MATH 2310H: Integrability of non negative functions

In all advanced mathematics courses we will be concerned with two separate questions in attacking any problem:
One: Does the problem have an answer? We say it has an answer if we can describe the answer in words, even if we do not know how to compute it numerically.
Two: How can we actually compute, or approximate, the answer as a number, for example as a decimal?

In elementary mathematics courses the first question is often passed over in silence, we are simply given a formula for finding the answer to our problem, and our job is only to practice using this formula over and over in different situations. We may be told the area of a circle of radius r is πr^2, and then asked to compute this for various values of r.

This course will be different, since we will be expected to understand how the method for solving our problem is developed, and understand also how to apply it in particular cases. For example before we are through you should understand what it means to say the area of a circle is πr^2 and also be able to convince someone that this formula is correct.

In 2210 we have been concerned with the basic question: "What is the area under the graph of f?", where $f \geq 0$ is a non negative real valued function defined on an interval [a,b]. It is in fact not at all clear that this question has an answer for every such function. Indeed there are functions for which our notion of area under the graph does not make sense. So the first question we have to ask ourselves about any function is "does this function <u>have</u> an area under its graph?". This might seem silly if you have only looked at very simple graphs, but I assure you there are some horrible looking graphs under which there is no area, at least the elementary Riemann integral notion given in this course does not work for them.

Of course to answer the question of whether a function has area under its graph, we have to really know what area means, so we can test each function out. There are many good ideas that were suggested by class members as to how to approach area, involving rectangles, trapezoids, triangles, but the simplest one, the one used by Riemann, involves only rectangles, and we will give that version here. The basic idea was this:

<u>First</u>: we will assume we know what the area of a rectangle is, namely the length of the base multiplied by the height.

<u>Second</u>: if the area under a graph is going to make sense, then we ought to be able to approximate it by sums of areas of rectangles, both from below and from above, as closely as we like.

<u>Third:</u> to describe the area in words when it exists, we will say that if the lower and upper approximations do get closer together, then the unique number in between that they are both getting closer to, i.e. their common limit, is going to be the area under the graph.

This also tells us when area does not make sense, i.e. if the upper sums and the lower sums <u>cannot</u> be made to get as close together as we want, i.e. if all upper and lower sums stay further apart than some fixed positive distance, then we will say that the area under that graph does not exist.

Fortunately it is rather hard to cook up a function whose area does not exist. The area will exist for virtually all functions you can think of or write down. So deciding whether there is an area will not be so tough, the answer is almost always yes in practice. Figuring out what the area is will be much harder. Fortunately there are many functions whose areas can be found very easily using the fundamental theorem of calculus.

We can write down precise formulas for the areas under graphs of all polynomials, basic trig functions, and the log and exponential functions. At least we can write down formulas for these areas involving those same functions, so if you know how to evaluate those functions then you can evaluate these areas precisely. For poynomials it is not a problem since it is so easy to compute with polynomials.

For trig and log, exponential functions of course we still haver the problem of how to compute their values which usually requires some approximation procedure itself. Still this is a great advance over the terrific labor that was needed by geniuses like Archimedes and Galileo, to find even simple areas, such as that under a parabola, before calculus was invented. But there are many, many, more functions whose area, although it exists, still cannot be precisely computed even using calculus.

This happens as soon as we start to combine even the familiar trig and polynomial functions, for example there is no simple formula for the area under the graph of $\exp(x^2)$, at least not one which involves only trig, log, exponential, and polynomial functions. I.e. for most functions we will be able to say yes the area exists, and we can describe it in words as the common limit of the upper and lower sums, but we still cannot compute precisely what specific number the area is equal to.

Nonetheless, we will virtually always be able to approximate any area as closely as we like, by taking an appropriate upper sum or lower sum. This is almost the only place where calculators are really useful in this subject. It is harmful and pointless to have a calculator do our thinking for us by finding the derivative or antiderivative of easy functions like $\sin(x)$ or $\cos(x)$, or x^5, but it is really essential to use a calculator, or even a computer, to approximate the area under a curve using 1,000 subdivisions and hence add up the areas of 1,000 rectangles, or even 20 of them.

