The Definite Integrals of Cauchy and Riemann

March 24, 2016

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 an error in Cauchy's proof.

Dirichlet gives a proof of Fourier series convergence that is valid for a function f with finite discontinuities and finite number of extrema in his 1829 paper. He then discusses the possibility of extending his proof for a function f with infinite extrema (in a bounded interval), but he doesn't hold much hope for functions with infinite discontinuities. To indicate why, he gives an example that quickly became famous in mathematical circles of his day. Here is Dirichlet in [D] discussing the Fourier series for a function with discontinuities.

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"

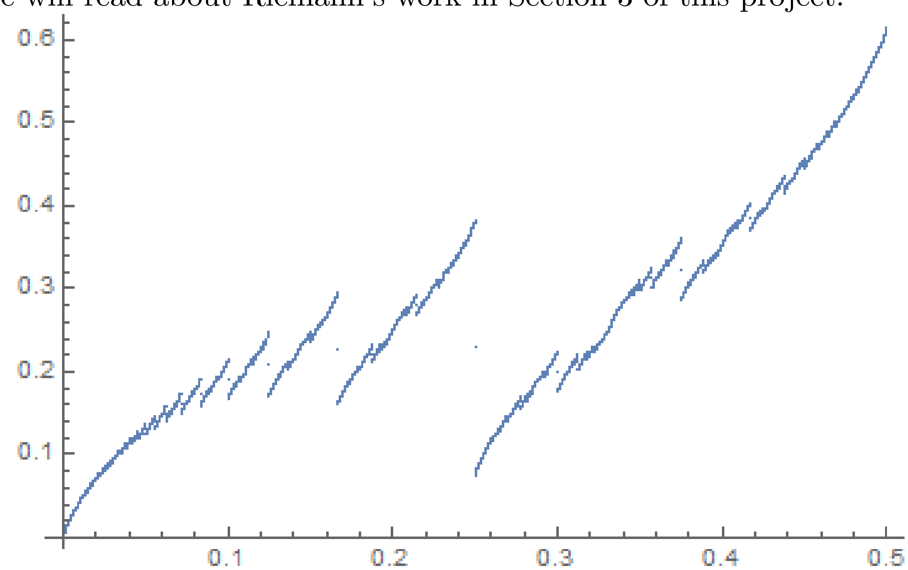
Exercise 1 *Consider the example function $\phi(x)$ Dirichlet gives in the excerpt. Dirichlet claims this function does not satisfy the condition:*

Condition 2 *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 Condition 2.

It is important to remember that in 1829 the only definition of the definite integral was given by Cauchy, and that 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 a integrable function with infinite discontinuities that does not satisfy Dirichlet's Condition 2 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 his 1823 *Calcul Infinitésimal* [C]. We will read his development over the course of three 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 3 *Make and label a diagram that graphically represents what is going on with Cauchy's construction of S in (3) for a fixed n value. Comment on how you think this relates to $\int_{x_0}^X f(x)\,dx$. Does this S formula remind you of something you've seen in your Introductory Calculus courses?*

Exercise 4 *Consider the example $f(x) = x^2 - 2$, $x_0 = 0$, $x_1 = 1/2$, $x_2 = 3/2$, $X = 2$, $n = 3$ Find the elements $x_1 - x_0, x_2 - x_1, x_3 - x_2$ for this example. Then calculate the sum S for this example. How close is S to $\int_{x_0}^X f(x)\,dx$?*

We will find it convenient to give a 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 notation $S(f, \mathcal{P})$ for Cauchy's sum S to indicate the dependence of S on f and $\mathcal{P}$.

Exercise 5 *Observe that Cauchy makes a bold claim at the very end of the excerpt:*

> ***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 6 *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 name it. Define $\operatorname{mesh}(\mathcal{P})$, the **mesh** of a partition $\mathcal{P}$, to be its maximum element value. For example, $\operatorname{mesh}(\mathcal{P}) = 1$ for the partition $\mathcal{P}$ in Exercise 4.

After making his Claim M, Cauchy gives a proof. Before we read it, we need to know that he uses the term "average" in an unusual way. He calls an **average** of given quantities $a_1, a_2, \ldots, a_n$ a new quantity between the smallest and the largest of those under consideration, and he denotes an average by $M(a_1, a_2, \ldots, a_n)$. For example, any value between 1 and 17 is an average of the values $1, 3, 4, 17$. Cauchy proves several results about his averages, including the following corollary, which we will need to follow his proof of Claim M.

