MA 8, WEEK 6: INTEGRATION, C^k-FUNCTIONS, NEWTON'S METHOD

1. Integration

Definition 1. A function f is **integrable** on the interval $[a, b]$ if and only if the following holds:

- For any $\epsilon > 0$,
- there is a partition $a = t_1 < t_2 < \ldots < t_{n-1} < t_n = b$ of the interval $[a, b]$ such that

$$\left(\sum_{i=1}^{n} \sup_{x\in(t_i,t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i) - \sum_{i=1}^{n} \inf_{x\in(t_i,t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i)\right) < \epsilon$$

One way to interpret the sums above is through the following picture:

[picture will be included later]

Specifically,

- think of the $(\sum \inf)$-sum as the area of the blue rectangles in the picture below, and
- think of the $(\sum \sup)$-sum as the area of the red rectangles in the picture below.
- Then, the difference of these two sums can be thought of as the area of the gray-shaded rectangles in the picture above.
- Thus, we're saying that a function $f(x)$ is **integrable** iff we can find collections of red rectangles – an "upper limit" on the area under the curve of $f(x)$ – and collections of blue rectangles – a "lower limit" on the area under the curve of $f(x)$ – such that the area of these upper and lower approximations are arbitrarily close to each other.

Note that the above condition is equivalent to the following claim: if $f(x)$ is integrable, we can find a sequence of partitions $\{P_n\}$ such that "the area of the gray rectangles with respect to the P_n partitions goes to 0" – i.e. a sequence of partitions $\{P_n\}$ such that

$$\lim_{n\to\infty} \left(\sum_{i=1}^{n} \sup_{x\in(t_i,t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i) - \sum_{i=1}^{n} \inf_{x\in(t_i,t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i)\right) = 0.$$

In other words, there's a series of partitions P_n such that these upper and lower sums both converge to the same value: i.e. a collection of partitions P_n such that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \sup_{x\in(t_i,t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \inf_{x\in(t_i,t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i).$$

Date: Nov 5, 2013.

If this happens, then we define

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^{n} \sup_{x\in(t_i,t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \inf_{x\in(t_i,t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i),$$

and say that this quantity is the **integral** of $f(x)$ on the interval $[a, b]$. For convenience's sake, denote the upper sums of $f(x)$ over a partition P as $U(f(x), P)$, and the lower sums as $L(f(x), P)$.

This discussion, hopefully, motivates why we often say that the integral of some function $f(x)$ is just "the area under the curve" of $f(x)$. Pictorially, we are saying that a function is integrable if and only if we can come up with a well-defined notion of area for this function; in other words, if sufficiently fine upper bounds for the area beneath the curve (the ($\sum$ sup)-sums) are arbitrarily close to sufficiently fine lower bounds for the area beneath the curve (the ($\sum$ inf)-sums.)

The definition of the integral, sadly, is a tricky one to work with: the sups and infs and sums over partitions amount to a ton of notation, and it's easy to get lost in the symbols and have no idea what you're actually manipulating. If you ever find yourself feeling confused in this way, just remember the picture above! Basically, there are three things to internalize about this definition:

- the area of the **red** rectangles corresponds to the upper-bound ($\sum$ sup)-sums,
- the area of the **blue** rectangles corresponds to the lower-bound ($\sum$ inf)-sums, and
- if these two sums can be made to be arbitrarily close to each other – i.e. the area of the gray rectangles can be made arbitrarily small – then we have a "good" idea of what the area under the curve is, and can say that $\int_a^b f(x)$ is just the limit of the area of those red rectangles under increasingly smaller partitions (which is also the limit of the area of the blue rectangles.)

The integral is a difficult thing to work with using just the definition: later on, we'll develop lots of tools to help us actually do nontrivial things with the integral. To illustrate how working with the definition goes, though, let's work two examples:

1.1. **Calculating the Integral.**

Example 2. The integral of any constant function $f(x) = C$ from a to b exists; furthermore,

$$\int_a^b c dx = C \cdot (b - a).$$

Proof. Pick any constant function $f(x) = C$. To use our definition of the integral, we need to find a sequence of partitions P_n such that $\lim_{n\to\infty} U(f(x), P_n) - L(f(x), P_n)$ goes to 0. How can we do this? Well: what kinds of partitions of $[a, b]$ into n parts even exist?