Actually things are a little better than this, since there is a “formula”: for the area under even $\exp(x^2)$, if you allow infinite formulas. These formulas of course cannot be evaluated fully, but can be used for approximations. This might sound bad, but to be completely honest, even evaluating a familiar function like $\sin(x)$ or e^x actually involves such infinite formulas. Really the only functions we can evaluate precisely are polynomial functions, and even then only for easy input values.

So the moral of the story is we must learn the actual definition of area in order to understand when we can compute areas the “easy way” using the fundamental theorem of calculus, and also how to approximate those areas when the easy way does not work for us.

To repeat: Our first goals are the following:
I) Learn the meaning of “area under a graph” according to Riemann.
II) Learn to recognize which functions have areas.
III) Learn how to calculate these areas the easy way, when possible, using the fundamental theorem of

calculus.
IV) Understand why the fundamental theorem of calculus is true, i.e. why the easy way works.
V) Learn what to do with functions which do have area, but on which the easy way for finding them does not work, i.e. learn how to approximate all areas, even hard ones.

Remark: All five goals above have practical uses, even step iv, of understanding why the FTC is true. If you understand why the FTC is true, you will then be able to apply the same idea to solve other problems, not necessarily involving area.

The following facts summarize what we have learned in class, and have explained to some extent.

Description of when the area under graph(f) makes sense.
Assume $f \geq 0$ is a bounded function with non negative real values on the interval [a,b], i.e. assume f does not get infinitely large on [a,b]. I.e. assume there is some big number B such that for every point x in [a,b], $f(x) \leq B$. This B is called a bound for the values of f on [a,b].

Define a partition of the interval [a,b] to be a finite sequence of points in [a,b], starting with a and ending with b, I.e. a partition of [a,b] into n subintervals, is a sequence of n+1 points $x_0 \leq x_1 \leq \ldots\ldots \leq x_{n-1} \leq x_n$, with $x_0 = a$, and $x_n = b$.

A partition of [a,b] divides up the interval [a,b] into n smaller intervals called subintervals, as follows: $[x_0, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$,......,$[x_{i-1}, x_i]$,......., $[x_{n-1}, x_n]$.

Given such a partition of [a,b], we construct an upper sum for f for this partition as follows: for each subinterval, choose an upper bound for f on that subinterval. I.e. for each index i between 1 and n, choose a number M_i such that $f(x) \leq M_i$ for every point x in $[x_{i-1}, x_i]$.
I.e. $f(x) \leq M_1$ for all x in $[x_0, x_1]$, and $f(x) \leq M_2$ for all x in $[x_1, x_2]$, etc...

Then form the corresponding upper sum from this partition and these upper bounds as follows: the upper sum is the sum of the areas of the rectangles with bases equal to the various subintervals $[x_{i-1}, x_i]$, and heights equal to the bounds M_i. Thus the upper sum U is the following:

$$M_1(x_1-x_0) + M_2(x_2-x_1) + M_3(x_3-x_2) + \ldots\ldots + M_n(x_n-x_{n-1}).$$

A very compact shorthand for this sum is: $\sum_{i=1}^{n} M_i(x_i - x_{i-1})$. Here you just write out one typical term $M_i(x_i-x_{i-1})$, and the symbol $\sum$ tells you to sum up a sequence of those, as i takes on all values from 1 to n. So the $\sum$ symbol and the 1 at the bottom and the n at the top tell you how many terms there are in your sum.

Now to form a lower sum for this same partition choose for each i, a lower bound for f on the subinterval $[x_{i-1}, x_i]$. I.e. let m_i be a number such that $0 \leq m_i \leq f(x)$ for every x in $[x_{i-1}, x_i]$. Then form the lower sum L =
$$m_1(x_1-x_0) + m_2(x_2-x_1) + m_3(x_3-x_2) + \ldots\ldots + m_n(x_n-x_{n-1}),$$

or in shorthand: $\sum_{i=1}^{n} m_i(x_i - x_{i-1})$. Then since the rectangle with base $[x_{i-1}, x_i]$. and height m_i lies entirely under that part of graph(f) over the subinterval $[x_{i-1}, x_i]$, the area of this rectangle should be ≤ the actual area under that part of the graph. Since this true for every subinterval, we should have $\sum_{i=1}^{n} m_i(x_i - x_{i-1}) \leq$ area under graph(f), provided the area under graph(f) makes sense. In the same way we should have (area under graph(f)) ≤ $\sum_{i=1}^{n} M_i(x_i - x_{i-1})$, if the area makes sense. Thus the actual area under the graph should be caught between all upper sums and all lower sums.

