Summer 2017

The Definite Integrals of Cauchy and Riemann

Dave Ruch
ruch@msudenver.edu

The Definite Integrals of Cauchy and Riemann

David Ruch*

December 18, 2017

1 Introduction

Rigorous attempts to define the definite integral began in earnest in the early 1800's. A major motivation at the time was the search for functions that could be expressed as Fourier series as follows:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos(kw) + b_k \sin(kw)\right) \quad \text{where the coefficients are:} \tag{1}$$

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt, \quad a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(kt)\,dt \quad , \quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(kt)\,dt.$$

Joseph Fourier (1768-1830) argued in 1807 that this series expansion was valid for *any* function f, and he used the expansion in his study of heat conduction. This ambitious claim was met with considerable skepticism among mathematicians, but it certainly motivated much research into the convergence of these infinite series.

One of the pioneers in this development was A. L. Cauchy (1789-1857). He made a study of the definite integral for continuous functions in his 1823 *Calcul Infinitésimal* [C], which we will read from in Section 2 of this project. Both Cauchy and Fourier attempted to prove that the Fourier series would converge to $f(x)$ under suitable conditions. Unfortunately, both proof attempts had flaws. J. Dirichlet (1805-1859) read their work, and in an 1829 paper [D] he set out to give a rigorous proof after pointing out an error in Cauchy's proof.

Dirichlet gave a proof of Fourier series convergence in his 1829 paper that is valid for a piecewise continuous function f with finitely many jump discontinuities[1] and a finite number of extrema. He then discussed the possibility of extending his proof for a function f with infinite extrema (in a bounded interval), but he didn't hold much hope for functions with infinite discontinuities. To indicate why, he gave an example that quickly became famous in mathematical circles of his day. The next passage is from Dirichlet's discussion of the Fourier series for a function with discontinuities.

*Department of Mathematical and Computer Sciences, Metropolitan State University of Denver, Denver, CO; ruch@msudenver.edu

[1] At each point of discontinuity, the one-sided limits exist and are finite.

If the points of discontinuity are infinite in number, the integral ... makes sense only when the function is given in such a way that, for any two values a and b where $-\pi < a < b < \pi$, we can find two values r and s, with $a < r < s < b$, such that the function is continuous in the interval from r to s. One readily feels the necessity of this restriction on considering that the various terms of the series [(1)] are definite integrals and on returning to the fundamental concept of an integral. One then sees that the integral of a function means something only when the function satisfies the condition set out above. One would have an example of a function which does not fulfil this condition, if one assumes $\phi(x)$ equal to a specific constant c when the variable x acquires a rational value, and equal to another constant d, when this variable is irrational. The function so defined has finite and determinate values for every value of x, and yet one does not know how to substitute it the series [(1)], seeing that the various integrals that enter into this series will lose all meaning in this case.

For the rest of project, we'll refer to this example function as "Dirichlet's function ϕ".

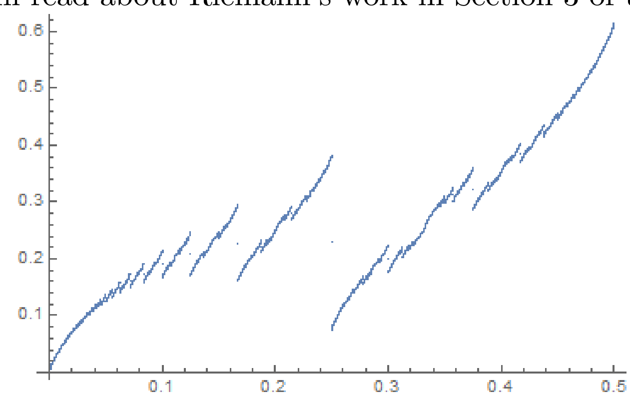
Exercise 1 *Consider the example function $\phi(x)$ Dirichlet gives in the excerpt. Dirichlet claims this function does not satisfy the condition:*

Dirichlet Condition. *For any two values a and b where $-\pi < a < b < \pi$, we can find two values r and s, with $a < r < s < b$, such that the function is continuous in the interval from r to s.*

First show that Dirichlet's function ϕ is not continuous at any rational x. Then prove it is not continuous at any irrational x. Finally, use these results to verify Dirichlet's claim that ϕ does not satisfy the Dirichlet Condition.

It is important to remember that in 1829 the only definition of the definite integral was the one given by Cauchy, and that definition was only for continuous functions. Thus we can see why Dirichlet felt "One readily feels the necessity of ... returning to the fundamental concept of an integral."

While the study of Fourier series raged on for the next couple decades, it wasn't until 1854 that Bernard Riemann developed a more general concept of the definite integral that could be applied to functions with infinite discontinuities. Amazingly, he also constructed an integrable function with infinite discontinuities that does not satisfy Dirichlet's Condition above - see the graph below. We will read about Riemann's work in Section 3 of this project.



2 Cauchy's Definite Integral

Most mathematicians before Cauchy's time preferred to think of integration as the inverse of differentiation: to evaluate $\int_a^b f(x)\,dx$ you found an antiderivative F of f and evaluated $F(b) - F(a)$. However, there was plenty of 18th century mathematics evaluating difficult integrals approximately using sums. Cauchy used many of their ideas in creating his new definition of the definite integral.

Cauchy was a professor at the École Polytechnique in Paris during the 1820's when he wrote two texts on the calculus. He developed his theory of the definite integral for continuous functions in his 1823 *Calcul Infinitésimal* [C]. We will read his development over the course of several excerpts in Section 2 of this project.

Excerpt A from Cauchy's Calcul Infinitésimal

Definite Integrals.