Corollary A. Given two sets of n quantities $y_1, \ldots, y_n$ and $a_1, a_2, \ldots, a_n$ where the $y_1, \ldots, y_n$ all have the same sign. Then

$$\frac{a_1 y_1 + \cdots + a_n y_n}{y_1 + \cdots + y_n} \text{ is an average of } a_1, a_2, \ldots, a_n$$

Equivalently,

$$a_1 y_1 + \cdots a_n y_n = (y_1 + \cdots y_n)\, M(a_1, a_2, \ldots, a_n)$$

for some average $M(a_1, a_2, \ldots, a_n)$ of the quantities $a_1, a_2, \ldots, a_n$.

Exercise 7 *Prove Corollary A.*

- *Let* $Y = \sum y_k$ *and note we just need to prove that* $\frac{1}{Y} \sum a_k y_k$ *is between* $\min(a_k), \max(a_k)$ *case all* $y_k \geq 0$, *from* $\min(a_k) \leq a_j \leq \max(a_k)$ *we get* $\min(a_k) \cdot y_j \leq a_j \cdot y_j \leq \max(a_k) \cdot y_j$ *and then sum and factor out max, min, then divide by* Y

$$\begin{aligned} \min(a_k) \sum y_j &\leq \sum a_j \cdot y_j \leq \max(a_k) \sum y_j \\ \min(a_k) &\leq \frac{1}{Y} \sum a_j \cdot y_j \leq \max(a_k) \end{aligned}$$

We now pick up with Cauchy's discussion directly after the last excerpt where he made his Claim M.

Excerpt B from Cauchy's Calcul Infinitésimal

If we were to suppose all the elements of the difference $X - x_0$ reduce to a single one, which would be this difference itself, we would simply have

$$S = (X - x_0)\, f(x_0)\,. \tag{4}$$

When, on the other had, we take the expressions in (2) for the elements of the difference $X - x_0$, the value of S, determined in this case by equation (3), is equal to the sum of the elements multiplied by an average between the coefficients

$$f(x_0)\,, \quad f(x_1)\,, \quad f(x_2)\,, \quad \ldots, \quad f(x_{n-1})\,.$$

Exercise 8 *Use Cauchy's Corollary A to help write out the proof details of his statement below (4) that the sum S in (3) is the sum of the elements multiplied by an average $M(f(x_0), \ldots, f(x_{n-1}))$.*

After this Excerpt B, Cauchy continues his discussion of $S = S(f, \mathcal{P})$ for partition $\mathcal{P} = \{x_0, x_1, x_2, \ldots, x_{n-1}, X\}$ He claims that from $S = (X - x_0)M(f(x_k))$ we can find a θ for which

$$S = (X - x_0) f\left[x_0 + \theta\,(X - x_0)\right] \tag{5}$$

in which $0 < \theta < 1$.

Exercise 9 *Prove Cauchy's claim (5) using the Intermediate Value Theorem and the previous exercise.*

Now Cauchy will use this reformulation of $S(f, \mathcal{P})$ in (5) on each of the subintervals (x_{k-1}, x_k) to get another expression for S when we slice up the interval (x_0, X) into even smaller subintervals. Here is how he describes it, directly after the last excerpt B.

Excerpt C 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

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

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

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

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

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

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 4 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, the sum in (3) is $S(f, \mathcal{P})$ and the sum in (7) is $S(f, \mathcal{P}')$.

Exercise 10 *Draw a diagram of the refinement that reflects Cauchy's construction to obtain (6). Then write out the details in modern notation.*

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

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

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

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 D from Cauchy's Calcul Infinitésimal

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

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

we will derive

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

then, by developing products,

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

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 $S\left(f,\mathcal{P}\right), S\left(f,\mathcal{Q}\right)$ 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 11 *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 $\left|S\left(f,\mathcal{P}\right)-S\left(f,\mathcal{P}'\right)\right|<\epsilon$.*

Exercise 12 *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 13 *Use Cauchy's ideas to give a modern proof of Lemma 11.*

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 E 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 14 *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 15 *Suppose we are given a continuous function g on $[a, b]$ and $\epsilon = 0.1$. Further, suppose we find the value d from Lemma 11 for $\epsilon/2$, and two partitions $\mathcal{P}_1$,$\mathcal{P}_2$ each with mesh less than d. Use Cauchy's reasoning and Lemma 11 to prove that*

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

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 16 *State and prove a modern version of Claim M that generalizes the previous exercise.*

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 16 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 3.3.

