REVIEW NOTES FOR FIRST MATH 1B EXAM

PETE L. CLARK

1. RIEMANN INTEGRALS

You should know the definition of the definite integral of a continuous function f on a closed interval $[a,b]$ as a limit of Riemann sums: a Riemann sum is obtained by dividing the interval $[a,b]$ into n subintervals of equal length $\Delta x = \frac{b-a}{n}$, by selecting a point x_i^* on the subinterval $[x_i, x_{i+1}]$ and approximating the value of the function on the entire subinterval by the value $f(x_i^*)$: thus we are approximating the area on $[x_i, x_{i+1}]$ by $\Delta x f(x_i^*) = \frac{b-a}{n} f(x_i^*)$; the entire Riemann sum is then $R(f,n,*) = \Sigma_{i=0}^{n-1} f(x_i^*) \frac{b-a}{n}$. Then, if f is continuous, no matter which sample points x_i^* are chosen, these Riemann sums will converge to a common value in the limit: $\lim_{n\to\infty} R(f,n,*) = \int_a^b f(x)dx$. This value is by definition the definite integral.

Although we probably had in mind a non-negative function $f(x)$ to start, the Riemann sum can be taken for a function which is both positive and negative; in this case $\int_a^b f(x)dx$ represents the *signed area*, i.e. the positive area - the negative area.

The fundamental theorem of calculus gives us a powerful method for computing these definite integrals (which would be anywhere from needlessly onerous to nigh impossible to compute directly from the definition): if $F(x)$ is an antiderivative of $f(x)$ – that is, a function such that $F' = f$ – then we know $\int_a^b f(x)dx = F(x)|_a^b = F(b) - F(a)$.

Note that an antiderivative of a continuous function on a closed interval $[a,b]$ is well-determined up to an additive constant: if F_1, F_2 are two functions such that $F_1' = f = F_2'$, then $(F_1 - F_2)' = (F_1' - F_2') = f - f = 0$, so $F_1 - F_2$ has zero derivative on $[a,b]$ hence is constant. It is important to remember this indeterminacy, e.g.

Question: A particle is moving along a line with the velocity function $v(t) = \sin t$. If its initial position is $x(0) = 0$, what is its position at any time t?

Solution: By definition $v(t) = x'(t)$, so that $x(t) = \int v(t) = -\cos t + C$. What is C? It is such that $0 = x(0) = -\cos(0) + C = -1 + C$, i.e. $C = 1$, and the position function is $x(t) = 1 - \cos(t)$. Notice that if we forgot to include the constant C our answer would have been wrong.

Another warning: many students become so enamored of the fundamental theorem of calculus that they forget all about the definition of the definite integral as the limit of Riemann sums. We instructors like to force you to remember the (geometric) definition of the definite integral by giving problems where the definite integral can be easily computed when finding the antiderivative is difficult or impossible. Our favorite example of this is the following

Fact: If $f(x)$ is an odd function: $f(-x) = -f(x)$, or (otherwise put) the graph of f is symmetric about the origin – then $\int_{-a}^a f(x)dx = 0$, since $\int_{-a}^0 f(x)dx = -\int_0^a f(x)dx$. For example, $\int_{-\pi}^{\pi} (\sin x)e^{x^2} dx$ would be a real pain to find the antiderivative (indeed I don't think I know how to do this using the methods of this course), but the definite integral is 0.

Similarly, you should have some idea of how to use Riemann sums to get an approximate value for $\int_a^b f(x)dx$ – although we would not ask you to numerically evaluate a Riemann sum on an exam, we might ask you to argue geometrically about the value of a definite integral using Riemann sums, for example

Question: Give a method for calculating $\int_0^1 e^{x^2} dx$ to within an accuracy of .000001.

Solution: First recall that e^{x^2} is our basic example of a function whose antiderivative we cannot write down, so we must proceed numerically. The idea of course is to divide into a bunch of subintervals and take a Riemann sum, but how will we know how far we are? We exploit the fact that $f(x) = e^{x^2}$ is an increasing function on $[0, 1]$ (check: $f'(x) = 2xe^{x^2} \geq 0$ on this interval), so that the left-endpoint sums will be lower-sums and the right-endpoint sums will be upper-sums (draw a picture!). yThat is $L_{eft}(f, n) := R(f, n, x_i^* = x_i) \leq \int_0^1 e^{x^2} dx \leq R_{ight}(f, n) := R(f, n, x_i^* = x_{i+1})$. Here $\Delta x = \frac{1}{n}$ and $x_i = \frac{i}{n}$, so the left endpoint sum is $\Sigma_{i=0}^{n-1} e^{(\frac{i}{n})^2} \frac{i}{n}$, whereas the right-endpoint sum is $\Sigma_{i=0}^{n-1} e^{(\frac{i+1}{n})^2} \frac{i}{n}$. We know the first quantity is always less than the second, so we get a calculational device (TI-85 or mathematical software package) to spit out these sums for increasing n until we find a value of n for which these two quantities differ by less than .000001.