Suppose that, the function $y = f(x)$, being continuous with respect to the variable x between two finite limits $x = x_0$, $x = X$, we denote by $x_1, x_2, \ldots, x_{n-1}$ new values of x interposed between these limits, which always go on on increasing or decreasing from the first limit up to the second. We can use these values to divide the difference $X - x_0$ into elements

$$x_1 - x_0, \quad x_2 - x_1, \quad x_3 - x_2, \quad \ldots, \quad X - x_{n-1}, \tag{2}$$

which will always be the same sign. This granted, consider that we multiply each element by the value of $f(x)$ corresponding to the origin of this same element, namely, the element $x_1 - x_0$ by $f(x_0)$, the element $x_2 - x_1$ by $f(x_1)$, $\ldots$, finally, the element $X - x_{n-1}$ by $f(x_{n-1})$; and, let

...

$$S = (x_1 - x_0) f(x_0) + (x_2 - x_1) f(x_1) + \cdots + (X - x_{n-1}) f(x_{n-1}) \tag{3}$$

be the sum of the products thus obtained. The quantity S will obviously depend upon: 1° the number of

elements n into which we will have divided the difference $X - x_0$; 2° the values of these same elements, and by consequence, on the mode of division adapted. Now, it is important to remark that, if the numerical values of the elements become very small and the number n very considerable, the mode of division will no longer have a perceptible influence on the value of S.

Exercise 2 *Consider the example* $f(x) = x^2 - 2$, $x_0 = 0$, $x_1 = 1/2$, $x_2 = 3/2$, $X = 2$, $n = 3$.

(a) Find the elements $x_1 - x_0, x_2 - x_1, x_3 - x_2$ *for this example. Then calculate the sum* S. *How close is* S *to* $\int_{x_0}^{X} f(x)\,dx$?

(b) Make and label a diagram that graphically represents what is going on with Cauchy's construction of S *in (3) for this example. Does the general* S *formula remind you of something you've seen in your Introductory Calculus courses?*

We will find it convenient to give a modern name to the set of values $\mathcal{P} = \{x_0, x_1, x_2, \ldots, x_{n-1}, X\}$. We will call $\mathcal{P}$ a **partition** of the interval $[a, b]$ and require the x_k values to be distinct. When Cauchy refers to the "mode of division", this is equivalent to choosing a partition for the interval. Also, rather than continuing to use the letter S for different things, a handy modern notation is to include the partition in the notation. We will use the modern notation $S(f, \mathcal{P})$ for Cauchy's sum S to indicate the dependence of S on f and $\mathcal{P}$.

Exercise 3 *Observe that Cauchy makes a bold claim at the very end of the excerpt that we will call Claim M for "**m**ode of division"*

> ***Claim M.*** *"the mode of division will no longer have a perceptible influence on the value of S."*

What two requirements does Cauchy place on this claim?

Exercise 4 *Write Cauchy's Claim M with modern terminology and quantifiers.*

You may have noticed in the last exercise that the maximum element value will be important, and so we will give it a modern name. Define mesh $(\mathcal{P})$, the **mesh** of a partition $\mathcal{P}$, to be its maximum element value. For example, mesh $(\mathcal{P}) = 1$ for the partition $\mathcal{P}$ in Exercise 2.

In order to prove his claim, Cauchy takes up the idea of partitioning each subinterval (x_{k-1}, x_k) and considering the corresponding sum $S(f, \mathcal{P}')$ for the new partition $\mathcal{P}'$ of $[x_0, X]$. From inside the first subinterval $[x_0, x_1]$ he chooses m points $\left\{x_j^1\right\}_{j=1}^m$ with

$$x_0 < x_1^1 < x_2^1 < \cdots < x_m^1 < x_1$$

and considers the sum

$$S_1 = \left(x_1^1 - x_0\right) f(x_0) + \left(x_2^1 - x_1^1\right) f\left(x_1^1\right) + \left(x_3^1 - x_2^1\right) f\left(x_2^1\right) + \cdots + \left(x_1 - x_m^1\right) f\left(x_m^1\right). \tag{4}$$

Cauchy uses some very clever algebra and the Intermediate value Theorem (IVT) with the continuity of f to show that

$$S_1 = f(c_1)(x_1 - x_0) \tag{5}$$

for some c_1 between x_0 and x_1. He carries out this process for each subinterval and then adds up the sums to show that

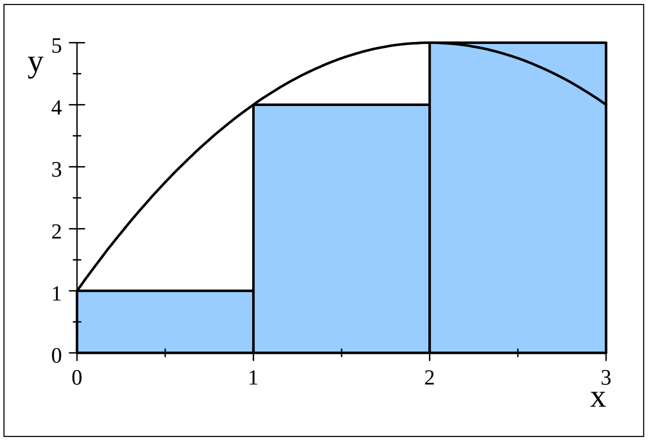
$$S(f, \mathcal{P}') = f(c_1)(x_1 - x_0) + f(c_2)(x_2 - x_1) + \cdots + f(c_n)(X - x_{n-1})$$

for $c_k \in (x_{k-1}, x_k)$ chosen according to the IVT.

Exercise 5 *Consider a function $f(x)$ on the interval $[0, 7]$, where part of the graph is given below. Let $x_0 = 0, x_1 = 3$, $m = 2$ where we partition the first subinterval $[x_0, x_1]$ as shown in the diagram.*

(a) Use the figure and a rectangle area argument to estimate the value of c_1 for (5) with this example. Explain from the graph why we can be sure such a c_1 exists, even without knowing a formula for $f(x)$.

(b) Now assume $f(x) = 5 - (x - 2)^2$. Calculate S_1 from the formula (4). Then find c_1 to one decimal place using algebra. Label c_1 on the diagram and compare with your estimate of c_1 in part (a).