One partition that often comes in handy is the uniform partition, where we break $[a, b]$ into n pieces all of the same length: i.e. the partition

$$P_n = \left\{ a, a + \frac{b-a}{n}, a + 2\frac{b-a}{n}, a + 3\frac{b-a}{n}, \ldots a + n\frac{b-a}{n} = b. \right\}$$

In almost any situation where you need a partition, this will work excellently! In particular, one advantage of this partition is that the lengths $(t_{i+1}-t_i)$ in the upper and lower sums are all the same: they're specifically $\frac{b-a}{n}$.

Let's see what this partition does for our integral. If we look at $U(f(x), P_n)$, where P_n is the uniform partition defined above, we have

$$\begin{aligned}
U(f(x), P_n) &= \sum_{i=0}^{n-1} \left(\sup_{x \in \left(a+i\frac{b-a}{n}, a+(i+1)\frac{b-a}{n}\right)} f(x) \right) \cdot \left(a + (i+1)\frac{b-a}{n} - a - i\frac{b-a}{n} \right) \\
&= \sum_{i=0}^{n-1} \left(\sup_{x \in \left(a+i\frac{b-a}{n}, a+(i+1)\frac{b-a}{n}\right)} f(x) \right) \cdot \left(\frac{b-a}{n} \right) \\
&= \sum_{i=0}^{n-1} C \cdot \left(\frac{b-a}{n} \right), \text{ because } f(x) \text{ is a constant function,} \\
&= C\frac{b-a}{n} + C\frac{b-a}{n} + \ldots + C\frac{b-a}{n} \\
&= C \cdot (b-a).
\end{aligned}$$

Similarly, if we look at $L(f(x), P_n)$, we have

$$\begin{aligned}
L(f(x), P_n) &= \sum_{i=0}^{n-1} \left(\inf_{x \in \left(a+i\frac{b-a}{n}, a+(i+1)\frac{b-a}{n}\right)} f(x) \right) \cdot \left(a + (i+1)\frac{b-a}{n} - a - i\frac{b-a}{n} \right) \\
&= \sum_{i=0}^{n-1} \left(\inf_{x \in \left(a+i\frac{b-a}{n}, a+(i+1)\frac{b-a}{n}\right)} f(x) \right) \cdot \left(\frac{b-a}{n} \right) \\
&= \sum_{i=0}^{n-1} C \cdot \left(\frac{b-a}{n} \right), \text{ because } f(x) \text{ is a constant function,} \\
&= C\frac{b-a}{n} + C\frac{b-a}{n} + \ldots + C\frac{b-a}{n} \\
&= C \cdot (b-a).
\end{aligned}$$

Therefore, we have that the limit

$$\lim_{n\to\infty} U(f(x), P_n) - L(f(x), P_n) = \lim_{n\to\infty} C \cdot (b-a) - C \cdot (b-a) = \lim_{n\to\infty} 0 = 0,$$

and consequently our integral exists and is equal to

$$\lim_{n\to\infty} U(f(x), P_n) = C \cdot (b-a).$$

□

2. Examples of Non-Integrable Functions

There are many functions that are not Riemann integrable. In fact, it turns out that "most" functions $f : \mathbf{R} \to \mathbf{R}$ are not Riemann integrable, for any reasonable way to measure elements of the set of functions from $\mathbf{R}$ to itself.

The simplest examples of non-integrable functions are, say,

$$\frac{1}{x}$$

defined on the an interval $[0, a]$ for some $a \in \mathbf{R}^+$ and

$$\frac{1}{x^2}$$

in an interval containing 0. Intuitively, both are not integrable, because the "area" that the integral represents is infinite.