2. GEOMETRIC APPLICATIONS OF INTEGRATION: AREA, VOLUME, ARCLENGTH

2.1. **Area between curves.** If $f(x)$ and $g(x)$ are two functions defined on $[a, b]$, we can, roughly speaking, integrate their difference to find the area between them. However, we must be careful about which function is on top: technically we define the area between f and g to be $\int_a^b |f(x) - g(x)| dx$, but what this really means is that we must break up the integral into regions where $f(x) - g(x)$ has constant sign – always positive or always negative. When $f(x) - g(x)$ is positive, $|f(x) - g(x)| = f(x) - g(x)$, so we integrate $f(x) - g(x)$. But when $f(x) - g(x)$ is negative, $|f(x) - g(x)| = g(x) - f(x)$, so we must integrate the latter.

Example: Find the area enclosed by $y = \sin x$ and $y = 0$ on $[0, 2\pi]$.

Solution: The answer is $\int_0^{2\pi} |\sin x| dx = \int_0^{\pi} \sin x dx + \int_{\pi}^{2\pi} - \sin x dx$ (because $\sin x < 0$ in the third and fourth quadrants).

2.2. **Volumes by slicing.** If we want to find the volume of a three-dimensional region whose cross-sectional areas are known to us, then we can simply integrate the cross-sectional area to get the volume:

$$V = \int_a^b A(x) dx$$

2.3. **Volumes of revolution.** If we are given a "region of revolution," i.e. a region obtained by revolving a region bounded by two curves about an axis parallel to the x or y axis, we have two methods for finding the volume of this region: the method of washers, and the method of shells.

Washers: If we are revolving around an axis parallel to the x-axis (aka $y = 0$) and we wish to integrate with respect to x, then we use the method of washers. Similarly, if we are revolving around an axis parallel to the y-axis and we wish to integrate with respect to y, then we use washers. Indeed washers are a special case of volumes by slicing: taking cross-sections perpendicular to the axis of integration, we will get disks (if there is only one curve) or the area between a larger disk and a smaller disk – a washer – if we are revolving the region between two curves. We need to measure the outer and inner radii: say $r_{max}(x)$ is the outer radius as a function of x and $r_{min}(x)$ is the inner radius as a function of x – when we are revolving around the y-axis we get r_{max} and r_{min} are just the top function and the bottom function, but in general we might need to adjust these two by some constant amount if the axis is of the form $y = a$, so it is better to think of it this way. The formula is then

$$V = \int_a^b \pi(r_{max}(x)^2 - r_{min}(x)^2) dx$$

when we integrate with respect to x and a similar formula is available for integrating with respect to y.

Shells: On the other hand, say we want to revolve around an axis parallel to the y-axis but we want to integrate with respect to x. Then we use the method of cylindrical shells: the idea is that we approximate the volume of revolution by a sequence of concentric cylinders, whose base radius is the distance from the axis of revolution, whose height is the height of the function at that point, and which is "differentially thin": the volume of the cylindrical shell of radius r, height h and thickness Δx is (exactly) $2\pi r h \Delta x$. This leads to a Riemann sum

$\Sigma_{i=0}^{n-1} 2\pi r(x_i^*) f(x_i^*)\Delta x$ and hence to the expression

$$V = \int_a^b 2\pi r(x) f(x) dx$$

when there is only one curve $y = f(x)$ or, if there are two,

$$V = \int_a^b 2\pi r(x)(f_{max}(x) - f_{min}(x))dx$$

As above, if we revolve around the y-axis precisely, then $r(x) = x$, but in general we could have $r(x) = \pm x + a$ if we revolve around an axis merely parallel to the y-axis – draw a picture to figure out what the radius should be as a function of x.

2.4. **arclength.** Let $f(x)$ be a *differentiable* function on $[a, b]$. (I wonder if anyone noticed the first time around that in all the other geometric applications it was enough for f to be continuous, but for arclength we need f to be differentiable. In fact there are continuous functions f whose arclength on an interval $[a, b]$ is infinite – i.e. if we layed out a piece of string on top of this function defined on a finite interval, we'd need an infinite amount of string. Try searching for "Koch curve" on google.com to see an example of this.) The idea is that on a small subinterval, $f(x)$ looks like a straight line with slope given by its derivative $f'(x)$, and the arclength of this small line is given by $(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 = (\Delta x)^2 + (f'(x_i)\Delta x)^2 = (\Delta x)^2(1 + f'(x_i)^2), so that \Delta s = \sqrt{1 + f'(x_i)^2)}$. This leads to a Riemann sum, and in the limit to the formula

$$\int_a^b \sqrt{1 + f'(x)^2} dx$$

for the arclength of a differentiable function f on $[a, b]$. However, it is a strange quirk of algebra that there are very few functions f such that $\sqrt{1 + f'(x)^2}$ can be integrated by the methods of this course – even trying to compute the arclength of an ellipse leads to something very tricky, called an elliptic integral and intensively studied by mathematicians in the 19th century. So it is, honestly, not very likely that you will see an arclength integral on the exam.

