satisfies $|x-y|<\delta(\varepsilon)$, then $|f(x)-f(y)|<\varepsilon$. So, fix $\varepsilon>0$, and let $\delta(\varepsilon)>0$ as guaranteed by the definition of continuity. So, for any h with $0<|h|<\delta(\varepsilon)$, we have

$$f(x)-\varepsilon=\frac{1}{h}\int_x^{x+h}(f(x)-\varepsilon)dt\leq\frac{1}{h}\int_x^{x+h}f(t)dt\leq\frac{1}{h}\int_x^{x+h}(f(x)+\varepsilon)dt=f(x)+\varepsilon.$$

Letting $h\to0$ and using the Squeeze Theorem, we conclude that, for all $\varepsilon>0$,

$$f(x)-\varepsilon\leq\lim_{h\to0}\frac{1}{h}\int_x^{x+h}f(t)dt\leq f(x)+\varepsilon.$$

Then, letting $\varepsilon\to0$ and using the Squeeze Theorem again, we get $\lim_{h\to0}\frac{1}{h}\int_x^{x+h}f(t)dt=f(x)$, as desired. □

Proof of of Theorem 5.1(ii). Suppose we have a partition of $[a, b]$. That is, we have

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

Then

$$\begin{aligned} f(b) - f(a) &= f(x_n) - f(x_0) \\ &= f(x_n) + [-f(x_{n-1}) + f(x_{n-1})] + \cdots + [-f(x_1) + f(x_1)] - f(x_0) \\ &= [f(x_n) - f(x_{n-1})] + [f(x_{n-1}) - f(x_{n-2})] + \cdots + [f(x_2) - f(x_1)] + [f(x_1) - f(x_0)] \\ &= \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \end{aligned}$$

By the Mean Value Theorem, there exists $y_i \in [x_i, x_{i-1}]$ such that, for $i = 1, \ldots, n$,

$$(x_i - x_{i-1}) f'(y_i) = f(x_i) - f(x_{i-1}).$$

Therefore,

$$f(b) - f(a) = \sum_{i=1}^{n} (x_i - x_{i-1}) f'(y_i). \qquad (*)$$

The right side of $(*)$ is a Riemann sum for f'. Since f' is continuous, we know that f' is integrable. So, letting $\max_{i=1,\ldots,n}(x_i - x_{i-1}) \to 0$ and applying Proposition 3.9, $(*)$ becomes our desired equality

$$f(b) - f(a) = \int_a^b f'(t) dt.$$

□

Proof of Corollaries. Let $f(x) = (1/(\alpha + 1)) x^{\alpha+1}$. Note that $f'(x) = x^\alpha$, and then apply the Fundamental Theorem, Theorem 5.1(ii). This proves the first corollary. For the second corollary, recall that $(d/dt) \log(t) = 1/t$ for $t > 0$. So, the Fundamental Theorem, Theorem 5.1(ii), says that $\int_1^x (1/t) dt = \log(x) - \log(1) = \log(x)$. □

6. Integration by Substitution

Theorem 6.1. ***(Change of Variables)*** *Let $g\colon [a, b] \to [c, d]$ be differentiable. Also, suppose that $f\colon [c, d] \to \mathbb{R}$ is continuous. Then*

$$\int_a^b f(g(x)) g'(x) dx = \int_{g(a)}^{g(b)} f(x) dx.$$

Example 6.2. The following integral may appear difficult if not impossible to evaluate, but Theorem 6.1 allows us to evaluate it. For $x > 0$, let $f(t) = e^t$, and let $g(x) = 1/x$. Applying Theorem 6.1 and then Theorem 5.1(ii),

$$\int_1^2 \frac{e^{1/x}}{x^2} dx = -\int_1^2 f'(g(x)) g'(x) dx = -\int_1^{1/2} e^t dt = \int_{1/2}^1 e^t dt = \int_{1/2}^1 \frac{d}{dt} e^t dt = e - \sqrt{e}.$$

Proof of Theorem 6.1. For $a < x \leq b$, let $F(x) = \int_a^x f(t) dt$. From the Fundamental Theorem of Calculus, Theorem 5.1(i), $F'(x) = f(x)$. Also, by the Chain Rule, $(d/dx)[F(g(x))] =$

$F'(g(x))g'(x) = f(g(x))g'(x)$. Note also that $f(g(x))g'(x)$ is integrable by Corollary 3.11. Putting everything together, we have

$$\begin{aligned}\int_a^b f(g(x))g'(x)dx &= \int_a^b \frac{d}{dx}[F(g(x))]dx = F(g(b)) - F(g(a)), \qquad \text{by Theorem 5.1(ii)}\\ &= \int_{g(b)}^{g(a)} \frac{d}{dx}F(x)dx, \qquad \text{by Theorem 5.1(ii)}\\ &= \int_{g(b)}^{g(a)} f(x)dx\end{aligned}$$

□

7. Integrals Involving Infinity

From Definition 3.5 and Definition 3.2, we can not yet integrate functions that are unbounded, and we can not yet integrate functions with unbounded domains. If the function is unbounded, then some of its Riemann sums could have an infinite value, so the integral technically does not exist. And a function technically cannot be integrated over an unbounded domain, since we only allowed finite partitions in our definition of the Riemann sum. However, there are easy ways to fix these two issues, as we now show.