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

Exercise 17 *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 16.

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 $\operatorname{mesh}(\mathcal{P}_1), \operatorname{mesh}(\mathcal{P}_2) < d$, then $|S(f, \mathcal{P}_1) - S(f, \mathcal{P}_2)| < \epsilon$.

Exercise 18 *Prove that your theorem from Exercise 16 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 infinite but isolated discontinuities and infinite 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, then f is **Riemann integrable** on $[a,b]$, and we will write $\int_a^b f$ for the definite integral.

Exercise 19 *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 20 *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** $t_k = x_k + \epsilon_k\delta_k$ 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 21 *What are the tags for the example in Exercise 19?*

Exercise 22 *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 23 *Using appropriate quantifiers and modern notation for tagged partitions and mesh, rewrite Riemann's definition 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 24 *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* $\operatorname{mesh}\left(\dot{\mathcal{P}}\right) < \delta$ *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 $\operatorname{mesh}(\mathcal{P}) < \delta$. 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 25 *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 26 *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 27 *Is Dirichlet's function ϕ Riemann integrable on $[0,1]$? Prove your assertion.*

Exercise 28 *Define function h to be 3 on $[0,1]$ and $h(x) = 4$ on $(1,2]$. Is h Riemann integrable on $[0,2]$? Prove your assertion.*

Exercise 29 *Prove that changing the value of $f(x)$ at a finite number of points in $[a,b]$ will not change whether f in integrable, and will not change the value of $\int_a^b f$ when it exists.*

Exercise 30 *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 gives his new definition of the definite integral, he now develops an alternate condition for the existence of $\int_a^b f$.

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

Exercise 32 *Consider the example from Exercise 19 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 33 *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 the preceding exercise that the tags are not relevant for $\operatorname{Fluc}(f, \mathcal{P})$.

Exercise 34 *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 35 *Consider the example from Exercise 19 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 36 *Now give a modern proof of Riemann's claim that (13) "must become infinitely small when the* δ*'s do" for integrable* f.

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 37 *A function f is Riemann integrable on $[a, b]$ if and only if for all $\epsilon > 0$ there exists $d > 0$ such that if $\mathcal{P}$ is a partition $\mathcal{P}$ 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 37, that the limit of these Riemann Sums is $\int_a^b f$. The details are given in the Supplementary Exercises, Section 3.3.

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 38 *Use Theorem 37 to prove that if f is continuous on $[a, b]$, then $\int_a^b f$ exists.*

Exercise 39 *Use Theorem 37 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 40 *Use Theorem 37 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,

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

As we shall see in Section 3.4, Riemann constructs a integrable function with infinite discontinuities that does not satisfy Dirichlet's Condition **??**. 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 w'eve 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 41 *Show that T is continuous at all irrationals and discontinuous at all rationals.*

Exercise 42 *Use Theorem 37 to prove that T is integrable.*

3.1 Riemann's new necessary condition for integrability

After Excerpt B, Riemann immediately set about creating a new necessary and sufficient condition for integrability. He will use this condition to define his integrable function with infinite discontinuities. Naturally these discontinuities can't be too huge: remember Dirichlet's function ϕ! Riemann needs to measure the contribution of these discontinuities to the overall fluctuation of the integrand f.

We pick up with Riemann right after Excerpt B, where he is considering a bounded, integrable function f.

Riemann Excerpt C

We suppose further, that Δ is the greatest value this sum [(13)] can reach, as long as all of the δ's are smaller than d. Then Δ will be a function of d, which is decreasing with d and becomes infinitely small with d. Now, if the total length of the intervals, in which the fluctuation is greater than σ, is L[1], then the contribution of these intervals to the sum $\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$ is clearly $\geq \sigma L$. Therefore one has

$$\sigma \leq \delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n \leq \Delta, \quad \text{hence } L \leq \frac{\Delta}{\sigma} \tag{16}$$

Now if σ is given, Δ/σ can always be made arbitrarily small by a suitable choice of d. The same is true for L, which yields:

In order for the sum S to converge whenever all the δ's become infinitely small, in addition to $f(x)$ being finite, it is necessary that the total length of the intervals, in which the fluctuations exceed σ, can be made arbitrarily small for any given σ by a suitable choice of d.