Another way the integrability can fail is if the integrand "jumps around" too much. For instance, an example given in class is

$$f(x) = \begin{cases} 1, & x \in \mathbf{Q} \\ 0, & x \notin \mathbf{Q}. \end{cases}$$

The reason that this is not Riemann integrable is because of the following strange factor: a single slice in the Riemann sum can be either its width or 0, depending on whether we evaluate it on a rational x or an irrational x on that interval. Therefore, no matter how we choose the intervals, we can have a Riemann integral that takes any value in $[0, b - a]$, where we are integrating f on the interval $[a, b]$.

Remark 1. Using a different form of integration (Lebesgue), that was alluded to in class, you *can* integrate this function. The argument using Lebesgue integration essentially says that since there are so many more irrational points than rational ones, you can ignore the values of the function at the rational points, so that the integral is 0.

This is beyond what we'll cover in Ma 1, and you won't need to know this for this class, but it's something that is crucial to modern mathematics and applications, especially in probability theory and anything that uses it (which is pretty much every scientific discipline these days).

3. C^k functions

If k is nonnegative integer, a function $f : \mathbf{R} \to \mathbf{R}$ is said to be a C^k **function**, denoted $f \in C^k(\mathbf{R}) = C^k(\mathbf{R}, \mathbf{R})$ if f has k continuous derivatives.

For instance, $C^0(\mathbf{R})$ is just the space of continuous functions, and C^1 is the space of continuously differentiable functions, and C^2 is the space of functions with continuous first and second derivatives. Hence, C^k functions can be derived as functions whose first through kth derivatives are continuous. Obviously, if $\ell > k$, any C^ℓ-function is also a C^k-function.

We also commonly let $C^\infty(\mathbf{R})$ denote the "smooth functions," that is, which consists of functions for which *all* orders of derivatives are continuous. Clearly these are contained in any $C^k(\mathbf{R})$.

Remark 1. It is customary, especially in physics or mathematical physics, to says that smooth functions just have continuous derivatives up to some desired order over the domain of definition. However, we will be more careful with language and notation in this class.

Intuitively, the graphs of the functions get "smoother" as k moves from 0 to ∞. Examples of C^k-functions that are not C^{k+1} functions are

$$f(x) = |x|^{k+1}$$

for even k, and

$$g(x) = x^{k+1} \sin\left(\frac{1}{x}\right).$$

Both do not have a $k+1$-st derivative at 0.

We saw above that things can get pretty bad with respect to integration. But can they be bad while staying relatively nice? We know that differentiability implies continuity, and we'll later learn about the relation between differentiability and integrability (which you've used for nice functions in your previous calculus classes). But is there a function that is continuous without being differentiable? Surprisingly, the answer turns out to be yes!

Example 2. This famous (class of) examples is due to Weierstrass. Before this example, it was widely believed that every continuous function was differentiable except on a set of isolated points.

For instance, a simple example of such a function is

$$f(x) = \sum_{n=0}^{\infty} \frac{\sin(2^n x)}{2^n} = \sin x + \frac{1}{2}\sin(2x) + \frac{1}{4}\sin(4x) + \cdots.$$

This series converges for all x, because it's absolutely convergent. It turns out that this is continuous for all x, but not differentiable at any x.

This function is 2π-periodic, that is, $f(x) = f(x+2\pi)$, and it looks very "bumpy" and so, you might guess it's not differentiable everywhere. However, you can prove this rigorously with some work.

There are many interesting things about such functions, but here are two in particular. First, as I indicated in the beginning, it turns out that "most" functions from $\mathbf{R}$ to $\mathbf{R}$ are some kind of Weierstrass function; and for functions from $\mathbf{R}^n \to \mathbf{R}^n$, tend to look like some kind of "Brownian motion" as you travel along a path. This is an example of the same phenomena that you experience with real numbers. "Most" real numbers are not rational, but almost all of the numbers you've seen in your life are rational.