3. Density and mass

In high school chemistry we learn that mass equals volume times density. This is true if the density is constant; in general, the density of a (one, two or three-dimensional) region can vary on the region according to a density function, and to find the mass we will need to integrate the density function over the region. (In general this would be a multiple integral and hence the subject of 21a; here we study special cases where the density function depends only on a single parameter.) For example, if we had a mass given by horizontal slices from a to b and a density function constant on each slice, $\rho(x)$, then an expression for the mass is

$$M = \int_a^b A(x)\rho(x)dx$$

The other example that we saw lots of was when we had shells: either circular, spherical, or cylindrical, and the density depended only on the radius. In all cases, we adjust the appropriate volume integral by throwing in a factor corresponding to the density function, which we think of as "weighting" the integral, giving us a mass. If we had a radially symmetric density function ρ on a disk of radius R we would then get

$$M = \int_0^R 2\pi r\rho(r)dr$$

Or a radially symmetric density function ρ on a ball of radius R would give rise to an expression

$$M = \int_0^R 4\pi r^2\rho(r)dr$$

(In this last formula we have slipped in the formula for the approximate volume of a spherical shell of radius r and thickness Δr, namely $4\pi r^2\Delta r$. As we discussed in class, this is not literally the volume of a cylindrical shell, but it is asymptotic to the volume as the thickness goes to 0, so it's an acceptable approximation.) Finally, if we have a volume of revolution and a radially symmetric density function (corresponding to finding the volume using cylindrical shells) $\rho(r)$, then an expression for the mass is

$$M = \int_a^b 2\pi r f(r)\rho(r)dr$$

4. AVERAGE VALUES

Let f be a continuous function on the closed interval $[a, b]$. We have a notion of the *average value* of f on this subinterval: f_{ave} should be a number such that $\int_a^b f_{ave}dx = \int_a^b f(x)dx$; that is, it is the unique y-value such that the area of the rectangle with that y-value and width $b - a$ is the same as the area under the curve. (Imagine a sculpture cut from a slab of ice in the shape of the region under $y = f(x)$. If we let the ice melt, the depth of the water is the average value.) But this tells us that f_{ave} must be equal to $\frac{\int_a^b f(x)dx}{b-a}$. Another interpretation of the average value is as the limit of averages of the function at finitely many many sample points x_i^* in $[x_i, x_{i+1}]$. Choosing n such sample points, the average (as understood in life generally) of the values of f at these points is $\frac{\Sigma_{i=0}^{n-1} f(x_i^*)}{n} = \Sigma_{i=0}^{n-1} f(x_i^*)\frac{1}{n}$. This looks very much like a Riemann sum, except that in a Riemann sum we would multiply the values of f at the sample points not by $\frac{1}{n}$ but by $\Delta x = \frac{b-a}{n}$. That is, if we multiplied this average through by $b - a$, we would have a Riemann sum for the function. Therefore in the limit we get $\int_a^b f(x)dx = f_{ave} \cdot (b - a)$. This last formula is a nice alternate interpretation of the integral: it is the average value of the function times the length of the interval on which it is defined. As we discussed in class, in many real-world situations this makes more sense dimensionally: e.g. if $f(x)$ measures temperature as a function of time, the units of the integral are degrees Fahrenheit × minutes (or some other unit of time). It is better to say that the integral of temperature over a time interval is the average temperature over that period of time multiplied by the length of time.

A nice property of the average value of a continuous function is that it is always attained: that is, there exists at least one x_0 in $[a, b]$ such that $f(x_0) = f_{ave}$. Notice that this is *not* true for averages of finite sets of values: if the 2002 World Series goes 7 games and Barry Bonds hits 5 home runs, then we know, without watching even a single game, that there was no game in which he hit precisely $\frac{5}{7}$ home runs! This property is true for continuous functions because the f_{ave} lies somewhere between the minimum value of the function and the maximum value of the function, and it is a fundamental property of continuous functions – the Intermediate Value Theorem – that if $f(x_1) < f(x_2)$ then every y-value between $f(x_1)$ and $f(x_2)$ is the value of the function at some point x.