Let's start with some exercises exploring Riemann's new concepts Δ & L, and inequality (16). Since $x_{k-1} < t_k < x_k$ and f is not assumed to be continuous, we must actually define Δ as a function of d a bit differently than Riemann does. We will need to use the supremum instead of maximum and define

$$\Delta(d) = \sup \{\text{Fluc}(f, \mathcal{P}) : \text{mesh}(\mathcal{P}) < d\} \tag{17}$$

Exercise 43 *Consider the example $f(x) = 2x^3 - 9x^2 + 12x + 1$ from Exercise 19, but with $a = 0, b = 4$.*

(a) Show that $\Delta(1) \geq 34$.

(b) Show that $\Delta(1) \geq 66$.

Exercise 44 *For integrable f, give a modern proof that $\Delta(d)$ "becomes infinitely small with d."*

Riemann next defines L to be the total length of the intervals in which the fluctuation is greater than σ. Observe that L is a function of σ and the partition $\mathcal{P}$, so we can write $L = L(\sigma, \mathcal{P})$ where convenient.

[1] Riemann uses s here; we have used L to suggest length and avoid confusing with the sum symbol S

Exercise 45 *Consider* $f(x) = 2x^3 - 9x^2 + 12x + 1$, $a = 0, b = 4$ *and partition* $\mathcal{P} = \{0, 1, 3, 4\}$ *from Exercise 43. Calculate* $L(6, \mathcal{P})$ *and* $L(2, \mathcal{P})$

Let's introduce some notation for working with the intervals in which the fluctuation is greater than σ. For any partition $\mathcal{P}$ and $\sigma > 0$, let J denote the set of indices k for which $D_k > \sigma$. We can then write

$$L(\sigma, \mathcal{P}) = \sum_{k \in J} \delta_k \quad \text{and} \quad L(\sigma, \mathcal{P}) + \sum_{k \notin J} \delta_k = b - a$$

For the previous example, you should verify that for $\sigma = 6$, $J = \{3\}$ and for $\sigma = 2$, $J = \{1, 2, 3\}$

Exercise 46 *Explain in your own words what Riemann means by "the contribution of these intervals to the sum* $\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$*". Find this contribution for the last example with* $\sigma = 6$*. Then verify this contribution is* $\geq \sigma L$ *with* $\sigma = 6$*. Also verify* $\sigma \leq \delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$ *for* $\sigma = 6$.

Exercise 47 *Now give a general proof that "the contribution of these intervals to the sum* $\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$*" is* $\geq \sigma L$.

Exercise 48 *Use your results to give a modern, general proof of Riemann's claim in (16) that* $L(\sigma, \mathcal{P}) \leq \dfrac{\Delta(d)}{\sigma}$ *for any partition* $\mathcal{P}$ *with* $\operatorname{mesh}(\mathcal{P}) < d$.

After developing his new concepts Δ & L and inequality (16), Riemann gives a new necessary condition for integrability at the end of Excerpt C.

Exercise 49 *The last sentence of Excerpt C is the theorem statement. Rewrite "In order for the sum* S *to converge whenever all the* δ*'s become infinitely small, in addition to* $f(x)$ *being finite, it is necessary that the total length of the intervals, in which the fluctuations exceed* σ*, can be made arbitrarily small for any given* σ *by a suitable choice of* d*." as a theorem with quantifiers and modern notation.*

Exercise 50 *Give a modern proof of the this theorem you wrote in modern notation in Exercise 49. Theorem 37 and Exercise 48 should be useful.*

The following exercises are not needed for the flow of Riemann's discussion, but will sharpen your skills in working with Riemann's new condition for integrability.

Exercise 51 *We know Dirichlet's function* ϕ *is not integrable so we can't expect the theorem to apply to this function. Show this directly. That is, show that "for all* $\epsilon > 0$ *and for all* $\sigma > 0$ *there exists a* $d > 0$ *such that if* $\operatorname{mesh}(\mathcal{P}) < d$ *then* $L(P, \sigma) < \epsilon$*" is false for Dirichlet's function* ϕ.