Secondly, the Weierstrass function is an example of one of the first "fractals" studied, although this term wasn't in common usage until Mandelbrot popularized it in the late 20th century. Weierstrass functions have "detail at every level," in that if you keep on zooming in onto the further, it does not ever get closer to being a straight line. More precisely, between any two points, the function is never monotone. This roughness, as Mandelbrot and others noticed, is a phenomenon that seems to be ubiquitous in nature, and it seems to be this kind of self-similarity that makes objects look "natural" (as opposed to artificial).

3.1. **Differentiability and Integrability Do Not Imply Each Other.** This is an important point, but is something obscured in the slew of definitions. There are differentiable functions that are not integrable, and there are integrable but non-differentiable functions.

For instance,

$$f(x) = \frac{1}{x}$$

is differentiable in (0,1), but it's not integrable in this interval. Why? We have

$$\int_0^1 f(x)\,dx = \int_0^1 \frac{1}{x}\,dx = \ln 1 - \ln 0$$

which is undefined.

Another more prosaic example is simply the function $f(x) = x$. This is certainly differentiable, but it is not an integrable function from $\mathbf{R}$ to $\mathbf{R}$ since its integral is infinite. However, it IS integral, on, say, an interval.

A Weierstrass function is an example of an integrable but non-differentiable function.

4. Newton's Method

Until now, we have covered several applications of derivatives including the Mean Value Theorem. These "applications" might not seem useful or *really* applicable to real world situations (even though they are!).

Newton's Method, however, will change that. Newton's method helps us approximate solutions to equations, which is very helpful when we cannot explicitly solve the equation.

So how does it work? Suppose we have a function $f(x)$ and we would like to find a solution for $f(x) = 0$. We make an initial guess, say x_0. If $f(x_0) = 0$, we have found a solution and we can stop here. If not, consider

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Looking closely at how x_1 is chose, we see that x_1 is in fact the x-intercept of the line $y = f(x_0) + f'(x_0)(x - x_0)$. But note that this is the equation of the tangent line to $f(x)$ at $f(x_0)$. Under certain conditions that we will see later, $f(x_1)$ is guaranteed to be closer to 0 than $f(x_0)$. In other words, x_1 is a better approximate of the solution to $f(x) = 0$ than x_0.

Continuing in the same way, we define for each n

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Again, under certain conditions, x_n will converge to a solution, x^*, of $f(x) = 0$.

Example 1. Let's see how we can use Newton's method to approximate $\sqrt{e}$. First, note that $\sqrt{e}$ is the root of $f(x) = x^2 - e$. Let's make our lives easier and compute x_n now: We have $f'(x) = 2x$ and thus,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - e}{2x_n} = \frac{x_n^2 + e}{2x_n}$$

Our initial guess is $x_0 = 1$ (which is obviously quite a terrible guess since $1^2 = 1$ is *clearly* not e). Using the recursive formula we obtained above, we get:

$$x_1 = \frac{1+e}{2} \approx 1.86 \quad ; \quad x_2 = \left(\left(\frac{1+e}{2}\right)^2 + e\right)/(1+e) \approx 1.66 \quad ; \quad x_3 = \cdots \approx 1.64876$$

The actual value of $\sqrt{e} \approx 1.64872$ so just after 3 iterations of Newton's method, we got a value that is right upto the fourth decimal place!

Even though Newton's method worked quite brilliantly in above example, Newton's method does not always work as we will see in our next example.

Example 2. Let's find the root of $f(x) = \sqrt[3]{x}$. For the moment, let's ignore the fact that the root of $f(x)$ is quite clearly $x = 0$.

Let $x_0 = 1$ be the initial guess. Again, we find the recursive formula for x_n.

$$f'(x) = \frac{1}{3}x^{-2/3} \quad \Rightarrow \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - 3\frac{x_n^{1/3}}{x_n^{-2/3}} = -2x_n$$

This means that when we set $x_0 = 1$ we get

$$x_1 = -2 \quad ; \quad x_2 = 4 \quad ; \quad x_3 = -8 \quad \cdots$$

In this example, instead of getting closer to the solution, Newton's method gives us a sequence of x_n's which get farther and farther away from the true solution $x = 0$ as n goes to infinity. So Newton's method fails quite spectacularly here.