Definition: We say the area under the graph of f makes sense, where $f \geq 0$ is a bounded non negative function on [a,b], if there is only one number which lies between all upper and all lower sums. If this is so, we say that number is the area under the graph.

It is possible to see, using the least upper bound property of real numbers, that the following condition is equivalent to saying the area makes sense.

Test for existence of area: The area exists if and only if we can make the upper and lower sums arbitrarily close together.
More precisely, the area exists if the following is true: given any positive number e, we can make the upper and lower sums closer together than e. I.e. if for any $e > 0$, we can find a partition, $a = x_0 \leq x_1 \leq .....\leq x_{n-1} \leq x_n = b$, of [a,b], and a set of upper and lower bounds $\{M_i\}$, $\{m_i\}$, for f on that partition, such that $\sum_{i=1}^{n} M_i(x_i - x_{i-1}) - \sum_{i=1}^{n} m_i(x_i - x_{i-1}) < e$, then the area exists.

Description of the area in words: In case the area exists, we can describe it in words by saying it is the unique number which lies between all upper and lower sums for f on [a.b]. I.e. If the area under graph f on the interval [a,b] exists, it equals the unique number I such that whenever $\sum_{i=1}^{n} m_i(x_i - x_{i-1})$ is any lower sum and $\sum_{i=1}^{n} M_i(x_i - x_{i-1})$ is any upper sum for f on [a,b], then

$$\sum_{i=1}^{n} m_i(x_i - x_{i-1}) \leq I \leq \sum_{i=1}^{n} M_i(x_i - x_{i-1}).$$

Remark: This description does not exactly tell you how to find the exact area in a finite amount of time, but it does tell you how to approximate it. I.e choose any partition and corresponding upper and lower sums. Then the area, if it exists, is somewhere between the numbers $\sum_{i=1}^{n} m_i(x_i - x_{i-1})$ and $\sum_{i=1}^{n} M_i(x_i - x_{i-1})$. For example, if f(x) = sin(x), and the interval is $[0,\pi]$, then 0 is a lower bound for f on all of $[0,\pi]$ and 1 is an upper bound, so with the simple partition $x_0 = 0$ and $x_1 = \pi$, I can say already that, if it exists, the area under that part of the graph of sin(x) is somewhere between $0.\pi = 0$, and $1.\pi = \pi$. Not very accurate, but it didn`t take very long either. Of course we do not know yet whether this area actually exists so do not know whether the approximation makes sense. For this we need some criterion that guarantees the area makes sense in certain good cases that we can recognize.

Basic criterion for existence of area

We saw in class based on suggestions of class members, that any increasing non negative function, and by the same argument any decreasing non negative function has an area under its graph. I.e. it is always possible to make the upper and lower sums as close together as you want. In fact if f is increasing say on [a,b] and you want to make the upper and lower sums closer together say than 1/100, you do it as follows. Let B = f(b) - f(a), be the distance the graph rises from the beginning to the end of the interval [a,b]. Then take a partition of your interval into n subintervals so small that the product of B times the length (b-a)/n of one subinterval, is less than 1/100. That will do it. I.e. let n be so large a number that $B\{(b-a)/n\} < 1/100$. Then the amount of area between the closest possible upper and lower sums for that partition, can be stacked up into a rectangle with height B and base (b-a)/n. Thus the upper and lower sums are closer together than the area of that rectangle. Since by taking n large enough that area becomes as small as you want we are done, and the area exists. Now that still does not tell you how to compute it exactly, but it does tell you how to approximate it as closely as you want. For example using this method (and a $12. calculator), I got that the area under $\exp(x^2)$ between 0 and 1, is between 1.4 and 1.5. To sum up:

Theorem: If $f \geq 0$ is any weakly increasing or weakly decreasing function on the interval [a,b], then the area under graph(f) does exist.