Definition 7.1. Let $f\colon (a,b] \to \mathbb{R}$ be continuous. Then, for any $\varepsilon > 0$, $f\colon [a+\varepsilon, b] \to \mathbb{R}$ is continuous, and the following integral exists by Theorem 3.10:

$$\int_{a+\varepsilon}^b f(x)dx.$$

If the limit $\lim_{\varepsilon\to 0^+} \int_{a+\varepsilon}^b f(x)dx$ exists, we define

$$\int_a^b f(x)dx = \lim_{\varepsilon\to 0^+} \int_{a+\varepsilon}^b f(x)dx.$$

Definition 7.2. Let $f\colon [a,\infty) \to \mathbb{R}$ be continuous, and let $g\colon (-\infty,\infty)$ be continuous. Then, for any $N > 0$, $f\colon [a,N] \to \mathbb{R}$ is continuous, $g\colon [-N,N] \to \mathbb{R}$ is continuous and the following integrals exists by Theorem 3.10:

$$\int_a^N f(x)dx, \quad \int_{-N}^N g(x)dx$$

If the limit $\lim_{N\to\infty} \int_a^N f(x)dx$ exists, we define

$$\int_a^\infty f(x)dx = \lim_{N\to\infty} \int_a^N f(x)dx.$$

If the limit $\lim_{N\to\infty} \int_{-N}^N g(x)dx$ exists, we define

$$\int_{-\infty}^\infty g(x)dx = \lim_{N\to\infty} \int_{-N}^N g(x)dx.$$

Example 7.3. The function $f(x) = x^{-1/3}$ is unbounded near zero, but its integral on the interval $[0, 1]$ is finite:

$$\int_0^1 x^{-1/3}dx = \lim_{\varepsilon \to 0^+} \int_\varepsilon^1 x^{-1/3}dx = \lim_{\varepsilon \to 0^+} [(3/2)x^{2/3}]_{x=\varepsilon}^{x=1} = \lim_{\varepsilon \to 0^+} [(3/2) - (3/2)\varepsilon^{2/3}] = 3/2.$$

Example 7.4. The function $f(x) = 1/x$ is unbounded near zero, and its integral on the interval $[0, 1]$ is infinite. Using Corollary 5.5,

$$\int_0^1 x^{-1}dx = \lim_{\varepsilon \to 0^+} \int_\varepsilon^1 x^{-1}dx = \lim_{\varepsilon \to 0^+} -\log(\varepsilon) = \lim_{\varepsilon \to 0^+} \log(\varepsilon^{-1}) = \lim_{y \to \infty} \log(y) = \infty.$$

8. Selected Exercises from the Textbook

Exercise 8.1. Let $f: \mathbb{R} \to \mathbb{R}$ be continuous with two continuous derivatives. Find all such f such that $f''(x) = 20x^3 - 12x^2 + 6x$.

Exercise 8.2. For a continuous function f, we know from Section 4, Property (9), that

$$(b - a) \min_{y \in [a,b]} f(y) \le \int_a^b f(x)dx \le (b - a) \max_{y \in [a,b]} f(y).$$

Use this property to estimate $\int_0^2 (x^3 - 3x + 3)dx$.

Exercise 8.3. Let $f, g: \mathbb{R} \to \mathbb{R}$ be integrable functions. Suppose $\int_0^9 f(x)dx = 5$ and $\int_0^9 g(x)dx = 7$. Find $\int_0^9 (3f(x) + 2g(x))dx$.

Exercise 8.4. Evaluate $\int_{-2}^3 (x^2 - 3)dx$.

Exercise 8.5. A manufacturing company owns a major piece of equipment that depreciates at the (continuous) rate $f = f(t) \ge 0$, where t is the time measured in months since its last overhaul. Because a fixed cost $A > 0$ is incurred each time the machine is overhauled, the company wants to determine the optimal time T (in months) between overhauls.

(a) Explain why $\int_0^t f(s)ds$ represents the loss in value of the machine over the period of time t since the last overhaul.

(b) Let $C = C(t)$ be given by

$$C(t) = \frac{1}{t}\left[A + \int_0^t f(s)ds\right].$$

What does C represent and why would the company want to minimize C?