Why did Newton's method work in the first example but fail in the second? More importantly, can we tell when Newton's method is going to work? Fortunately, the answer is yes.

Theorem 3. *Let I be an interval and f a function which is twice continuously differentiable on I. Suppose that for every $x \in I$, we have $|f''(x)| < M$ and $|f'(x)| > \frac{1}{K}$. Then if we pick $x_0 \in I$, and we define the sequence $\{x_n\}$ by*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

If $x_n \in I$ for each N and $|f(x_0)| < \frac{r}{K^2M}$, then we have

$$|f(x_n)| \leq \frac{r^{2^n}}{K^2M}$$

5. More on differentiability

5.1. **C^n functions and smooth functions.** Recall that a function $f(x)$ is called $C^n(I)$ if for all $k \leq n$, the k-th derivative of $f(x)$ is defined for all $x \in I$ and is continuous on I. If $f(x)$ is $C^n(I)$ for all n, then $f(x)$ is said to be *smooth* and is denoted $f(x) \in C^\infty(I)$.

Proposition 1. *Let I be an interval and let $f \in C^n(I)$ and $g \in C^m(I)$ where $n \leq m$. Then,*

- *$f \in C^k(I)$ for all $k \leq n$*
- *$f' \in C^{n-1}(I)$*
- *$\alpha f + \beta g \in C^n(I)$ for all $\alpha, \beta \in \mathbb{R}$*

Now, suppose $f, g \in C^\infty(I)$. Then,

- *$fg \in C^\infty(I)$*
- *The composition of f and g, $f \circ g$ is smooth*
- *Suppose further that $g(x) \neq 0$ for all $x \in I$. Then, $\frac{f}{g} \in C^\infty(I)$.*

Examples 2. Lots of the functions we are familiar with are smooth. For example,

- polynomials $f(x) = a_nx^n + \cdots + a_1x + a_0$ are in $C^\infty(\mathbb{R})$;
- trigonometric functions $\sin(x), \cos(x), \tan(x), \cdots \in C^\infty(\mathbb{R})$;
- the function $f(x) = e^x$ is smooth.

Example 3. You might be thinking, "well are there any continuous functions that is not smooth?" The answer is: of course there are! If all continuous functions were smooth, we wouldn't have defined what it means for a function to be C^n!

Consider $f(x) = |x|$. We already know that f is continuous, i.e. $f \in C^0(\mathbb{R})$. But note that the first derivative of f is *not* defined at $x = 0$.

$$\text{Left derivative: } f'(x) = -1 \quad ; \quad \text{Right derivative: } f'(x) = 1$$

Hence, $f(x) \notin C^1(\mathbb{R})$.

Example 4. Consider the function

$$f(x) = \begin{cases} x^2 & x \geq 0 \\ -x^2 & x < 0 \end{cases}$$

$f(x)$ is clearly continuous, and thus, $f \in C^0(\mathbb{R})$.

Is $f(x)$ C^1? Since $f(x)$ is defined to be $\pm x^2$ away from $x = 0$ and since we know polynomials are smooth, we only need to check what happens at $x = 0$. But the left derivative is $f'(x) = -2x \Rightarrow f'(x) = 0$ and the right derivative is $f'(x) = 2x \Rightarrow f'(0) = 0$. Therefore, the left and the right derivatives of f agree at $x = 0$ and thus, we conclude that f does have a continuous derivative which is defined on all of $\mathbb{R}$. Hence, $f \in C^1(\mathbb{R})$.

Is $f \in C^2$? We saw above that

$$f'(x) = \begin{cases} 2x & x \geq 0 \\ -2x & x < 0 \end{cases}$$

Therefore, we get

$$f''(x) = \begin{cases} 2 & x > 0 \\ -2 & x < 0 \end{cases}$$

Note that $f''(0)$ is not defined since the left and the right derivatives do not agree. Therefore, $f \notin C^2(\mathbb{R})$.