Next we will read Cauchy's description of this process of partitioning each subinterval and deriving a new formula for the sum S.

Excerpt B from Cauchy's Calcul Infinitésimal

To pass from the mode of division that we have just considered, to another in which the numerical values of the elements of $X - x_0$ are even smaller, it will suffice to partition each of the expressions in (2) into new elements. Then, we should replace, in the second member of equation (3), the product $(x_1 - x_0) f(x_0)$ by a sum of similar products, for which we can substitute an expression of the form

$$(x_1 - x_0) f[x_0 + \theta_0 (x_1 - x_0)], \tag{6}$$

θ_0 being a number less than unity. ...

By the same reasoning, we should substitute for the product $(x_2 - x_1) f(x_1)$, a sum of terms which can be presented under the form

$$(x_2 - x_1) f[x_1 + \theta_1 (x_2 - x_1)],$$

θ_1 again denoting a number less than unity.

By continuing in this manner, we will finally conclude that, in the new mode of division, the value of S will be of the form

$$\begin{aligned} S \;=\; & (x_1 - x_0) f[x_0 + \theta_0 (x_1 - x_0)] \\ & + (x_2 - x_1) f[x_1 + \theta_1 (x_2 - x_1)] + \cdots \\ & + (X - x_{n-1}) f[x_{n-1} + \theta_{n-1} (X - x_{n-1})]. \end{aligned} \tag{7}$$

In modern terminology, we define a **refinement of partition** to describe what Cauchy calls the new mode of division in which we partition each of the expressions in (2) into new elements. If we let $\mathcal{P} = \{x_0, x_1, x_2, \ldots, x_{n-1}, X\}$ be the original partition and let $\mathcal{P}'$ be the refinement, $\mathcal{P}'$ will include X and all the x_k plus some additional values between x_0 and X. For example, a refinement of the partition $\mathcal{P} = \{0, 1/2, 3/2, 2\}$ in Exercise 2 is $\mathcal{P}' = \{0, 1/3, 1/2, 7/8, 1, 3/2, 2\}$.

If we let $\mathcal{P} = \{x_0, x_1, x_2, \ldots, x_{n-1}, X\}$ be Cauchy's original partition and let $\mathcal{P}'$ be a refinement, then in modern terminology the sum in (3) is $S(f, \mathcal{P})$ and the sum in (7) is $S(f, \mathcal{P}')$.

2.1 Comparing $S(f,\mathcal{P})$ and $S(f,\mathcal{P}')$ for refinement $\mathcal{P}'$.

Let's reflect briefly on what Cauchy cleverly created with his expression (7) for the sum $S(f,\mathcal{P}')$ with *refined* partition $\mathcal{P}'$. He now has

$$\begin{aligned} S(f,\mathcal{P}) &= (x_1-x_0)\,f(x_0)+(x_2-x_1)\,f(x_1)+\cdots+(X-x_{n-1})\,f(x_{n-1}) \qquad \text{and} \\ S(f,\mathcal{P}') &= (x_1-x_0)f[x_0+\theta_0(x_1-x_0)]+\cdots+(X-x_{n-1})f[x_{n-1}+\theta_{n-1}(X-x_{n-1})] \end{aligned}$$

which are both expressions in terms of the original partition $\mathcal{P}$ values $x_0, x_1, \ldots, x_{n-1}, X$. Then he can work more easily with the difference $S(f,\mathcal{P})-S(f,\mathcal{P}')$, which is allegedly tiny, in his proof of Claim M. Let's see how he does it.

Excerpt C from Cauchy's Calcul Infinitésimal

If in this last equation [(7)] we let

$$\begin{aligned} f[x_0+\theta_0(x_1-x_0)] &= f(x_0)\pm\epsilon_0, \\ f[x_1+\theta_1(x_2-x_1)] &= f(x_1)\pm\epsilon_1, \\ &\cdots\cdots\cdots \\ f[x_{n-1}+\theta_{n-1}(X-x_{n-1})] &= f(x_{n-1})\pm\epsilon_{n-1} \end{aligned} \tag{8}$$

we will derive

$$\begin{aligned} S = {} & (x_1-x_0)[f(x_0)\pm\epsilon_0] \\ & +(x_2-x_1)[f(x_1)\pm\epsilon_1]+\cdots \\ & +(X-x_{n-1})[f(x_{n-1})\pm\epsilon_{n-1}]\ ; \end{aligned} \tag{9}$$

then, by developing products,

$$\begin{aligned} S = {} & (x_1-x_0)\,f(x_0)+(x_2-x_1)\,f(x_1)+\cdots+(X-x_{n-1})\,f(x_{n-1}) \\ & \pm\epsilon_0(x_1-x_0)\pm\epsilon_1(x_2-x_1)\pm\cdots\pm\epsilon_{n-1}(X-x_{n-1})\,. \end{aligned} \tag{10}$$

Add that, if the elements $x_1-x_0, x_2-x_1, , \ldots, X-x_{n-1}$ have very small numerical values, each of the quantities $\pm\epsilon_0, \pm\epsilon_1, \ldots, \pm\epsilon_{n-1}$ will differ very little from zero; and as a result, it will be the same for the sum

$$\pm\epsilon_0(x_1-x_0)\pm\epsilon_1(x_2-x_1)\pm\cdots\pm\epsilon_{n-1}(X-x_{n-1})\,,$$

which is equivalent to the product of $X-x_0$ by an average between these various quantities. This granted, it follows from equations (3) and (10), when compared to each other, that we will not significantly alter the calculated value of S for a mode of division in which the elements of the difference $X-x_0$ have very small numerical values, if we pass to a second mode in which each of these elements are found subdivided into several others.

