THE RIEMANN-DARBOUX INTEGRAL (REVISED: 2/16/05)

PETE L. CLARK

1. Introduction

Let $f : [a, b] \to \mathbb{R}$ be a function. Recall from calculus that there is supposed to be a number, the *definite integral* $\int_a^b f$, obtained as the limit of a sequence of approximations. We divide the interval $[a, b]$ up into subintervals $\Delta_i = [x_{i-1}, x_i]$ and replacing f on each of its subintervals by a constant function $f(x_i^*)$, where x_i^* is some point in Δ_i. The sum of the areas under the rectangles is given by $\sum_i f(x_i^*)\Delta_i$, which is regarded as an approximation to the "true area under the curve $y = f(x)$." Moreover we are told that by dividing the interval into sufficiently many subintervals, in such a way (we are now listening very carefully!) that the maximum length of each subinterval approaches zero, then this "sequence" of Riemann sums approaches a number, which is by definition $\int_a^b f$. (Many less ambitious calculus texts will tell us to divide $[a, b]$ into n equally spaced subintervals.)

What are we to make of this recipe? First, that the limiting process involved is both vague and sufficiently complicated so that it is not immediately apparent how to turn it into a rigorous "For all $\epsilon > 0$, there exists ... such that ... $< \epsilon$" style definition. For instance, in contrast to the limit of a sequence, a series or a function at a point, there are *choices* involved here in the endpoints of the partition $\{x_0, x_1, \ldots, x_n\}$ and also the values x_i^* on each subinterval, and there is legitimate doubt how we are supposed to quantify over all these choices – i.e., do we require convergence of the Riemann sums associated to just one sequence of partitions (with maximal subinterval length approaching zero) – which raises the question of whether our definition would change if we made another such choice – or do we somehow quantify over all possible sequences of partitions (notice that the set of all partitions of $[a, b]$ is uncountable, so this is certainly not the limit of an ordinary sequence), and if so how?

The second thing that may strike us is that in calculus the only functions that we actually integrate are very nice ones: almost certainly they are continuous (in our wildest calculus dreams we might admit functions with finitely many points of discontinuity). For a sufficiently "nice" function the idea of the area under the curve has at least an intuitive meaning, and it is worth hoping that any reasonable approximation process will get us arbitarily close to this platonic "area" under the curve. However, consider functions like $f_1(x) = 0$ for rational x, 1 for irrational x and $f_2(x) = \frac{1}{q}$ when $0 \neq x = p/q$, and 0 otherwise. The graphs of these functions in the xy-plane cannot be said to be curves at all, so we do not really know – even in the approximate sense – what the "area under y = f$_i(x)$" ought to mean: perhaps

$\int_a^b f_i$ does not exist at all.

Our first order of business is to make this limiting process precise and in so doing arrive at a notion of "integrable functions." The doubts of the previous paragraphs are justified in that there are several plausible ways of enunciating the limiting process. We will show that two different notions integrability are *equivalent* in the sense that the same number $\int_a^b f$ is arrived at for all three and moreover that exactly the same class of functions are integrable under both processes.

2. An axiomatic approach

It seems strange that many calculus books, giving at best incomplete details on what it means for a function to be "integrable," nevertheless include a proof of the fundamental theorem of calculus relating antiderivatives to indefinite integrals. This section serves to explain this phenomenon by giving an axiomatic approach to the Riemann integral: if we admit that the expression $\int_a^b f$ satisfies certain properties – all "obvious" from the geometric interpretation as the area under $y = f(x)$ on $[a, b]$ – then it is forced to satisfy the fundamental theorem. Definition: An *integral* I is a collection of *bounded* functions $\mathcal{R}[a, b]$ defined on each closed interval $[a, b]$ together with an assignment $I : ([a, b], f) \mapsto I([a, b], f) \in \mathbb{R}$ of a real number to each such function, "the integral of f on $[a, b]$," and satisfying the following conditions:

• $\mathcal{R}[a, b]$ contains every continuous function on $[a, b]$.
• If $a < c < b$, then a function f is in $\mathcal{R}[a, b]$ if and only if its restrictions to the subintervals $[a, c]$ and $[c, b]$ are in $\mathcal{R}[a, c]$ and $\mathcal{R}[c, b]$, and $I([a, b], f) = I([a, c], f) + I([c, b], f)$.
• If $f_1 \leq f_2 \leq f_3$ are three functions in $\mathcal{R}[a, b]$ then $I([a, b], f_1) \leq I([a, b], f_2) \leq I([a, b], f_3)$.
• For any constant function C on $[a, b]$, $I([a, b], C) = (b - a)C$.