5.2. **Constructing smooth functions.** Sometimes, we want to construct functions that behave the way we want. For example, we might require that a function return 0 for all $x < 0$. Now, we will see how to construct functions that satisfy certain conditions AND are smooth.

Example 5. Let's construct a smooth function $f(x)$ such that $f(x) = 0$ for all $x \leq 0$ and $f(x) \geq 0$ for all $x \geq 0$.

We could try

$$f_1(x) = \begin{cases} 0 & x \leq 0 \\ x & x > 0 \end{cases} \quad \text{or} \quad f_2(x) = \begin{cases} 0 & x \leq 0 \\ x^2 & x > 0 \end{cases}$$

But neither f_1 and f_2 is smooth: the first derivatives for both functions are not defined at $x = 0$.

Instead, consider

$$f(x) = \begin{cases} 0 & x \leq 0 \\ e^{-\frac{1}{x}} & x > 0 \end{cases}$$

Note that since $\lim_{x \to 0} e^{-\frac{1}{x}} = 0$, $f(x)$ is continuous.

Moreover,

$$f'(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{x^2} e^{-\frac{1}{x}} & x > 0 \end{cases}$$

We recall that $e^{-\frac{1}{x}} = o(x^n)$ for all n. Therefore, $f'(0) = 0$ and we see that $f \in C^1(\mathbb{R})$.

Since $e^{-\frac{1}{x}} = o(x^n)$ for all n, all the higher order derivatives at $x = 0$ exist and thus, $f(x) \in C^\infty(\mathbb{R})$.

Note also that $f(x) = 0$ for all $x \leq 0$ and $f(x) \geq 0$ for all $x \geq 0$ as desired.

Example 6. Now, suppose we want to construct a function $g(x)$ such that:

- $g(x)$ is smooth
- $g(x) = 0$ for all $x \leq 0$
- $g(x) = 1$ for all $x \geq 1$
- $0 < g(x) < 1$ for all $x \in (0, 1)$

Consider the function

$$g(x) = \frac{f(x)}{f(x) + f(1-x)}$$

We claim that $g(x)$ satisfies all the properties listed above.

First, note that $f(x) + f(1-x) \neq 0$ for all x: Suppose $f(x) + f(1-x) = 0$. Since $f(x), f(1-x) \geq 0$ for all x, this implies that $f(x) = 0$ and $f(1-x) = 0$. However, $f(x) = 0$ means that $x \leq 0$ whereas $f(1-x) = 0$ implies $x \geq 1$. This is a contradiction and thus, $f(x) + f(1-x) \neq 0$ for all x.

Moreover, letting $h(x) = 1 - x$, we see that $f(1-x) = f(h(x))$ is smooth since composition of smooth functions is smooth.

We see that $g(x)$ is a ratio of smooth functions where the denominator is never zero. Hence, $g(x) \in C^\infty(\mathbb{R})$.

Let $x \leq 0$. Then,

$$g(x) = \frac{f(x)}{f(x) + f(1-x)} = \frac{0}{0 + f(1-x)} = 0$$

Let $x \geq 1$. Then,

$$g(x) = \frac{f(x)}{f(x) + f(1-x)} = \frac{f(x)}{f(x) + 0} = \frac{f(x)}{f(x)} = 1$$

Let $x \in (0, 1)$. Since $f(x) \geq 0$ for all x and $f(x) + f(1-x) > 0$ for all x, we conclude that $g(x) \geq 0$ for all x and when $f(x) \neq 0$, $g(x) > 0$. In particular, if $x \in (0, 1)$, $f(x) \neq 0$ and thus, $g(x) > 0$. We also know that since $f(1-x) \geq 0$, $f(x) \leq f(x) + f(1-x)$. Moreover, when $f(1-x) \neq 0$, we get $f(x) < f(x) + f(1-x)$. Again, for $x \in (0, 1)$, we get $g(x) = \frac{f(x)}{f(x)+f(1-x)} < 1$.

Therefore, we have constructed $g(x)$ which satisfies all the desired properties.