Notice that Cauchy is not yet comparing the sums $S(f,\mathcal{P}), S(f,\mathcal{Q})$ for two *arbitrary* partitions $\mathcal{P}$,$\mathcal{Q}$ with small mesh. For now he is working only with refinements. Let's rewrite what Cauchy actually proved in modern terminology with quantifiers as a lemma.

Lemma 6 *Suppose f is continuous on $[a,b]$. For any $\epsilon > 0$, we can find $d > 0$ such that if* $\operatorname{mesh}(\mathcal{P}) < d$ *and $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then $|S(f,\mathcal{P}) - S(f,\mathcal{P}')| < \epsilon$.*

Exercise 7 *A key to Cauchy's proof is his claim that "each of the quantities $\pm\epsilon_0, \pm\epsilon_1, \ldots, \pm\epsilon_{n-1}$ will differ very little from zero". What property of f allows him to say this?*

Exercise 8 *Use Cauchy's ideas to give a modern proof of Lemma 6.*

2.2 Comparing $S(f,\mathcal{P}_1)$ and $S(f,\mathcal{P}_2)$, and defining the definite integral

Cauchy is now ready to consider two "modes of division of the difference $X - x_0$, in each of which the elements of this difference have very small numerical values." That is, he wants to compare the sums $S(f,\mathcal{P}_1), S(f,\mathcal{P}_2)$ for two arbitrary partitions $\mathcal{P}_1$,$\mathcal{P}_2$ with small mesh.

Excerpt D from Cauchy's Calcul Infinitésimal

We can compare these two modes to a third, chosen so that each element, whether from the first or second mode, is found formed by the union of the various elements of the third. For this condition to be fulfilled, it will suffice that all the values of x interposed in the first two modes between the limits x_0, X are employed in the third, and we will prove that we alter the value of S very little by passing from the first or from the second mode to the third, and by consequence, in passing from the first to the second. Therefore, when the elements of the difference $X - x_0$ become infinitely small, the mode of division will no longer have a perceptible influence on the value of S; and, if we decrease indefinitely the numerical values of these elements, by increasing their number, the value of S will eventually be substantially constant, or in other words, it will finally attain a certain limit which will depend uniquely on the form of the function $f(x)$, and the extreme values x_0, X attributed to the variable x. This limit is what we call a definite integral.

Exercise 9 *Explain what Cauchy means by " it will suffice that all the values of x interposed in the first two modes between the limits x_0, X are employed in the third". Illustrate for the example where* $\mathcal{P}_1 = \{1, 2, 3.5, 5\}$ *and* $\mathcal{P}_2 = \{1, 1.7, 2.9, 4.7, 4.8, 5\}$.

Exercise 10 *Suppose we are given a continuous function g on $[a,b]$ and $\epsilon = 0.1$. Further, suppose we find the value d from Lemma 6 for $\epsilon/2$, and two partitions $\mathcal{P}_1$,$\mathcal{P}_2$ each with mesh less than d. Use Cauchy's reasoning and Lemma 6 to prove that*

$$|S(g,\mathcal{P}_1) - S(g,\mathcal{P}_2)| \leq 0.1$$

Now we just need to generalize the previous exercise to finally give a modern equivalent to Cauchy's Claim M, that "when the elements of the difference $X - x_0$ become infinitely small, the mode of division will no longer have a perceptible influence on the value of S".

Exercise 11 *State and prove a modern version of Claim M that generalizes Exercise 10.*

After convincing us of Claim M, Cauchy then goes on to define the definite integral $\int_a^b f$ as a limit, but he is not terribly precise about this limit. His basic idea is to choose any sequence of sums $S(f, \mathcal{P}_n)$ with $\lim_{n\to\infty} \text{mesh}(\mathcal{P}_n) = 0$. Then your theorem from Exercise 11 can be used to show the sequence $\{S(f, \mathcal{P}_n)\}$ is a Cauchy sequence in $\mathbb{R}$ and therefore has a limit, which we define to be the definite integral $\int_a^b f$. The formal details of this discussion can be explored in the Supplementary Exercises, Section 3.1.

Many of Cauchy's ideas will work for finding integrals of functions with discontinuities, but he uses continuity in a couple crucial spots.

Exercise 12 *Reflect on Cauchy's development of the definite integral for continuous functions. Where did he use continuity? Which ideas would make sense even for functions with discontinuities?*

To illustrate the problems with integrating functions with lots of discontinuities, we now look at Dirichlet's function ϕ and the theorem you proved in Exercise 11.

Theorem M. Suppose g is continuous on $[a, b]$. For any $\epsilon > 0$, we can find $d > 0$ such that if $\mathcal{P}_1$,$\mathcal{P}_2$ are partitions with $\text{mesh}(\mathcal{P}_1), \text{mesh}(\mathcal{P}_2) < d$, then $|S(f, \mathcal{P}_1) - S(f, \mathcal{P}_2)| < \epsilon$.

Exercise 13 *Prove that your theorem from Exercise 11 is not true for Dirichlet's function ϕ.*

While we won't prove it here, the condition in Theorem M turns out to be necessary and sufficient for a function f to be integrable. We will see similar ideas developed - with some twists - by Riemann in the next section.

3 Riemann's Definite Integral

Cauchy's 1823 development of the definite integral for continuous functions was not extended to non-continuous functions for another three decades. While Dirichlet and others continued to research the problem of Fourier series convergence, no one looked hard at the definite integral itself until 1854, when Dirichlet's student Bernard Riemann took up the issue.

Riemann (1826-1866) was born near Hanover, Germany and studied mathematics at the University of Göttingen and Berlin University with strong influence by C. Gauss and Dirichlet. Despite his early death from tuberculosis, Riemann made major contributions in geometry, number theory, and complex analysis, in addition to his work with Fourier series and the definite integral that bears his name.

Remember from the project introduction that Dirichlet was hoping to extend his Fourier series convergence proof to the case where there are infinitely many but isolated discontinuities and infinitely many extrema. This clearly motivated his student Riemann to develop and use a more general definition of the definite integral, as we shall now see.