Theorem 1. *(Fundamental theorem of calculus, first version) Let $f \in \mathcal{R}[a, b]$, and define, for all $x \in [a, b]$, $F(x) := I([a, x], f)$.*
Then the function $F : [a, b] \to \mathbb{R}$ is continuous. Moreover, if f is continuous at a point x_0, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

Proof: By assumption, there exists M such that $|f(x)| \leq M$ on $[a, b]$. Suppose $|x_2 - x_1| < \delta$; we may well assume $x_1 < x_2$. Then

$$|F(x_2) - F(x_1)| = |I([x_1, x_2], f)| \leq I([x_1, x_2], M) = (x_2 - x_1)M.$$

Thus F is uniformly continuous with $\delta = \frac{\epsilon}{M}$.

If we assume that f is continuous at x_0, for any $\epsilon > 0$ we can find a $\delta > 0$ such that $|f(x) - f(x_0)| < \epsilon$ on $(x_0 - \delta, x_0 + \delta)$. Then if $|x - x_0| < \delta$ we have $\frac{F(x)-F(x_0)}{x-x_0} = \frac{1}{x-x_0} I([x_0, x], f)$. Since $f(x) \in (f(x_0) - \epsilon, f(x_0) + \epsilon)$, we have $\frac{1}{x-x_0} I([x_0, x], f) \in (f(x_0) - \epsilon, f(x_0) + \epsilon)$. Therefore, $\lim_{x \to x_0} \frac{F(x)-F(x_0)}{x-x_0} = f(x_0)$, completing the proof of the theorem.

Corollary 2. *There is at most one integral with $\mathcal{R}[a, b] = \mathcal{C}[a, b]$ equal to the set of continuous functions on $[a, b]$: $I([a, b], f) = F(b) - F(a)$, where $F(x)$ is any function such that $F' = f$.*

Proof: The existence of any integral I guarantees that every continuous function f on $[a,b]$ has an antiderivative, namely $F(x) := I([a,x], f)$. Now let I_2 be any other integral, and consider $F_2(x) := I_2([a,x], f)$. We have that $(F - F_2)' = 0$, so by the Mean Value Theorem $F_2 = F + C$. It is easy to check that $\lim_{x \to a^+} F_2(x) = \lim_{x \to a^+ F(x)=0}$, so it must be that $C = 0$ and $F = F_2$. Also, choosing any $x < a$, we have $I([a,b], f) = I([x,b], f) - I([x,a], f) = F(b) - F(a)$.

It remains to be shown that an integral exists. The integral that we are going to discuss in this course is the *Riemann integral*, which is indeed a precision of what one learns in calculus class. However, there is a subtlety here: by our definition, two integrals are the same if they are defined on the same collection of functions and yield the always yield the same number. But this is rather abstract: in practice, an integral is given by some sort of *limiting procedure* which when applied to *any* function $f : [a,b] \to \mathbb{R}$ may or may not yield a number. (This is an analogy to an infinite sequence, an infinite series or the left / right / limit of a function at a point – we can always ask for this number, and it may or may not exist.) When the particular limiting process converges, the function is said to be *integrable*. A highly interesting question to ask is: precisely what is the collection of functions integrable with respect to a given integral process? (We require at least that all continuous functions be integrable, but we may get a much larger supply, and as we will see there are good reasons to hope that the supply is much larger than just the continuous functions.) Then again, it may be that two different processes lead to the same collection of integrable functions and assign the same number to each integrable function – in other words, the corresponding integrals are the same in the previous sense. However, the processes themselves might be phrased differently, and it may be easier to compute with one process than the other.