Exercise 52 *Consider* $f(x) = 2x^3 - 9x^2 + 12x + 1$, $a = 0, b = 4$ *and partition* $\mathcal{P} = \{0, 1, 3, 4\}$ *from Exercise 43. Given* $\sigma = 2$ *and* $\epsilon = 1$*, and* $d = 0.1$ *, show that for* $\operatorname{mesh}(\mathcal{P}) < d$ *we have* $L(\sigma, \mathcal{P}) < \epsilon$.

3.2 A new sufficient condition for integrability

In the previous Excerpt C, Riemann gives a necessary condition for integrability in terms of the total length of the intervals, in which the fluctuations exceed $\sigma > 0$. He next uses this condition to give a sufficient condition for integrability.

Riemann Excerpt D

This theorem also has a converse:

If the function $f(x)$ is always finite, and by infinitely decreasing the δ's together, the total length L of the intervals in which the fluctuation of the function is greater than a given number σ always becomes infinitely small, then the sum S converges as the δ's become infinitely small together.

For those intervals in which the fluctuations are $> \sigma$ make a contribution to the sum $\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$, less than L times the largest fluctuation of the function between a and b, which is finite (by agreement). The contribution of the remaining intervals is $< \sigma(b-a)$. Clearly one can now choose σ arbitrarily small and then always determine the size of the intervals (by agreement) so that L is also arbitrarily small. In this way the sum $\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$ can be made as small as desired. Consequently the value of the sum S can be enclosed between arbitrarily narrow bounds.

Thus we have found necessary and sufficient conditions for the sum S to be convergent when the quantities δ tend together to zero, or equivalently, for the existence of the integral of $f(x)$ between a and b.

We saw in Exercise 49 that Riemann's necessary theorem "In order for the sum S to converge whenever all the δ's become infinitely small, in addition to $f(x)$ being finite, it is necessary that the total length of the intervals, in which the fluctuations exceed σ, can be made arbitrarily small for any given σ by a suitable choice of d" could be rewritten as:

Suppose f is integrable and bounded on $[a,b]$. Then for all $\epsilon > 0$ and for all $\sigma > 0$ there exists a $d > 0$ such that if $\operatorname{mesh}(\mathcal{P}) < d$ then $L(\mathcal{P}, \sigma) < \epsilon$.

Exercise 53 *Rewrite Riemann's converse theorem in Excerpt D using quantifiers and modern terminology.*

Let's rewrite Riemann's proof of this theorem using modern notation over the next exercises.

Exercise 54 *Suppose f is bounded on $[a,b]$ and $\epsilon > 0$ is fixed. Use Riemann's ideas to show that you can find $\sigma > 0$ so that for any partition $\mathcal{P}$,*

$$\sum_{k \notin J} \delta_k D_k < \frac{\epsilon}{2}$$

Exercise 55 *Suppose f is bounded on $[a,b]$ and $\epsilon, \sigma > 0$ are fixed. Use Riemann's ideas to show that you can find $d > 0$ so that for if* $\operatorname{mesh}(P) < d$ *then*

$$\sum_{k \in J} \delta_k D_k < \frac{\epsilon}{2}$$

Exercise 56 *Use the previous two exercises together with Theorem 37 to write a modern proof of Riemann's converse theorem.*

Over the last two subsections, we have explored Riemann's necessary and sufficient condition for the integrability of f in terms of $L(\mathcal{P},\sigma)$, the total length L of the intervals in which the fluctuation of the function is greater than a given number σ. Riemann will next use this result to show the integrability of a function with infinite discontinuities that he designed.

3.3 Supplementary exercises on the $\mathrm{Fluc}(f,\mathcal{P})$ 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 37 using $\mathrm{Fluc}(f,\mathcal{P})$ instead of $S(f,\mathcal{P}_1)-S(f,\mathcal{P}_2)$ and we proved the necessity but not the sufficiency in Section 3. In the exercises below, you will prove the sufficiency. That is:

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

$$\text{if } \mathcal{P} \text{ is a partition with } \operatorname{mesh}(\mathcal{P})<d, \text{ then } \mathrm{Fluc}(f,\mathcal{P})<\epsilon \tag{18}$$

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. $\mathrm{Fluc}(f,\mathcal{P}')\le \mathrm{Fluc}(f,\mathcal{P})$
2. $\left|S\left(f,\dot{\mathcal{P}}'\right)-S\left(f,\dot{\mathcal{P}}\right)\right|\le \mathrm{Fluc}(f,\mathcal{P})$ 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 57 *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 (18) we can construct, for each $n\in\mathbb{N}$, a $d_n>0$ so that:

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

Now 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 58 *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. This is our candidate for the integral of f! We will show this using the properties of $\operatorname{Fluc}(f, \mathcal{P})$.