All excerpts in this section are from Riemann's 1854 paper [R].

Riemann Excerpt A

Vagueness still prevails in some fundamental points concerning the definite integral. Hence I provide some preliminaries about the concept of a definite integral and the scope of its validity.

Hence first: What is one to understand by $\int_a^b f(x)\,dx$?

In order to establish this, we take a succession of values $x_1, x_2, \ldots, x_{n-1}$ between a and b arranged in succession, and denote, for brevity, $x_1 - a$ by δ_1, $x_2 - x_1$ by $\delta_2, \ldots,$ $b - x_{n-1}$ by δ_n, and a positive number less than 1 by ϵ. Then the value of the sum

$$S = \delta_1 f(a + \epsilon_1\delta_1) + \delta_2 f(x_1 + \epsilon_2\delta_2) + \delta_3 f(x_2 + \epsilon_3\delta_3) + \cdots + \delta_n f(x_{n-1} + \epsilon_n\delta_n) \tag{11}$$

depends on the selection of the intervals δ and the numbers ϵ. If this now has the property, that however the δ's and ϵ's are selected, S approaches a fixed limit A when the δ's become infinitely small together, this limiting value is called $\int_a^b f(x)\,dx$.

If we do not have this property, then $\int_a^b f(x)\,dx$ is undefined. ... if the function $f(x)$ becomes infinitely large ... then clearly the sum S, no matter what degree of smallness one may prescribe for δ, can reach an arbitrarily given value. Thus it has no limiting value, and by the above $\int_a^b f(x)\,dx$ would have no meaning.

Observe that Riemann frequently writes ϵ or δ where he clearly means a set of ϵ_k or δ_k values.

From hereon, we will say that if $\int_a^b f(x)\,dx$ exists according to Riemann's definition in Excerpt A, then f is **Riemann integrable** on $[a, b]$, and we will write $\int_a^b f$ for the definite integral.

Exercise 14 *Consider the example with* $f(x) = 2x^3 - 9x^2 + 12x + 1,\ a = 0, b = 3,$ *partition* $\mathcal{P} = \{0, 1, 3\}$, *and* $\epsilon_1 = 1/2,\ \epsilon_2 = 3/4$. *Make and label a diagram that graphically represents what is going on with Riemann's construction of S.*

Exercise 15 *Riemann has read Cauchy's work on the definite integral. Compare and contrast Riemann's definition of the sum S in (11) with Cauchy's definition of sum S in (3) and Cauchy's reworked formulation of S in (7).*

We've seen that in order to calculate the sum S for Riemann, we need to keep track of the ϵ_k values as well as the partition values x_k. For ease of notation, we will name the $x_{k-1} + \epsilon_k \delta_k$ values **tags** $t_k = x_{k-1} + \epsilon_k \delta_k$ and call the combined set of x_k and t_k values a **tagged partition**, writing $\dot{\mathcal{P}} = \{x_k, t_k\}_{k=1}^n$ for the tagged partition (with $x_0 = a$, $x_n = b$). Then we can write $S\left(f, \dot{\mathcal{P}}\right)$ for the sum S in (11) and call $S\left(f, \dot{\mathcal{P}}\right)$ a **Riemann sum**.

Exercise 16 *What are the tags for the example in Exercise 14?*

Exercise 17 *Give a general inequality that relates the tags t_k and partition values x_k in Riemann's definition of $\int_a^b f$.*

Exercise 18 *Using appropriate quantifiers and modern notation for tagged partitions and mesh, rewrite Riemann's definition in Excerpt A for the existence of $\int_a^b f$.*

After his definition of $\int_a^b f$, Riemann discusses the case where "the function $f(x)$ becomes infinitely large". You will use his ideas in the next exercise to give a modern proof of the following theorem:

Theorem B. If $f(x)$ is not bounded on $[a, b]$ then f is not Riemann integrable on $[a, b]$.

Exercise 19 *Assume, for the sake of contradiction, that f is unbounded but integrable with $A = \int_a^b f$. Since f is integrable, using $\epsilon = 1$ we can find $\delta > 0$ such that for any tagged partition $\dot{\mathcal{P}}$ of $[a, b]$ with* mesh$\left(\dot{\mathcal{P}}\right) < \delta$ *we have*

$$\left| S\left(f; \dot{\mathcal{P}}\right) - A \right| < 1. \tag{12}$$

(a) Let $\mathcal{P}$ be a partition $\{x_k\}_{k=1}^n$ of $[a, b]$ with mesh$(\mathcal{P}) < \delta$*. Explain why f must be unbounded on at least one subinterval of $[a, b]$, say $[x_{j-1}, x_j]$.*

Now we will choose tags $\{t_k\}_{k=1}^n$ for $\mathcal{P}$ to get a contradiction to (12). Choose $t_k = x_k$ except for $[x_{j-1}, x_j]$ where f is unbounded. Then choose t_j so that

$$|f(t_j)| > \frac{1}{x_j - x_{j-1}} \left(|A| + 1 + \left| \sum_{k \neq j} f(t_k)(x_k - x_{k-1}) \right| \right)$$

(b) Use part (a) and (12) to obtain a contradiction. The triangle inequality

$$\left| \sum_{k=1}^n f(t_k)(x_k - x_{k-1}) \right| \geq |f(t_j)(x_j - x_{j-1})| - \left| \sum_{k \neq j} f(t_k)(x_k - x_{k-1}) \right|$$

may be helpful.

The following exercises are not needed for the flow of Riemann's discussion, but will sharpen your skills in working with Riemann sums and Riemann's definition of the definite integral.