As an analogy, suppose that two excellent students in a computer science complete the same assignment independently – they will of course write two different programs. Suppose the assignment is to write the program which, given any element s of a certain input set S, returns an output $f(s)$. If the two programs are both correct, they will yield the same function f, but they certainly may do it in different ways (and it is possible that one of the programs takes much longer to run and uses up much more memory than the other). Moreover, if the instructor then experiments with the programs by inputting values outside the input set S, then even the students themselves might not know which inputs will be accepted and what the outputs will be. In our situation, the assigned range of inputs is $\mathcal{C}[a,b]$, the "students"' names are Riemann and Darboux, and it will turn out to be true – but not obvious – that the acceptable range of inputs, i.e., the Riemann-integrable and Darboux-integrable functions, are the same enlargement of the continuous functions, and the integrals themselves are the same. However, it is in practice easier to show that a function is Darboux-integrable than Riemann-integrable, so knowing that the supply of integrable functions is the same will be important in characterizing the set of Riemann-integrable functions.

3. The Riemann integration process

In this section we will give Riemann's integration process and show that it gives an integral in the sense of the preceding section. From this follows the unconditional version of the fundamental theorem of calculus, and especially, the fact that every continuous function has an antiderivative.

We begin by formalizing the notion of a partition together with a choice of point on each subinterval. A *partition* of the closed interval $[a, b]$ is just a finite sequence of points $a = x_0 < x_1 < \ldots < x_n = b$ – so to give a partition is the same as giving any finite subset of $[a, b]$ containing a and b, but we always write the points of the partition in increasing order. We view a partition as dividing $[a, b]$ into subintervals $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$. The *mesh* $||P||$ of a partition $P = \{x_i\}$ is $\max_{i=1}^n x_i - x_{i-1}$, i.e., the largest length of a subinterval comprising the partition. If P and P' are two partitions so that every point of P is also a point of P' (more succinctly $P \subset P'$), then we say P' *refines* P. If P_1, P_2 are partitions, then a partition P_3 which refines both P_1 and P_2 is called a *common refinement* of P_1 and P_2. Common refinements always exist, e.g. $P_3 := P_1 \cup P_2$.

A *tagged partition* (P, τ) is a partition $a = x_0 < x_1 < \ldots < x_n = b$ together with, for all $1 \leq i < n$ a choice of x_i^* such that $x_{i-1} \leq x_i^* \leq x_i$. We say that (P', τ') refines (P, τ) if the underlying partition P' refines P (i.e., we ignore the tags completely and use the previous notion of refinement). To a function f and a tagged partition (P, τ) we associate the Riemann sum

$$R(f, P, \tau) = \sum_{i=1}^{n} f(x_i^{\star})(x_i - x_{i-1}).$$

Definition: (Riemann integrable, Riemann integral) A function $f : [a, b] \to \mathbb{R}$ is said to be Riemann integrable if there exists a number L with the following property: for every $\epsilon > 0$ there exists a $\delta > 0$ such that for all tagged partitions (P, τ) with $||P|| < \delta$, $|R(f, P, \tau) - L| < \epsilon$. If this number exists, it is called the Riemann integral of f on $[a, b]$, and denoted (for now) $R \int_a^b f$.

Exercise 41: Let $f : [0, 1] \to \mathbb{R}$ be the function which takes the value 1 at every rational number and the value 0 at every irrational number. (Note that f is bounded.) Show that f is not Riemann-integrable.

Exercise 42: Suppose that $f : [a, b] \to \mathbb{R}$ is a Riemann-integrable function with the following property: there exists a number L and a sequence of tagged partitions (P_n, τ_n), with $||P_n|| \to 0$, such that $R(f, P_n, \tau_n) = L$. Show that $\int_a^b f = L$.

Example: Let $f : [0, 1] \to \mathbb{R}$ be the function which is zero at $x = 0$ and at every irrational number, while for a rational number written in lowest terms as $\frac{p}{q}$, takes the value $\frac{1}{q}$. We claim that the function is Riemann integrable, with integral 0. (Note that for any partition P we can choose a tagging consisting of irrational numbers x_i^*, whose Riemann sum is zero. Thus, by Exercise 42, if the function is Riemann-integrable, then its integral must be zero.) Because the Riemann sums are visibly non-negative, it suffices to show: for every $\epsilon > 0$, there exists a $\delta > 0$ such that if (P, τ) is any tagged partition of mesh less than δ, $R(f, P, \tau) < \epsilon$. The idea for this is that the function f takes on values at most $\frac{1}{N}$ only on a controlled number of subintervals. More precisely, it can take the value $1/1$ on at most one subinterval (i.e., only on the last one and only if $x_n^* = 1$), it can take the value $\frac{1}{2}$ on at most two subintervals (actually only one, but we are making an estimate), the value $\frac{1}{3}$ on at most three subintervals, and for every N the value $\frac{1}{N}$ on at most N subintervals. If we then consider the contribution to the Riemann sum coming from the subintervals on which the function could take the value $\frac{1}{k}$ for some $k \leq N$, then we find that this part of the sum has at most k terms, each of which is a function value which is at most $\frac{1}{k}$ times a subinterval of length at most δ: so we get $k \cdot \frac{1}{k} \cdot \delta$, for a total contribution of δ. Summing from $k = 1$ to N we find that the contribution of the Riemann sum on all subintervals on which the function could be at most $\frac{1}{N}$ is $N\delta$. On the other subintervals the contribution is at most $\frac{1}{N}$ times the total length of these subintervals, which is certainly no more than the total length of $[0,1]$, i.e., 1. Thus we get:

$$R(f, P, \tau) \leq N\delta + \frac{1}{N}.$$

Recall that we want this sum to be at most ϵ by taking δ sufficiently small: N is an auxiliary quantity that we introduced, which we can take to be any positive integer we want. By taking N sufficiently large so that $\frac{1}{N} < \frac{\epsilon}{2}$ and then choosing δ to be sufficiently small so that $N\delta < \frac{\epsilon}{2}$ (you can unwind this to give an explicit recipe for δ in terms of ϵ if you so desire), we get that $R(f, P, \tau) < \epsilon$ so that f is Riemann integrable.

Note that this function f has the property that its limit at any point $x \in [0,1]$ is equal to zero. In particular it is continuous precisely at the irrational numbers and at $x = 0$, so at all but a countably infinite set of points.

Exercise 43*: Let $X \subset [a,b]$ be any countably infinite subset, which we enumerate as $(x_n)_{n=1}^{\infty}$. Let $(a_n)_{n=1}^{\infty}$ be any real sequence such that $a_n \to 0$. Define a function $f : [a,b] \to \mathbb{R}$ by $f(x_n) = a_n$, and otherwise $f(x) = 0$.
a) Show that f is Riemann-integrable, with $\int_a^b f = 0$.
b) Show that for all $c \in [a,b]$, $\lim_{x \to c} f(x) = 0$. In particular, f is discontinuous precisely at those points x_n for which $a_n \neq 0$.

Theorem 3. *The Riemann integral on the collection of Riemann-integrable functions is an integral in the sense of Section 1.*

Proof: We must show that every Riemann-integrable function is bounded and that the four properties of the integral hold. We will separate this into five different proofs.

Proposition 4. *A Riemann-integrable function $f : [a,b] \to \mathbb{R}$ is bounded.*

Proof: We will show the contrapositive, namely that an unbounded function is not Riemann-integrable. If f is unbounded, then either there exists a sequence of distinct points $\{x_k\} \subset [a,b]$ with $f(x_k) \to +\infty$ or with $f(x_k) \to -\infty$; without much loss of generality, we'll assume the former. Let P_n be the uniform partition into subintervals of length $\frac{b-a}{n}$. For each n, there exists at least one subinterval I_{i_0} in the partition such that $I_{i_0} \cap \{x_k\}$ is infinite. If we then take any tags in all the other subintervals, we get the incomplete Riemann sum $\sum_{i \neq i_0} f(x_i^*)\Delta_i = S$, say. To complete the Riemann sum we must choose a tag x_i^* in I_{i_0}, and then $R(f, P, \tau) = S + f(x_i^*)\frac{b-a}{n}$. But by taking $x_i * = x_k$ for sufficiently large k, we can make the entire Riemann sum as large as we want. Since $||P_n|| \to 0$, we certainly are not getting convergence to any finite number.

Proposition 5. *If C is any real number, $R\int_a^b C = C(b-a)$.*

Exercise 44: Prove Proposition 5.

Proposition 6. *If $f_1 \leq f_2 \leq f_3$ are three Riemann integrable functions, we have*

$$R\int_a^b f_1 \leq R\int_a^b f_2 \leq R\int_a^b f_3.$$

Exercise 45: Prove Proposition 6.

In order to prove the next result, we need the following result, which will allow us to show that a function is or is not Riemann-integrable without making explicit reference to what the integral is.