We also saw that you can integrate a function by breaking it up into smaller intervals and working on each interval separately. I.e.

Theorem: If $f \geq 0$ is bounded and non negative on [a,b], and if we choose any points $a = c_0 \leq c_1 \leq .....\leq c_n = b$, (i.e. essentially a partition), then the area under graph(f) makes sense on the whole interval [a,b] if and only if it makes sense on all the subintervals $[c_{i-1},c_i]$, and the area on [a,b] is obtained by adding up the areas for all the subintervals $[c_0, c_1]$, $[c_1, c_2]$, ..., $[c_{n-1}, c_n]$.

Taken in combination, these two results yield:

Theorem: If f is piecewise monotone on [a.b], i.e. if there is a partition of [a,b] such that f is monotone on each subinterval separately, then the area of f makes sense on each subinterval and thus on the whole interval [a.b].

Since we know from differential calculus and our study of graphing that the graphs of all polynomials are always monotone except for a finite number of changes of direction, this says in particular.

Corollary: If f is any polynomial which is non negative on the interval [a.b], then the area under the graph of f exists.

The following principle is basic but is harder to prove than the one above, so we will just assume it: all continuous functions have area over any interval where they are non negative.

Theorem: If $f \geq 0$ on [a,b] and f is continuous on [a,b], then the area under graph(f) exists.

We also discussed the principle that there is zero area under a single point, and so we can change the value of f at one point, or at a finite number of points without changing the area. This gives us the following principle:

Theorem: If $f \geq 0$ is a bounded non negative function on [a,b], and if there is some partition of [a,b] such that f is continuous on each open subinterval (x_{i-1}, x_i), then the area under graph(f) exists, and it may be computed on each subinterval separately and the results added.

The most powerful way of finding areas is based on looking for an area function. I.e. assume the area under graph(f) exists between a and b. Then for every point x in [a,b], the area also exists between a and x. This defines an area function A(x) for x in [a,b], whose value is A(x) = the area under that part of the graph of f lying over the subinterval [a,x]. Then A(a) = 0, because there is no area between a and a, while the full area under graph(f) = A(b) = the area between a and b. This function is very important for the following reasons. If we can guess this function somehow from looking at f, we can find the area under graph(f). This function turns out not to be so hard to guess for the very simple reason that it is frequently an antiderivative of f, as we saw in several examples, but not in all. What we saw was the area function was an antiderivative at least in those cases where f was continuous. I.e. we have the following theorem called the <u>fundamental theorem of calculus</u>.

Theorem (FTC): Let $f \geq 0$ be a bounded function on [a,b] whose area exists, for example any piecewise monotone function. Then
(i) the area function define by A(x) = the area under that part of the graph of f lying over the subinterval [a,x], is a continuous function on [a,b] such that A(a) = 0, and A(b) = the area under graph(f).
(ii) A(x) is differentiable at those points x of [a,b] where f was continuous, and at such points, A'(x) = f(x).

We will discuss later why this is true. For now let us use it.
In particular, this theorem says that if f is continuous, then f always has an antiderivative, and the area function A is one. Moreover A is the unique antiderivative of f on [a,b] with value zero at a.
I.e. if f is continuous on [a,b] then its area function is the unique function A such that A'(x) = f(x) for all x in [a,b] and A(a) = 0.

This often makes it very easy to guess the area function A. For example, if $f(x) = x^4$, then the area function on the interval [1,2] is $A(x) = x^5/5 - 1/5$, where the 1/5 is put in to make A(1) = 0. Thus the area between 1 and 2 is A(2) = 32/5 - 1/5 = 31/5. Now that was easier than Archimedes method! The moral is the following: to compute area under graph(f) easily, try to guess the area function. If f is continuous on all of [a,b] just look for an antiderivative on all of [a,b] which has value zero at a. If f is not continuous on all of [a,b] try to break up the interval into subintervals such that f can be made continuous on each subinterval by changing only the values at the endpoints. Then use the antiderivative method on each subinterval separately and add the results.