It is useful to remember that the average value of a linear function $f(x) = cx + d$ on the interval $[a, b]$ always occurs at the midpoint of the interval, so is $f(\frac{a+b}{2}) = c\frac{a+b}{2} + d$.

5. WORK

In physics we learn that work equals force times distance (to be precise, it is the component of the force that is in the direction of motion; e.g. whirling a ball around on the end of a string does no work; however this subtlety is for the physicists to sort out: for us it will go without saying that the force is in the direction of motion). As for density and average value, this is only literally true if the force is constant; if the force varies along an interval $[a, b]$ then we define

$$W = \int_a^b F(x)dx$$

i.e., work is the integral of force over distance. That's really all there is to say: now go back and review springs being pulled, nonviolent activists pulling themselves up the sides of buildings using a heavy chain that they absorb as they go, etc.

I will help you out with one kind of work problem, the pumping problem. We have a fluid-filled tank in a certain shape (not a cylinder, because that would be too freaking easy; a cone or a sphere often arise), and we want to pump the fluid out through the tank. To calculate the work being done, we give ourselves a coordinate, y starting at the top of the tank ($y = 0$) and becoming positive as we go down the tank, up to a maximum value of $y = b$. We imagine the tank is divided into thin cylindrical cross-sections with cross-sectional area $A(y)$; the work done to move each cylindrical piece of fluid is equal to its weight, which is $\rho A(y)\Delta y$, times the distance that we have to move that layer, which is y. (Here ρ is a constant which gives the weight density of the fluid.) So we get a Riemann sum of the form $\Sigma \rho A(y) y dy$. Taking the width of the slices to zero, we get an integral; we can even assume that the tank is only filled up to $y = a$ and then we get the expression

$$W = \int_a^b \rho A(y) y dy$$

Question: What if we used a different coordinate system, namely, we put $u = 0$ at the bottom of the tank and took positive u-values in the upward direction? (Whether this seems more or less natural than the first version is entirely a matter of taste.)

Solution: Write $B(u)$ for the cross-sectional area at u. Say the top of the tank has height b. Then the amount of water that is empty from the last solution was $b - a$, so our new integral will extend from 0 to $b - a$. Moreover, at height u we must pump the water a distance of $b - u$, so the integral is

$$W = \int_0^{b-a} \rho B(u)(b - u) du$$

Exercise: Show that these two integrals have the same value. (Hint: $B(u) = A(b - u)$; why is this? Now make the substitution $y = b - u$ to get from the second integral to the first.) This answers a question raised by CA Rob Hanna.

6. INTEGRATION BY SUBSTITUTION

I hope you all know by now what a u-substitution is, and that it is obtained by integrating the chain rule $(f(g(x)))' = f'(g(x))g'(x)$; I won't rehearse this here. Unfortunately I do not know how to tell you what is the right u-substitution to make, or whether or not you should make a u-substitution, or integrate by parts – this is something that you can only learn by experience. I will remind you how a u-substitution works for definite integrals:

$$\int_a^b f(g(x))g'(x) dx = F(g(x))|_a^b$$

so in $u = g(x)$ notation when we do a definite integral, instead of doing the corresponding indefinite integral and back-substituting $u = g(x)$ at the end, it comes to the same (but is faster) to change the x-limits to u-limits and then forget that u was originally some function of x:

$$\int_a^b f(u)u'(x) dx = \int_{u(a)}^{u(b)} f(u) du$$

While there is nothing wrong with subtituting back for u – so if you have your heart set against the above equation, you don't need to use it – what you should most definitely *not* write is something like

$$(WRONG!) \int_a^b f(u)u'(x) dx = \int_a^b f(u) du = F(u(b)) - F(u(a))$$

This is not the end of the world, because the first expression is equal to the last expression (so your final answer will be correct), but the middle expression entails a confusion of u and x that drives calculus instructors crazy (and may lead to loss of points).

6.1. **Trig substitutions.**

When you are in a tough situation, be alert to the possibility of making an inverse subsitution: $x = h(u)$ (instead of $u = g(x)$) – i.e. we are implicitly making the subsitution $u = h^{-1}(x)$. The magic of an inverse subsitution is that it always suceeds: in an ordinary substitution you must find (or somehow arrange to find) an extra factor of $g'(x)$ in your expression, or you cannot make the substitution at all. For an inverse substitution $x = h(u)$ we get $dx = h'(u)du$, which can always be done.

Example: Find $\int \frac{x^{\frac{1}{3}}}{1+x^{\frac{1}{2}}} dx$.