Proposition 7. *(Cauchy condition for Riemann integrability) A function $f : [a,b] \to \mathbb{R}$ is Riemann-integrable if and only if for every $\epsilon > 0$, there exists a $\delta > 0$ such that for any two tagged partitions (P_1, τ_1), (P_2, τ_2) of mesh less than δ, $|R(f, P_1, \tau_1) - R(f, P_2, \tau_2)| < \epsilon$.*

Exercise 46: Prove Proposition 7. (Suggestion: look in §5.2 of your textbook for part of the proof.)

As we saw in class, the following simple result will allow us not to worry so much when changing tagged partitions by adding a point.

Lemma 8. *(The "It doesn't matter" Lemma) Let $f : [a.b] \to \mathbb{R}$ be a function with $\sup_x |f(x)| = M < \infty$. Let (P, τ) be a tagged partition of mesh at most δ.*
a) Let τ' be the tagging obtained by changing any single sample point x_i^* to w_i^*. Then $|R(f, P, \tau) - R(f, P, \tau')| \leq 2M\delta$.
b) If we form the partition $P \cup \{c\}$, we must change tagging a bit, since the subinterval $[x_{i-1}, x_i]$ containing c now splits into two subintervals. The minimal way to change the tagging would be to throw away x_i^* and take any two sample points $y_i^* \in [x_{i-1}, c]$ and $z_i^* \in [c, x_i]$. Letting τ_c denote this new tagging, we have $|R(f, P, \tau) - R(f, P \cup \{c\}, \tau_c)| \leq 3M\delta$.

Proof: For part a), the difference of the two Riemann sums is precisely

$$|(x_i - x_{i-1})(f(x_i^*) - (x_i - x_{i-1})f(w_i^*)| = (x_i - x_{i-1})|f(x_i^*) - f(w_i^*)|.$$

Since the first factor is at most δ and the second is at most $2M$, part a) follows. Similarly in part b), the difference of the two Riemann sums comes out to $|(x_i - x_{i-1})f(x_i^*) - (c - x_{i-1})f(y_i^*) - (x_i - c)f(z_i^*)|$, which similar reasoning shows to be at most $3M\delta$. The result follows.

Now we are ready to show:

Proposition 9. *Let $a < c < b$ and $f : [a,b] \to \mathbb{R}$. Then f is Riemann integrable on $[a,b]$ if and only if it is Riemann integrable on $[a,c]$ and on $[c,b]$, and if these conditions hold then $R\int_a^b f = R\int_a^c f + R\int_c^b f$.*

Remark: The notation for a tagged partition gets tedious. If in a proof, the choice of tagging is not important, we will allow ourselves to write P^* as an abbreviation for (P, τ).

Proof: Assuming that f is Riemann integrable on each of $[a, b]$, $[a, c]$ and $[c, b]$, the second statement follows easily. Indeed, we can take a sequence of partitions P_n of $[a, b]$ with mesh approaching zero and each including c and arbitrary tags τ. Then we have

$$R\int_a^b f = \lim_{n\to\infty}\sum_{i=1}^{n} f(x_i^*)\Delta_i = \lim_{n\to\infty}\left(\sum_{i=1}^{i_0} f(x_i^*)\Delta_i + \sum_{i=i_0}^{n} f(x_i^*)\Delta_i\right) =$$

$$R\int_a^c f + R\int_c^b f.$$

Now assume f is Riemann integrable on $[a, c]$ and on $[c, b]$. Using the same argument as above, we get the convergence to $R\int_a^c f + R\int_c^b f$ of any sequence of Riemann sums of mesh approaching zero which includes c as an element of the partition. In general, given any tagged partition P^* with mesh less than δ, an application of the It Doesn't Matter Lemma allows us to put an extra point c into the partition, at a cost of changing the Riemann sum by at most $3M\delta$. Thus as $\delta \to 0$, Riemann sums over paritititions which did not include c get arbitrarily close to Riemann sums over partitions including c, so f is Riemann integrable on $[a, b]$.