Exercise 20 *Use Riemann's definition to prove the following: Suppose g is Riemann integrable on $[a, b]$ and $c \in \mathbb{R}$. Then cg is Riemann integrable on $[a, b]$.*

Exercise 21 *Use Riemann's definition to prove the following: Suppose f, g are Riemann integrable on $[a, b]$. Then $f + g$ is Riemann integrable on $[a, b]$.*

Exercise 22 *Is Dirichlet's function ϕ Riemann integrable on $[0, 1]$? Prove your assertion.*

Exercise 23 *Define function $h(x) = 3$ on $[0, 1]$ and $h(x) = 4$ on $(1, 2]$. Is h Riemann integrable on $[0, 2]$? Prove your assertion.*

Exercise 24 *Prove that changing the value of $f(x)$ at a finite number of points in $[a, b]$ will not change whether f is integrable, and will not change the value of $\int_a^b f$ when it exists.*

Exercise 25 *Use Riemann's definition to prove the following: Suppose $f(x) \geq 0$ on $[a, b]$ and f is Riemann integrable on $[a, b]$. Then $\int_a^b f \geq 0$.*

After Riemann gave his new definition of the definite integral, he developed an alternate condition for the existence of $\int_a^b f$. Recall from Excerpt A that the Riemann sum (11) is

$$S = \delta_1 f(a + \epsilon_1 \delta_1) + \delta_2 f(x_1 + \epsilon_2 \delta_2) + \delta_3 f(x_2 + \epsilon_3 \delta_3) + \cdots + \delta_n f(x_{n-1} + \epsilon_n \delta_n).$$

Riemann Excerpt B

Let us examine now, secondly, the range of validity of the concept, or the questions: In which cases can a function be integrated, and in which cases can it not?

We suppose that the sum S converges if the δ's together become infinitely small. We denote by D_1 the greatest fluctuation of the function between a and x_1, that is, the difference of its greatest and smallest values in this interval, by D_2 the greatest fluctuation between x_1 and $x_2, \ldots$, by D_n that between x_{n-1} and b. Then

$$\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n \tag{13}$$

must become infinitely small when the δ's do.

Exercise 26 *Consider the example from Exercise 14 with $f(x) = 2x^3 - 9x^2 + 12x + 1$, $a = 0, b = 3$, and partition $\mathcal{P} = \{0, 1, 3\}$. Calculate D_1, D_2 and the "fluctuation" (13) for this partition $\mathcal{P}$. Are the tags relevant for (13)?*

Exercise 27 *Try to give a brief "big picture" summary of this excerpt.*

Exercise 28 *Since f is not assumed to be continuous in general, we must actually define the D_k a bit differently than Riemann does. Explain why. Then give a definition of the D_k using set notation.*

Note the expression in (13) appears frequently in Riemann's discussion, and roughly measures the total fluctuation of f across the entire partition $\mathcal{P}$. We will name this expression $\operatorname{Fluc}(f, \mathcal{P})$, a function of f and $\mathcal{P}$:

$$\operatorname{Fluc}(f, \mathcal{P}) = \delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n \tag{14}$$

We saw in Exercise 28 that the tags are not relevant for $\operatorname{Fluc}(f, \mathcal{P})$.

Exercise 29 *Use quantifiers and* $\operatorname{Fluc}(f, \mathcal{P})$ *to rewrite Riemann's claim that the fluctuation (13) "must become infinitely small when the δ's do" for integrable f.*

Exercise 30 *Consider the example from Exercise 14 with $f(x) = 2x^3 - 9x^2 + 12x + 1$, $a = 0, b = 3$. For fixed $\epsilon = 0.1$, find a $d > 0$ such that for any partition $\mathcal{P}$ with* $\operatorname{mesh}(\mathcal{P}) < d$, *you can guarantee that* $\operatorname{Fluc}(f, \mathcal{P}) < \epsilon$.

Exercise 31 *Now give a modern proof of Riemann's claim that (13) "must become infinitely small when the δ's do" for integrable f, using Exercises 28, 29 and 18.*

Observe that what Riemann is stating here is an indirect condition for integrability that doesn't involve $\int_a^b f$ itself: if f is integrable, then for any $\epsilon > 0$ we can find $d > 0$ so that for any partition $\mathcal{P}$ with $\operatorname{mesh}(\mathcal{P}) < d$ we are guaranteed that the total fluctuation of f across $\mathcal{P}$ is less than ϵ. It turns out this condition is necessary and sufficient, which we record as a theorem.

Theorem 32 *A function f is Riemann integrable on $[a, b]$ if and only if for all $\epsilon > 0$ there exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of $[a, b]$ with* $\operatorname{mesh}(\mathcal{P}) < \delta$ *then*

$$\operatorname{Fluc}(f, \mathcal{P}) < \epsilon \tag{15}$$

You have shown the necessity of this condition (15) for integrability. The proof of sufficiency is a bit technical. The basic idea is much the same as we outlined in Cauchy Section 2.2. We construct a sequence of partitions with mesh approaching zero and Riemann sums that converge, and prove, using Theorem 32, that the limit of these Riemann Sums is $\int_a^b f$. The details are given in the Supplementary Exercises, Section 3.1.

This characterization of integrability is very powerful. In the next two exercises you will use it to give fairly easy proofs that all continuous and monotone functions are integrable.

Exercise 33 *Use Theorem 32 to prove that if f is continuous on $[a, b]$, then $\int_a^b f$ exists.*

Exercise 34 *Use Theorem 32 to prove that if f is monotone on $[a, b]$, then $\int_a^b f$ exists.*

It may seems obvious that $\int_a^b f = \int_a^c f + \int_c^b f$ for $a < c < b$, but the technical proof is challenging.

Exercise 35 *Use Theorem 32 to prove the following theorem.*

Split Interval Theorem. Let $f : [a, b] \to \mathbb{R}$ and $a < c < b$. Then f is Riemann integrable over $[a, b]$ if and only if f is Riemann integrable over both $[a, c]$ and $[c, b]$. In this case,

$$\int_a^b f = \int_a^c f + \int_c^b f$$

Later in his paper, Riemann constructs a integrable function with infinite discontinuities that does not satisfy Dirichlet's Condition 1 (the graph is displayed in the project Introduction). Well after Riemann's work, the mathematician Carl Thomae (1840-1921) devised another function with infinite discontinuities that is easier to show is integrable with the tools we've developed so far.