(c) Assume that $\lim_{s \to \infty} f(s) = \infty$. Show that C has a minimum value at one of the numbers $t = T$ where $C(T) = f(T)$.

Exercise 8.6. A high-tech company purchases a new computing system whose initial value is V. The system will depreciate at the rate $f = f(t)$ and will accumulate maintenance costs at the rate $g = g(t)$, where t is the time measure in months. The company wants to determine the optimal time to replace the system.

(a) Let

$$C(t) = \frac{1}{t}\int_0^t [f(s) + g(s)]ds.$$

Show that the critical numbers of C occur at the numbers t where $C(t) = f(t) + g(t)$.

(b) Suppose

$$f(t) = \begin{cases} \frac{V}{15} - \frac{V}{450}t & \text{, if } 0 < t \leq 30 \\ 0 & \text{, if } t > 30 \end{cases},$$

and suppose $g(t) = \frac{Vt^2}{12900}$ for $t > 0$. Determine the length of time T for the total depreciation $D(t) = \int_0^t f(s)ds$ to equal the initial value V.

(c) Determine the absolute minimum of C on $(0, T]$.

(d) Sketch the graphs of C and $f + g$ in the same coordinate system, and verify the result of part (a) in this case.

Exercise 8.7. Find the average value of the function $h(x) = (\cos(x))^4 \sin(x)$ on the interval $[0, \pi]$.

Exercise 8.8. Let $a > 0$. Evaluate $\int_0^a x\sqrt{a^2 - x^2}dx$.

Exercise 8.9. State whether or not the statement is True or False. Justify your answer. Let $a < b$.

(1) If $f, g\colon [a, b] \to \mathbb{R}$ are continuous, then

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

(2) If $f, g\colon [a, b] \to \mathbb{R}$ are continuous, then

$$\int_a^b (f(x)g(x))dx = \left(\int_a^b f(x)dx\right)\left(\int_a^b g(x)dx\right).$$

(3) If $f, g\colon [a, b] \to \mathbb{R}$ are continuous, and if $f(x) \geq g(x)$ for all $x \in [a, b]$, then

$$\int_a^b f(x)dx \geq \int_a^b g(x)dx.$$

(4) If $f, g\colon [a, b] \to \mathbb{R}$ are continuous, and if $f(x) > g(x)$ for all $x \in [a, b]$, then

$$\int_a^b f(x)dx > \int_a^b g(x)dx.$$

(5) All continuous functions have antiderivatives.

(6) If $f\colon (a, b) \to \mathbb{R}$ is continuous, then $\int_a^b f(x)dx$ exists.

9. Selected Problems

Problem 9.1. One of the integrals below has a finite value. The other does not. Evaluate the one with a finite value, and explain why the other does not have a finite value.

$$(a) \quad \int_0^2 \frac{1}{\sqrt{x}}dx, \qquad\qquad (b) \quad \int_{-1}^3 \frac{1}{x^2}dx$$

Problem 9.2. Evaluate

$$\lim_{n\to\infty} \frac{1^5 + 2^5 + \cdots + n^5}{n^6},$$

by showing that the limit is $\int_0^1 x^5dx$.

Problem 9.3. MIT 18.01SC Final, Q 7(b) What is $\lim_{h\to 0} \frac{1}{h}\int_2^{2+h} \sin(x^2)dx$

Problem 9.4. What is $\lim_{h\to 0}\int_2^{2+h}\sin(x^2)dx$?

Problem 9.5. MIT 18.01SC Exam3, Q1 Compute the area between the curves $x = y^2 - 4y$ and $x = 2y - y^2$.

Problem 9.6. MIT 18.01SC Exam3, Q3(b) Suppose $f\colon \mathbb{R}\to\mathbb{R}$ is continuous and

$$x\sin(\pi x) = \int_0^{x^2} f(t)dt$$

Find $f(4)$.

Problem 9.7. (Challenge question) Come up with a bounded function $f\colon [0,1]\to[0,1]$ such that the integral $\int_0^1 f(x)dx$ does not exist. Hint: try using the function you found in the notes from the first few weeks of class.

Problem 9.8. (Challenge question) The following formula comes from Chapter 6, but it is so useful that it should be mentioned. This formula allows us to move around a derivative inside an integral. Let $f,g\colon [a,b]\to\mathbb{R}$ be differentiable functions. Use the Fundamental Theorem of Calculus and the product rule to derive the **integration by parts formula**:

$$\begin{aligned}\int_a^b f'(x)g(x)dx &= [f(x)g(x)]_{x=a}^{x=b} - \int_a^b f(x)g'(x)dx\\ &= [f(b)g(b) - f(a)g(a)] - \int_a^b f(x)g'(x)dx.\end{aligned}$$

COURANT INSTITUTE, NEW YORK UNIVERSITY, NEW YORK NY 10012
E-mail address: heilman@cims.nyu.edu