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 59 *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 60 *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 61 *Use the exercises above to prove that if* f *satisfies (18) then* f *is integrable on* $[a, b]$.

3.4 Riemann's integrable function with infinite discontinuities

With his new definition of the definite integral in hand, Riemann wanted to show that many more functions could be integrated, which would in turn expand the set of functions that could have a sensible Fourier series expansion. In particular, he wanted to construct an integrable function that met Dirichlet's Condition 2, which we discussed in Exercise 1.

After developing his necessary and sufficient conditions for integrability (see Exercises 49 and 53), Riemann was ready for his construction.

Here is a notation note we need to read his excerpt below: Riemann uses $f(x+0)$ for a right-hand limit where we would write $\lim_{z\to x^+} f(z)$, and he uses $f(x-0)$ where we would write $\lim_{z\to x^-} f(z)$.

Riemann Excerpt E

We consider functions which are discontinuous infinitely often between any two numbers, no matter how close.

Since these functions have never been considered before, it is well to start from a particular example. Designate, for brevity, $E(x)$ to be the excess of x over the closest integer[2], or if x lies in the middle between two (and thus the determination is ambiguous) the average of the two numbers $1/2$ and $-1/2$, hence zero. Furthermore, let n be an integer and let p an odd integer, and form the series

$$f(x) = \frac{E(x)}{1} + \frac{E(2x)}{4} + \frac{E(3x)}{9} + \cdots = \sum_{n=1}^{\infty} \frac{E(nx)}{n^2} \tag{19}$$

It is easy to see that the series converges for each value of x. When the argument continuously decreases to x, as when it continuously increases to x, the value always approaches a fixed limit. Indeed, if $x = \frac{p}{2n}$ (where n and p are relatively prime $[\text{and } p \text{ is odd}]^3$)

$$\begin{aligned} f(x+0) &= f(x) - \frac{1}{2n^2}\left(1 + \frac{1}{9} + \frac{1}{25} + \cdots\right) = f(x) - \frac{\pi^2}{16n^2}, \\ f(x-0) &= f(x) + \frac{1}{2n^2}\left(1 + \frac{1}{9} + \frac{1}{25} + \cdots\right) = f(x) + \frac{\pi^2}{16n^2}; \end{aligned} \tag{20}$$

in all other cases $f(x+0) = f(x)$ and $f(x-0) = f(x)$.

Hence this function is discontinuous for each rational value of x which in lowest terms is a fraction with even denominator. Thus while f is discontinuous infinitely often between any two bounds, the number of jumps greater than a fixed number is always finite.

Riemann makes quite a few claims in this excerpt, all correct, but some of which are very difficult to verify rigorously with the tools we have developed so far. Nevertheless, the details are worth exploring. In fact, this example motivated other mathematicians to develop more mathematical theory.

By the "excess of x over the nearest integer", Riemann means literally "x minus the nearest integer" except at odd multiples of $1/2$. For example, $E(10/7) = 3/7$, $E(5/6) = -1/6$, and $E(7/2) = 0$.

Riemann also uses the fact $1+\frac{1}{9}+\frac{1}{25}+\cdots = \pi^2/8$, which Euler had shown long before Riemann's era.

Let's start with some exercises to get a sense of the behavior of Riemann's amazing function f. Recall a graph of f was given in the project introduction.

Exercise 62 *(a) Sketch the graphs of* $\frac{E(x)}{1}, \frac{E(2x)}{4}, \frac{E(3x)}{9}$ *and note where they have discontinuities.*

[2] Riemann uses simply (x) instead of $E(x)$ for this excess function. We give this function the name E for clarity.

[3] Riemann did not explicitly state that p must be odd

(b) *Then give a general description of the graph of the partial sum function* $P_3(x) = \dfrac{E(x)}{1} + \dfrac{E(2x)}{4} + \dfrac{E(3x)}{9}$ *and list its points of discontinuity.*

(c) *Verify that these points are all of the form* $p/2n$ *where* n *and* p *are relatively prime and* p *is odd.*

Exercise 63 *Explain in your own words why adding more terms* $E(nx)/n^2$ *to* $P_3(x)$ *will result in discontinuities "at each rational value of* x *which in lowest terms is a fraction with even denominator" for function* f.