Solution: We would like to make the substitution $u = x^{\frac{1}{6}}$ but we do not see a factor of $u' = \frac{1}{6}x^{-\frac{5}{6}}$ anywhere. But now swoops down the inverse substitution: $x = u^6$, for which we get $dx = 6u^5 du$, and, using $x^{\frac{1}{3}} = u^2$, $x^{\frac{1}{2}} = u^3$, we get $\int \frac{u^2(6u^5 du)}{1+u^3} = 6 \int \frac{u^7}{u^3+1}$. Actually, this leads to a rational function which is not the easiest thing in the world to integrate – I did this example in one class but not the other, so some of you will remember that it was a good bit of work – but it illustrates the method.

The most famous inverse substitution is the trigonometric substitution. My take on this is very simple: if you see an expression of the form $\sqrt{a^2 - x^2}$ in an integrand, you should consider making the (inverse!) substitution $x = a \sin\theta$. Why? Because then $a^2 - x^2 = a^2 - a^2 \sin^2\theta = a^2(1 - \sin^2\theta) = a^2\cos^2\theta$, so that $\sqrt{a^2 - x^2} = a\cos\theta$, and the square root has gone away.
Similarly, if you see an expression with $a^2 + x^2$ in it (with or without a squareroot), you should consider making the substitution $x = a\tan\theta$, because then $a^2 + x^2 = a^2 + a^2\tan^2\theta = a^2(1 + \tan^2\theta) = a^2\sec^2\theta$ (recall that by dividing the familiar identity $\sin^2\theta + \cos^2\theta = 1$ by $\cos^2\theta$ we get $\tan^2\theta + 1 = \sec^2\theta$.
(Someday – not this Thursday – you might see an integrand containing the expression $\sqrt{x^2 - a^2}$ and then you ought to at least give some thought to the substitution $x = a\sec\theta$. I leave it to you to figure out why.)

Example: Find $\int \frac{dx}{(1+x^2)^2}$.

Solution: First I remark that this sort of integral could come up as a result of doing partial fractions with a repeated quadratic factor, so we are supposed to know how to do it. (Honestly, though, if you are asked to integrate a rational function with a *repeated* irreducible quadratic factor in the denominator, your calculus karma must be running very low.) Anyway, it's a good illustration of the trig sub: take $x = \tan\theta$, so $(1+x^2)^2 = (\sec^2\theta)^2$ and $dx = \sec^2\theta d\theta$. The integral becomes $\int \frac{\sec^2\theta d\theta}{\sec^4\theta} = \int \cos^2\theta d\theta$. We now use (e.g.) the identity $\cos^2\theta = \frac{1}{2}(1+\cos 2\theta)$, so the integral becomes $\int \frac{1}{2}(1 + \cos 2\theta)d\theta = \frac{1}{2}\theta + \frac{1}{4}\sin 2\theta$. Now we must change back to the x-variable (the difficulty in doing this is the price we pay for the sure-thing inverse substitution). We have $\theta = \arctan x$, so the first term is just $\frac{1}{2}\arctan x$. For the second term, we should use the further identity $\sin 2\theta = 2\sin\theta\cos\theta$, and figure out what $\sin\arctan x$ and $\cos\arctan x$ are. For this, we draw a little right triangle with angle θ and such that the tangent is x, so that the legs have length x and 1 and therefore the hypotenuse has length $\sqrt{1+x^2}$. This tells us that $\sin\arctan x = \sin\theta = \frac{x}{\sqrt{1+x^2}}$ and $\cos\arctan x = \cos\theta = \frac{1}{\sqrt{1+x^2}}$, so the final answer is $\frac{1}{2}\arctan x + \frac{1}{2}\frac{x}{1+x^2}$.

7. INTEGRATION BY PARTS

Again, I hope you know that the basic formula is $\int udv = uv - \int vdu$. If there is a definite integral, don't forget to evaluate the uv part between the limits: $\int_a^b udv = uv|_a^b - \int_a^b vdu$.

Like the u-substitution, learning when and how to integrate by parts is largely a matter of experience, but somehow I feel more able to tell you some patterns in this case (as I did in class that day...). The first pattern is when you have $\int x^n f(x)dx$, where $f(x)$ is a function you could find antiderivatives of all the live-long day. Then you take $u = x^n, dv = f(x)dx$, so $du = nx^{n-1}dx, v = F(x)$, and you get

$$\int x^n f(x)dx = x^n F(x) - n\int x^{n-1}F(x)dx$$

Since the power of x has been reduced by 1, progress has been made. Eventually (by taking more and more antiderivatives of f) there will be no power of x at all. So e.g. you can do $\int x^2 e^x dx$ by integrating by parts twice and (as was posed as an extra credit problem solved by several students) $\int x^{17}e^x dx$ by integrating by parts 17 times.