The final, and trickiest, step is to show that if f is Riemann integrable on $[a, b]$ it is Riemann integrable on both subintervals.[1] Again we will argue by contraposition; namely we shall assume that f is not Riemann-integrable on both $[a, c]$ and $[b, c]$ and show that it is not Riemann-integrable on $[a, b]$. Without loss of generality, we will assume that it is not Riemann-integrable on $[a, c]$. From the Cauchy criterion, this means that there exists $\epsilon > 0$ such that for all $\delta > 0$, there exist two tagged partitions P^*, Q^* of $[a, c]$ each of mesh less than δ such that $|R(f, P^*) - R(f, Q^*)| \geq \epsilon$. The idea is to extend P^* and Q^* to tagged partitions of the entire interval $[a, b]$ in such a way so as to preserve the discrepancy between the values of the two Riemann sums. But this can be done by choosing any tagged partition S^* of $[c, b]$ of mesh less than δ and then extending both P^* and Q^* to tagged partitions of $[a, b]$ by adding S^*: formally, take $A^* := P^* \cup S^*$ and $B^* := Q^* \cup S^*$. Because of this choice, we have $|R(f, A^*) - R(f, B^*)| = |R(f, P^*) - R(f, Q^*)| \geq \epsilon$. This can be done for any $\delta > 0$, and by the Cauchy criterion again, we conclude that f is not Riemann-integrable on $[a, b]$.

The following result, which is certainly a theorem in its own right, establishes the property (I1) for the Riemann integral and completes the proof of Theorem 3.

Theorem 10. *Every continuous function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.*

Proof: Since $[a, b]$ is a closed interval, f is uniformly continuous. For each $\epsilon > 0$, fix a $\delta(\epsilon)$ such that $|x - x'| < \delta$ implies $|f(x) - f(x')| < \epsilon$. For any partition P of $[a, b]$, let $x_i^m and x_i^M \in [x_{i-1}, x_i]$ be such that $f(x_i^m)$ and $f(x_i^M)$ are respectively the minimum and maximum values of f on $[x_{i-1}, x_i]$. These give two taggings of P; we

[1] Consider that if $\sum_n a_n + b_n$ converges, it does not follow that $\sum_n a_n$ and $\sum_n b_n$ converge.

abbreviate the respective Riemann sums by $L(f,P)$ and $U(f,P)$. Note that they have the property that for any tagging τ of P,

$$L(f,P) \leq R(f,P,\tau) \leq U(f,P).$$

Step 2: If (P,τ) is a tagged partition with $||P|| < \delta$, then

$$U(f,P) - L(f,P) = \sum_{i=1}^{n} (f(x_i^M) - f(x_i^m))(x_i - x_{i-1}) \leq \sum_{i=1}^{n} \epsilon(x_i - x_{i-1}) = (b-a)\epsilon.$$

Let (P_n, τ_n) be a sequence of partitions such that $P_n \subset P_{n+1}$ and $||P_n|| \to 0$. (For instance, we could take P_n to be the uniform partition into 2^n subintervals.) We claim that $R(f, P_n, \tau_n)$ is a Cauchy sequence. Indeed, if $m, n \geq N$, then $R(f, P_n, \tau_n)$ and $R(f, P_m, \tau_m)$ both lie in the interval $[L(f, P_N), U(f, P_N)]$, whose length goes to zero with N. (Here we have used the fact that if P' is a refinement of P, $L(f, P') \geq L(f, P)$ and $U(f, P') \leq U(f, P)$, since, when we split a interval into subintervals, the maximum value on each subinterval is less than or equal to the maximum value on the entire interval; you are asked to show this in detail below.) Thus $R(f, P_n, \tau_n)$ converges to some number L, which we must show is the integral.

Step 3: Let (P,τ) be any tagged partition with $||P|| < \delta$. Choose P_n such that $||P_n|| < \delta$, and let Q be the common refinement of P and P_n. Give it any tagging τ_Q. Then

$$|R(f,P,\tau) - L| \leq |R(f,P,\tau) - R(f,Q,\tau_Q)| + |R(f,Q,\tau_Q) - L|.$$

It is easy to see that both terms approach zero as $n \to \infty$ and $\delta \to 0$. Indeed, $R(f,P,\tau)$ and $R(f,Q,\tau)$ both lie in $[L(f,P), U(f,P)]$, hence differ from each other by at most $(b-a)\epsilon$. Similarly, since Q is a refinement of P_n, $R(f,Q,\tau_Q)$ lies in the interval $[L(f,Q), U(f,Q)]$, which is contained in the interval $[L(f,P_n), U(f,P_n)]$, which by construction contains L and has length approaching zero. This completes the proof.