Exercise 64 *Justify Riemann's claim that the series (19) converges for each fixed value of* x*. Your Introductory Calculus knowledge of series convergence should be helpful.*

Exercise 65 *Justify Riemann's claim that "*f *is discontinuous infinitely often between any two bounds.*

Let's now consider Riemann's claims about the jumps in $f(x)$ at the discontinuities.

Exercise 66 *To verify (20), suppose* $x = \dfrac{p}{2q}$
where q *and* p *are relatively prime and* p *is odd.*

(a) *Find the following:*

$$\lim_{z\to x^+} E(z) - E(x) \quad \textit{and} \quad \lim_{z\to x^-} E(z) - E(x)$$

(b) *Find the following when* n *is an odd multiple of* q, $n = mq$ *for odd* $m \in \mathbb{N}$.

$$\lim_{z\to x^+} \frac{E(nz)}{n^2} - \frac{E(nx)}{n^2} \quad \textit{and} \quad \lim_{z\to x^-} E(nz) - E(nx)$$

(c) *Find the following when* n *is not an odd multiple of* q :

$$\lim_{z\to x^+} \frac{E(nz)}{n^2} - \frac{E(nx)}{n^2} \quad \textit{and} \quad \lim_{z\to x^-} E(nz) - E(nx)$$

(d) *Find the following:*

$$\sum_{n=1}^{\infty} \lim_{z\to x^+} \left(\frac{E(nz)}{n^2} - \frac{E(nx)}{n^2}\right) \quad \textit{and} \quad \sum_{n=1}^{\infty} \lim_{z\to x^-} \left(\frac{E(nz)}{n^2} - \frac{E(nx)}{n^2}\right)$$

We can use this exercise to justify Riemann's claims about the jumps in $f(x)$ at the discontinuities if we allow ourselves to interchange the limit and infinite sum:

$$\sum_{n=1}^{\infty} \lim_{z\to x^+} \left(\frac{E(nz)}{n^2} - \frac{E(nx)}{n^2}\right) = \lim_{z\to x^+} \sum_{n=1}^{\infty} \frac{E(nz)}{n^2} - \frac{E(nx)}{n^2} \tag{21}$$

This uses the concept of *uniform convergence*, which is beyond the scope of this project. Accepting the validity of (21) we can use the previous exercise to find $\lim_{z\to x^+} f(x) - f(x)$ when $x = \frac{p}{2q}$ where q and p are relatively prime and p is odd, and thus verify Riemann's jumps (20).

Exercise 67 *Fill in the details justifying (20) using (21) and the previous exercise.*

Exercise 68 *Justify Riemann's claim that "f is discontinuous infinitely often between any two bounds, the number of jumps greater than a fixed number is always finite."*

Exercise 69 *Use the interchange of limit and infinite sum in (21) to justify Riemann's claim that "$f(x+0) = f(x)$ and $f(x-0) = f(x)$" and thus f is continuous for all x other than rational x which in lowest terms is a fraction with even denominator.*

Now that we have a handle on how Riemann's function f behaves, let's see how he shows it is integrable.

Riemann Excerpt F

The function is everywhere integrable. Besides finiteness, it has two properties, that for each value of x it has limiting values $f(x+0)$ and $f(x-0)$ on both sides, and that the number of jumps greater than or equal to a given number σ is always finite. Applying our above investigation, as an obvious consequence of these conditions, d can be taken so small that in the intervals which do not contain these jumps, the fluctuations are smaller than σ, and the total length of the intervals which do contain these jumps will be arbitrarily small.

Exercise 70 *Consider Riemann's function f on any interval $[a, b]$.*

(a) Use Riemann's discussion in Excerpt F to write a modern proof that for all $\epsilon > 0$ and for all $\sigma > 0$ there exists a $d > 0$ such that if $\mathcal{P}$ is a partition with $\operatorname{mesh}(\mathcal{P}) < d$ *then $L(\mathcal{P}, \sigma) < \epsilon$.*

(b) Use part (a) to explain why Riemann's function 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 handing 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.