Another pattern is that if you see $\ln x$ (or even $(\ln x)^n$) in an expression, it is often a good idea to make it be the u function that gets differentiated. An example of this was on the techniques test: $\int(\ln x)^2$. Take $u = (\ln x)^2, dv = dx$. Then $du = 2(\ln x)\frac{1}{x}dx$ and $v = x$, to get $\int(\ln x)^2 dx = x(\ln x)^2 - \int(2(\ln x)\frac{1}{x})xdx$; the last term works out nicely to be $2\int(\ln x)dx$, which of course you should integrate by parts but really you should memorize (I do not repeat my non/sobriety story in print!). The answer is

$$x(\ln x)^2 - 2x\ln x + 2x$$

Beware that multiple integrations-by-parts introduces multiple minus signs; many of you had $-2x$ instead of $-(-2x)$ on the exam...

The last pattern is that sometimes you integrate by parts twice and return, not to where you started, but to *minus* where you started. This is great – moving the desired integral to the other side, you get that twice the integral you're trying to find is the expression you've gotten so far. I'll leave you to look at the classic example (see it again for the first time) of $\int \sin x e^x dx$.

8. INTEGRATION OF RATIONAL FUNCTIONS

When $f(x) = \frac{p(x)}{q(x)}$ is the quotient of two polynomials – a rational function – we have a systematic method for finding the antiderivative, in theory at least (in practice, the algebra involved becomes increasingly intricate). There is a three-step process:
Step 1: Make sure that $\frac{p(x)}{q(x)}$ is a proper rational function, i.e. that the degree of the denominator exceeds that of the numerator. That is, if this is not the case, then long-divide the denominator by the numerator to get a polynomial (great!) plus a rational function with denominator $q(x)$. As an example, our function $\frac{x^7}{1+x^3}$ from above is improper, and dividing we get $f(x) = \frac{x^7}{x^3+1} = x^4 - x + \frac{x}{x^3+1}$.

Step 2: Via a partial fractions decomposition, we can split up our integral into a sum of integrals of the form $\int \frac{C}{(ax+b)^n}$ and/or $\int \frac{Ax+B}{(ax^2+bx+c)^n}$, where in the second case the quadratic function $ax^2 + bx + c$ has no real roots (is "irreducible" over the real numbers). So we must first convince ourselves that we know how to integrate these two basic expressions. The first one is easy; just make $u = ax + b$ and you're all set. As for the second, by completing the square in the denominator – i.e. writing the denominator in the form $c(x-r)^2 + a^2$ (which can be done since there are no real roots), and then making the substitution $u = x - r$, it's enough to worry about how to find $\int \frac{Ax+B}{(x^2+a^2)^n} dx$. We split this up into $A \int \frac{x}{(x^2+a^2)^n} dx + B \int dx(x^2+a^2)^b$. Now the first term succumbs to the substitution $u = x^2 + a^2$ whereas the second gets the trig sub $x = a \tan \theta$ – we did an example of this above. (In practice, you will probably see no worse than $\int dxx^2 + a^2$, and you should remember that this is going to come out as $\frac{1}{a} \arctan \frac{x}{a}$.)

Step 3: Next we have to discuss the form of the partial fractions decomposition, which you need to memorize (sorry!). But there's really not that much to remember: the denominator will factor into a bunch of factors $(x - r_i)^{n_i}$ for various roots r_i (linear factors, possibly occurring with multiplicity greater than 1) together with factors $(ax^2 + bx + c)^{n_i}$ (irreducible quadratic factors, possibly occurring with multiplicity greater than 1). For each factor of the form $x - r_i)^{n_i}$, you write a sum of terms $\frac{C_1}{x-r_i} + \frac{C_2}{(x-r_i)^2} + \ldots + \frac{C_n}{(x-r_i)^n}$; i.e. the numerator is just a constant, even as the multiplicity increases. For each factor of the form $(\mathrm{ax}^2 + bx + c)^n$ you write a sum of terms $\frac{A_1x+B_1}{ax^2+bx+c} + \frac{A_2x+B_2}{(ax^2+bx+c)^2} + \ldots + \frac{A_nx+B_n}{(ax^2+bx+c)^n}$.
Then you must solve for these coefficients, which you do by multiplying through by the entire denominator and getting an equality of two polynomials. Then you set each of the coefficients of the polynomial on the lefthand side equal to each of the coefficeitns of the polynomial on the righthand side, obtaining a system of linear equations for the unknown A's, B's and C's. (Yes, this could be a real pain in the butt; hopefully the total degree of the denominator will be at most 3 or so if you are asked to this on an exam.) In case there are linear factors $(x - r_i)$ you will save at least a little time by plugging in $x = r_i$ on both sides, since every term on the righthand side except one will contain $(x - r_i)$ as a factor. (In the happy state when the denominator is a product of distinct linear factors $(x - r_1) \cdots (x - r_n)$, plugging in x equals each of the roots will very quickly tell you what all the coefficients are.)