Theorem 11. *(Fundamental theorem of calculus for the Riemann integral) Let* $f : [a,b] \to \mathbb{R}$ *be a Riemann integrable function. Define* $F(x) := R\int_a^x f$.
a) F *is continuous on* $[a,b]$*, and at every point* x_0 *at which* f *is continuous,* F *is differentiable with* $F'(x_0) = f(x_0)$.
b) Let $f : [a,b] \to \mathbb{R}$ *be a differentiable function. Suppose that* f' *is Riemann-integrable. Then* $R\int_a^b f' = f(b) - f(a)$.

Remark: Part a) has already been observed; it is a consequence of Theorem 3 and the abstract fundamental theorem. If f' is itself continuous, then part b) follows automatically (Corollary 2). The new point here – a subtle one, to be sure – is that there exist derivatives f' which are not continuous. Nevertheless f' might be Riemann-integrable, and if so the Fundamental Theorem still holds.

Proof of part b): Let P be any partition of $[a,b]$. We claim that there is a tagging τ such that $R(f,P,\tau) = G(b) - G(a)$. Indeed, we apply the Mean Value Theorem to G on each subinterval $[x_{i-1}, x_i]$, obtaining a point x_i^* such that

$$G(x_i) - G(x_{i-1}) = (x_i - x_{i-1})G'(x_i^*) = (x_i - x_{i-1})f(x_i^*).$$

These x_i^*'s form a tagging of P, and summing both sides from $i = 1$ to n, we get $G(b) - G(a) = G(x_n) - G(x_0) = R(f, P, \tau)$! Now taking a sequence of partitions with mesh approaching zero, the only possible limit of the Riemann sums is $G(b) - G(a)$.

Example: Consider $f(x) = x$ on $[0, 1]$. Since f is continuous, by Theorem 1 it is Riemann integrable. Since the integral of an integrable function can be computed via any sequence of Riemann sums with mesh approaching zero, we'll take, a la freshman calculus, uniform spacing and right endpoint sums. We get then

$$\int_0^1 f = \lim_{n\to\infty} \sum_{i=1}^{n} (\frac{i}{n})\frac{1}{n} =$$

$$\lim_{n\to\infty} (\frac{1}{n})^2 \sum_{i=1}^{n} i = \lim_{n\to\infty} \frac{n^2+n}{2n^2} = \frac{1}{2}.$$

On the other hand, since we happen to know that $F(x) = \frac{1}{2}x^2$ is an antiderivative for f, we can (of course!) use Corollary 2 to obtain $R\int_0^1 f = F(1) - F(0) = \frac{1}{2}$.

Exercise 47: Let $f : [a, b] \to \mathbb{R}$. Suppose that there exists a sequence of tagged partitions (P_n, τ_n) with $||P_n|| \to 0$ and such that $R(f, P_n, \tau_n) \to L$. Must f be Riemann integrable? (Hint: Think about Exercise 41.)

Because of Exercise 41, we know that not every bounded function is Riemann integrable. However, there are certainly some discontinuous Riemann integrable functions:

Proposition 12. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. If f is Riemann integrable on $[a, c]$ for all $c < b$, then f is Riemann integrable on $[a, b]$ and $R\int_a^b f = \lim_{c\to b^-} R\int_a^c f$.*

Proof: Fix $\delta > 0$. Suppose $|f| \le M$ on $[a, b]$. Let P^*, Q^* be two tagged partitions of mesh less than δ, and consider the difference $|R(f, P^*) - R(f, Q^*)|$. We want both partitions to have a common point c which is very close to the right endpoint, so we will choose a point c which is sufficiently close to b so as to lie in the interior of the last subinterval of both partitions; note this guarantees $b - c < \delta$. Let A^* be a tagged partition that results from adding the point c to P^* (and minimally changing the tags, as in the It Doesn't Matter Lemma); similarly, let B^* be a tagged partition that results from adding the point c to Q^*. Two applications of part b) of the Lemma gives $|(R(f, A^*) - R(f, B^*)) - (R(f, P^*) - R(f, Q^*)| \le 6M\delta$. Now, discarding the last subinterval of A^* and B^* we get tagged partitions C^* and D^* for f on $[a, c]$; since f is assumed Riemann-integrable on every subinterval, there exists some choice of δ such that $|R(f, C^*) - R(f, D^*)| < \epsilon$. But by reasoning very similar to part a) of the It Doesn't Matter Lemma, this differs from the difference between $R(f, A^*)$ and $R(f, B^*)$ by at most $2M\delta$, so overall we have

$$|R(f, P^*) - R(f, Q^*)| < \epsilon + 8M\delta,$$

a quantity which goes to zero with δ. Hence f is Riemann-integrable on $[a, b]$. Moreover, the argument shows that the difference between $\int_a^b f$ and $\int_a^c f$ goes to zero as $b - c \to 0$, which means precisely that $\lim_{c\to b} \int_a^c f = \int_a^b f$.