Thomae's Function. Define $T(x) : [0, 1] \to \mathbb{R}$ by $T(x) = 0$ for irrational x , $T(0) = 1$, and $T(m/n) = 1/n$ for rational $x = m/n$ where m/n is in reduced form.

Exercise 36 *Show that T is continuous at all irrationals and discontinuous at all rationals.*

Exercise 37 *Use Theorem 32 to prove that T is integrable.*

3.1 Appendix: Supplementary exercises on the $\text{Fluc}(f, \mathcal{P})$ sufficiency condition

We saw in Section 2 that Cauchy defined the definite integral $\int_a^b f$ for continuous f in a rather imprecise way as a limit of sums $S(f, \mathcal{P})$. He also showed that if two partitions $\mathcal{P}_1, \mathcal{P}_2$ had sufficiently small mesh, then we could make the difference $S(f, \mathcal{P}_1) - S(f, \mathcal{P}_2)$ arbitrarily small.

Riemann also gave a condition for integrability in Theorem 32 using $\text{Fluc}(f, \mathcal{P})$ instead of $S(f, \mathcal{P}_1) - S(f, \mathcal{P}_2)$ and in Section 3 we proved the necessity but not the sufficiency of that condition. In the exercises below, you will prove the sufficiency. That is:

If for all $\epsilon > 0$ there exists $d > 0$ such that

$$\text{if } \mathcal{P} \text{ is a partition with } \text{mesh}(\mathcal{P}) < d, \text{ then } \text{Fluc}(f, \mathcal{P}) < \epsilon \tag{16}$$

holds, then f is Riemann integrable on $[a, b]$.

To carry out this proof, a "Fluctuation Refinement Lemma" will be useful:

Fluctuation Refinement Lemma. Suppose f is bounded on $[a, b]$ and that partition $\mathcal{P}'$ is a refinement of $\mathcal{P}$. Then

1. $\text{Fluc}(f, \mathcal{P}') \leq \text{Fluc}(f, \mathcal{P})$
2. $\left| S\left(f, \dot{\mathcal{P}}'\right) - S\left(f, \dot{\mathcal{P}}\right) \right| \leq \text{Fluc}(f, \mathcal{P})$ for any tags of $\mathcal{P}'$ and $\mathcal{P}$.

A complete proof by induction on the number of additional points in refinement is appropriate here. For ease of notation, the following exercise is for the case where $\mathcal{P}'$ adds just one point to $\mathcal{P}$ between a and x_1.

Exercise 38 *Prove this lemma for the case* $\mathcal{P} = \{a, x_1, x_2, \ldots, x_{n-1}, x_n\}$ *and* $\mathcal{P}' = \{a, x', x_1, x_2, \ldots, x_{n-1}, x_n\}$.

Now we don't yet have a candidate for $\int_a^b f$, so we will construct one using a Cauchy sequence of Riemann sums. To do this, first note that by (16) we can construct, for each $n \in \mathbb{N}$, a $d_n > 0$ so that:

1. $d_n \leq d_{n-1}$, and
2. for any partition $\mathcal{P}$ with $mesh\left(\mathcal{P}\right) < d_n$ we have $\operatorname{Fluc}\left(f, \mathcal{P}\right) < 1/n$

Next define a sequence of tagged partitions $\left\{\dot{\mathcal{P}}_n\right\}$ by

1. $\mathcal{P}_{n+1}$ is a refinement of $\mathcal{P}_n$, and
2. $\operatorname{mesh}\left(\dot{\mathcal{P}}_{n+1}\right) \leq \operatorname{mesh}\left(\dot{\mathcal{P}}_n\right) < d_n$.

We will see that any tags will do for the $\dot{\mathcal{P}}_n$.

Exercise 39 *Prove that* $\left\{S\left(f, \dot{\mathcal{P}}_n\right)\right\}$ *is a Cauchy sequence in* $\mathbb{R}$.

Now let A denote $\lim_{n\to\infty} S\left(f, \dot{\mathcal{P}}_n\right)$, the limit for this Cauchy sequence of real numbers. This is our candidate for the integral of f! We will show this using the properties of $\operatorname{Fluc}\left(f, \mathcal{P}\right)$.

If $\dot{Q}$ is an arbitrary tagged partition with small mesh, we need to show its Riemann sum $S\left(f, \dot{Q}\right)$ is close to A. To do this, we will show $S\left(f, \dot{Q}\right)$ is close to some $S\left(f, \dot{\mathcal{P}}_K\right)$ where K is large enough to guarantee that $\left|S\left(f, \dot{\mathcal{P}}_K\right) - A\right|$ is tiny. The following exercises will be useful.

Exercise 40 *Let* $\dot{Q}$ *be a tagged partition. For* $K \in \mathbb{N}$ *and any tags of partition* $\mathcal{P}_K$, *choose* $\dot{\mathcal{P}}^*$ *to be a refinement of both* $\dot{Q}$ *and* $\dot{\mathcal{P}}_K$ *with any tags. Then show that*

$$\left|S\left(f, \dot{Q}\right) - A\right| \leq \left|S\left(f, \dot{Q}\right) - S\left(f, \dot{\mathcal{P}}^*\right)\right| + \left|S\left(f, \dot{\mathcal{P}}^*\right) - S\left(f, \dot{\mathcal{P}}_K\right)\right| + \left|S\left(f, \dot{\mathcal{P}}_K\right) - A\right|.$$

Exercise 41 *Fix* $\epsilon > 0$. *Choose* $K > 1/3\epsilon$. *Choose* d *appropriately and use the Fluctuation Refinement Lemma and above exercises to show that*

$$\left|S\left(f, \dot{Q}\right) - A\right| < \epsilon.$$

Exercise 42 *Use the exercises above to prove that if* f *satisfies (16) then* f *is integrable on* $[a, b]$.