Example: Let's finish the example above; we're supposed to be finding an antiderivative of $f(x) = \frac{x^7}{x^3+1} = x^4 - x + \frac{x}{x^3+1}$. We factor $x^3 + 1 = (x+1)(x^2 - x + 1)$. The second factor has no real roots, because the discriminant of this quadratic polynomial is $b^2 - 4ac = 1 - 4(1)(1) < 0$. So what we are going for is

$$\frac{x}{(x+1)(x^2-x+1)} = \frac{C}{x+1} + \frac{Ax+B}{x^2-x+1}$$

Multiplying through, we get the equality of polynomials

$$x = (x^2 - x + 1)C + (x + 1)(Ax + B)$$

So take $x = -1$ to get $-1 = ((-1)^2 - (-1) + 1)C = 3C$, so $C = -\frac{1}{3}$; that's some progress anyway. Now multiplying out the righthand side gives

$$0x^2 + 1x + 0 = (A - \frac{1}{3})x^2 + (A + B + \frac{1}{3})x + (B - \frac{1}{3})$$

so we find $A - \frac{1}{3} = 0$, so $A = \frac{1}{3}$ and $B - \frac{1}{3} = 0$, so $B - \frac{1}{3}$; as a check, the middle term comes out to be $1 = \frac{1}{3} + \frac{1}{3} + \frac{1}{3}$, which looks good. So $f(x) = x^4 - x - \frac{1}{3}\frac{1}{x+1} + \frac{1}{3}\frac{x+1}{x^2-x+1}$. We are still not quite homefree; we need to complete the square in the denominator to get $x^2 - x + 1 = (x^2 - x + \frac{1}{4}) + 1 - \frac{1}{4} = (x - \frac{1}{2})^2 + (\frac{\sqrt{3}}{2})^2$. So in the integral $\frac{1}{3}\frac{x+1}{(x-\frac{1}{2})^2+(\frac{\sqrt{3}}{2})^2}dx$, make the substitution $u = x - \frac{1}{2}$ to get $\int \frac{1}{3}\frac{u+\frac{1}{2}}{u^2+a^2}dx$, where $a = \frac{\sqrt{3}}{2}$. Finally, we split up this integral; for the piece $\frac{1}{3}\int \frac{u}{u^2+a^2}du$ we take $w = u^2 + a^2$, $dw = 2udu$, getting $\frac{1}{6}\ln(u^2 + a^2)$. The last piece integrates to $\frac{1}{6a}\arctan(\frac{u}{a})$, so the entire answer is

$$\frac{1}{5}x^5 - \frac{1}{2}x^2 - \frac{1}{3}\ln|x+1| + \frac{1}{6}\ln((x - \frac{1}{2})^2 + 3/4) + \frac{1}{3\sqrt{3}}\arctan(\frac{x - \frac{1}{2}}{\frac{\sqrt{3}}{2}}).$$

9. IMPROPER INTEGRALS

If f is a continuous function on a closed interval $[a, b]$, then $\int_a^b f(x)dx$ always exists. If however, we try to find the area under a curve over an infinite interval $[a, \infty)$ or $(-\infty, -\infty)$ OR if the function f is discontinuous at at least one point on the interval $[a, b]$, then the finiteness of the area is by no means assured. Integrals of these types are called *improper*, to emphasize that there is an (additional) limiting process taking place, and that the limit may or may not exist. As our basic example, we define $\int_a^\infty f(x)dx$ as $\lim_{b\to\infty} \int_a^b f(x)dx$, i.e. as $\lim_{b\to\infty} F(b) - F(a)$. That is, what is in question is the limiting behavior of the antiderivative as the variable approaches infinity. Recall that there are three essentially different types of limiting behavior:
1) convergence: $\lim_{x\to\infty F(x)}$ exists and is finite. Examples: $f_1(x) = x^{-p}$ for $p > 1$, $f_2(x) = e^{-x}$; $f_3(x) = p(x)/q(x)$ any rational function such that the degree of the denominator is at least *two more* than the degree of the numerator.
2) divergence to infinity: $\lim_{x\to\infty} F(x) = +\infty$ or $-\infty$: the former means that for any fixed y-value N, then the values of $F(x)$ will be at least N for all sufficiently large x. (Divergence to $-\infty$ is the same except that the values of $F(X)$ will be less than or equal to $-N$ for all sufficiently large x.) Examples: $f_1(x) = x^{-p}$ for $p \leq 1$; $f_2(x) = 3$; $f_3(x) = p(x)$ is any polynomial function (the improper integral will diverge to $\pm\infty$ according to whether the leading term is positive or negative).