Remark: Later we will interpret this result as saying that any improperly integrable *bounded* function on a finite interval is already Riemann-integrable.

Corollary 13. *Let $f : [a, b] \to \mathbb{R}$ be bounded with finitely many points of discontinuity. Then f is Riemann integrable.*

Exercise 48: Prove Corollary 12.

4. A pause for reflection

We've done quite a bit of hard work already, so it is probably a good idea to stop and recall what we have shown and where we are going next. Our original goal was to show that the "abstract fundamental theorem of calculus" was non-vacuous by showing that there does exist an integral defined at least on the collection of all continuous functions. We did this: by showing that every continuous function is Riemann-integrable, we showed in particular that every continuous function is the derivative of some other function. Moreover, the class of Riemann-integrable functions includes bounded functions with only finitely many discontinuities, so we have made good on our promise to show that all the functions that are encountered in calculus are Riemann-integrable.

In calculus class we would now do two things: first, we would show how many practical problems can be solved by computing limits of Riemann sums, i.e., finding integrals of functions. Second, since it is very tedious to compute integrals from the limiting definition, we would develop a range of antidifferentiation techniques that will allow us to antidifferentiate at least some of the familiar functions. (But not all: there exist functions, like e^{x^2}, which we showed to have antiderivatives, but for which it can be proved that the antiderivative does not live in the class of functions generated by rational functions, sines, cosines, exponentials and logarithms by the familiar operations. It is far from clear how one could show this – really this is its own branch of mathematics, **differential Galois theory**, closer to algebra than analysis – but it is true. Later we will see why our inability to write down antiderivatives of such functions does not trouble us much.)

However, since we have all had calculus class already, this is not the direction in which we are going to go. We are studying integration *theory.* And there are still some theoretical questions that remain, the foremost one being: what can we say about the class $\mathcal{R}[a, b]$ of Riemann-integrable functions: it is clearly larger than the class of continuous functions and smaller than the class of all bounded functions. We would like it at least to be a robust class: for instance, one thing that we have *not* showed yet is that the product of Riemann-integrable functions is integrable (if you think about the fact that this was not true for series without some restrictions, you will appreciate that this is not likely to be so easy). In fact, a major goal will be to show that all monotone functions are Riemann-integrable, even those with infinitely many discontinuities (cf. Exercise 47).

As we have already seen, the definition of Riemann-integrability is stringent and complicated enough so that working directly from it is rather difficult. It turns out that there is a slightly more recent alternative definition of integrability, due to Darboux, which is easier to check than Riemann-integrability. In the next section we will start from scratch with the Darboux integral, which we will check satsifies are axioms for an integral. From this it follows immediately that the Darboux integral agrees with the Riemann integral on all continuous functions. However, it will be much easier to show that monotone functions are Darboux-integrable. We will then have two different integrals with advantages and disadvantages to each, and we will show that we get the best of both worlds because the class of Darboux-integrable functions is exactly the same as the class of Riemann-integrable functions. Finally, we will give a necessary and sufficient condition for a function to be (either Riemann or Darboux) integrable, due to Lebesgue. The proof of this theorem belongs in a more advanced course, but it is nevertheless enlightening to see the answer.

Exercise 49*: Let $X = (x_n)_{n=1}^{\infty}$ be any countably infinite subset of $[0,1]$. For $x \in [0,1]$, let I_x be the subset of the positive integers consisting of all n such that $x_n \leq x$. Consider the function f whose value at a point $x \in [0,1]$ is $\sum_{n \in I_x} 2^{-n}$. Show that f is an increasing function whose set of discontinuities is precisely X.

For instance, we could take X to be the rational numbers between 0 and 1 (together with some fixed bijection to the positive integers) and the corresponding function is continuous at every irrational number and discontinuous at every rational number. We will show that this function is Riemann-integrable!