4 Conclusion

Riemann's definite integral raised new questions about the nature of $\int_a^b f$ as well as answering some old ones. On the one hand, he showed that you could integrate a function that has an infinite number of discontinuities densely packed into a bounded interval. This was mind-boggling for many mathematicians of his era! His necessary and sufficient conditions give new insight into how much a function can fluctuate at discontinuities and still remain integrable.

On the other hand, new questions about rules for handling integrals and infinite series occur naturally from his work. For example, can you evaluate his function legitimately by interchanging the integration and infinite sum? That is, can you integrate term by term:

$$\sum_{n=1}^{\infty} \int_a^b \frac{E(nx)}{n^2}\, dx \overset{???}{=} \int_a^b \sum_{n=1}^{\infty} \frac{E(nx)}{n^2}\, dx$$

This general question does not have an easy answer, and mathematicians in the 1800's had examples where term by term integration works fine, and other examples where it does not. The mathematician Henri Lebesgue (1875-1941) became convinced that an entirely new type integral was needed, and developed his own theory of integration, largely developed in his 1902 thesis. The Lebesgue integral has become very important in many fields of mathematics and statistics, and is frequently studied in graduate school.

References

[C] Cauchy, A.L. 1823. Résumé des leçons données à l'École royale polytechnique sur le calcul infinitésimal. Paris: De Bure.

[D] Dirichlet, J. P. G. Lejeune 1829. Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données, Journal de Crelle 4 p. 157-169.

[R] Riemann, B. 1854. Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe. Habilitationsschrift, Universität Göttingen.

5 Instructor Notes

This project is designed to introduce the definite integral with some historical background for a course in Analysis. The project starts with Cauchy's definite integral. Cauchy gives a more detailed development than Riemann, even if some aspects are specific to continuous functions. The first few Cauchy excerpts ease students into the ideas and notation of partitions, mesh, and refinements. It may interest students to see that Riemann's definition for a Riemann sum is identical to Cauchy's reworked formulation (7).

Project Content Goals

1. Learn the basics of Cauchy's definite integral, including the concepts of partition, mesh, refinements and (Riemann) sums.
2. Learn the basics of Riemann's definite integral definition.
3. Develop elementary properties of the Riemann integral.
4. Learn about and use Riemann's "fluctuation" condition for integrability.

Preparation of Students

Students have done a rigorous study of limits, continuity and derivatives for real-valued functions.

Preparation for the Instructor

This is roughly a two week project under the following methodology (basically David Pengelley's "A, B, C" method described on his website):

1. Students do some advanced reading and light preparatory exercises before each class. This should be counted as part of the project grade to ensure students take it seriously. Be careful not to get carried away with the exercises or your grading load will get out of hand! Some instructor have students write questions or summaries based on the reading.
2. Class time is largely dedicated to students working in groups on the project - reading the material and working exercises. As they work through the project, the instructor circulates through the groups asking questions and giving hints or explanations as needed. Occasional student presentations may be appropriate. Occasional full class guided discussions may be appropriate, particularly for the beginning and end of class, and for difficult sections of the project. I have found that a "participation" grade suffices for this component of the student work. Some instructors collect the work. If a student misses class, I have them write up solutions to the exercises they missed. This is usually a good incentive not to miss class!
3. Some exercises are assigned for students to do and write up outside of class. Careful grading of these exercises is very useful, both to students and faculty. The time spent grading can replace time an instructor might otherwise spend preparing for a lecture.

If time does not permit a full implementation with this methodology, instructors can use more class time for guided discussion and less group work for difficult parts of the project.

Section 1 Introduction

This material is included mostly to motivate the need for a rigorous treatment of the definite integral, especially for integrands with discontinuities. Analysis students have most likely encountered Dirichlet's function ϕ while studying continuity, but may not be aware of Dirichlet's motivations in creating this nowhere-continuous function.

Section 2 Cauchy

Cauchy's development of the integral with "new modes of division" (partition refinements) is quite useful for developing techniques for working with the Riemann integral, especially since Riemann does not spend much time developing properties of the integral.

Cauchy's argument in Excerpt C uses the uniform continuity of integrand f; this will be needed again in a Section 3 exercise showing that continuous functions are Riemann integrable.

Section 3 Riemann

Since Riemann does not spend much time developing properties of the integral, some elementary properties are inserted between Excerpts A and B. While these are not essential for reading the rest of Riemann's work, instructors may sample the set for classroom examples or homework problems.

A detailed exploration of Riemann's "fluctuation" expression (13) is important. He explicitly uses this fluctuation in a necessary condition for a function being integrable. He doesn't show the sufficiency (which is difficult and left to Supplementary Exercises Section 3.1), but uses it later in his paper. Some modern authors develop *oscillation* expressions very much like Riemann's fluctuation expression. It is interesting to note Riemann uses the maximum of various expressions where a modern treatment requires a supremum.

Riemann's work through Excerpt B, summarized in Theorem 32, can be used to develop a great number of integration properties, some of which are given in the exercises. Thomae's function is given as a relatively simple example of an integrable function despite being discontinuous on the rationals.

LaTeX code of this entire PSP is available from the author by request to facilitate preparation of 'in-class task sheets' based on tasks included in the project. The PSP itself can also be modified by instructors as desired to better suit their goals for the course.

Acknowledgments

The development of this project has been partially supported by the National Science Foundation's Improving Undergraduate STEM Education Program under Grants No. 1523494, 1523561, 1523747, 1523753, 1523898, 1524065, and 1524098. Any opinions, findings, and conclusions or recommendations expressed in this project are those of the author and do not necessarily reflect the views of the National Science Foundation.

For more information about TRIUMPHS, visit http://webpages.ursinus.edu/nscoville/TRIUMPHS.html.