Important: If $f(x) \geq 0$ for all x, then the area function $F(x)$ is increasing (why?). If we take the limit of an increasing function at ∞ either 1) or 2) above must occur: the only question is whether the function remains bounded, i.e. whether there exists a fixed constant M such that $F(x) \leq M$ for all x. If $F(x)$ remains bounded, it converges; otherwise it diverges to ∞. This makes improper integrals of positive functions much nicer to deal with; in particular the following cannot occur:

3) Divergence due to oscillation: This means that neither 1) nor 2) above occurs: the values of $F(x)$ do not settle down to any fixed value nor do they get arbitrarily large. Example: $f(x) = \cos x$, so $F(x) = \sin x$; the signed area accumulated (positive area - negative area) grows as large as 1 and as small as -1 for arbitrarily large x.

Principle of comparison: Say $0 \leq f(x) \leq g(x)$ are functions defined on $[a, \infty)$. Geometrically we see that the region under the smaller function $f(x)$ is contained in the region under the larger function $g(x)$, so certainly the area under $f(x)$ should be less than or equal to the area under $g(x)$ (and indeed, if f and g are continuous, then the area under f will be strictly less than the area under g unless $f = g$ the whole time). That means, that just as for "proper integrals" we have $\int_a^\infty f(x)dx \leq \int_a^\infty g(x)dx$. Keeping in mind that these improper integrals either converge or are infinite, then (under the reasonable convention that any real number is less than $+\infty$) we can make sense of this inequality even when one or both terms are infinite. We get the following conclusions:

a) If $\int_a^\infty g(x)dx < \infty$, then $\int_a^\infty f(x)dx < \infty$.
b) If $\int_a^\infty f(x)dx = \infty$, then $\int_a^\infty g(x)dx = \infty$.

As an example, we know that $\int_1^\infty e^{-x^{100}}dx$ converges, because for all $x \geq 1$, $x \leq x^{100}$, so $e^x \leq e^{x^{100}}$ and taking reciprocals reverses the inequality: $e^{-x} \geq e^{-x^{100}}$. Since we can compute $\int_1^\infty e^{-x}dx = \frac{1}{e}$, we can conclude by comparison that $\int_1^\infty e^{-x^{100}}dx$ converges, and indeed that it is less than $\frac{1}{e}$.

9.1. **Integrals on** $(-\infty, \infty)$**.** If $f(x)$ is defined on $(-\infty, \infty)$ (i.e. on the entire real line), then we can still ask for the area under the curve, but we have to be careful how we define it. The definition we choose is to split up the integral as follows: choose any number a; then

$$\int_{-\infty}^{\infty} f(x)dx := \int_{-\infty}^{a} f(x)dx + \int_{a}^{\infty} f(x)dx$$

That is, we require *both* of these integrals to converge separately; in particular if one of them is $+\infty$ and the other is $-\infty$, then we regard this as doubly bad – the integral is divergent for "two reasons". (In class we examined $\int_{-\infty}^{\infty} xdx$ and discussed whether or not the answer should be 0; now, while studying for the exam, we just reiterate that this is our definition of improper integral.)

9.2. **Discontinuities at a point.** If $f(x)$ is discontinuous at a point c on the interval $[a, b]$ (c could be an endpoint or a point in the interior of the interval) then again the well-definedness of the area under the curve is not assured. Say for example that $f(x)$ is defined on $[0, 1]$ and is continuous except at 0 – e.g. $\frac{1}{x^p}$ for any $p > 0$ fits the bill. Then $\int_0^1 f(x)dx$ is *improper*, and by this symbol we mean $\lim_{t \to 0^+} \int_t^1 f(x)dx$. As a very important example, recall that we found in class that $\int_0^1 \frac{dx}{x^p}$ was convergent (finite) exactly when $p < 1$. If the discontinuity is in the interior of the interval of definition, then we need to split up the integral and take limits as we approach the bad point both on the left and on the right.

As a fair warning, this is the point in the semester when we sometimes try to sneak in integrals that are improper because the function is discontinuous somewhere on the integral to see whether or not you'll notice this. (We can be pretty mean sometimes.) For example, consider $\int_{-1}^{1} \frac{dx}{x}$. This is an odd function on an interval symmetric about 0, so right away the integral is 0 right? Wrong! The function $f(x)$ is discontinuous at 0, and if we approach 0 from either direction we build up infinite area. (Notice that the fact that we have the "same infinite regions" on the left and on the right, one building up infinite positive area and the other building up infinite negative area, is not enough to get us off the hook; BY DEFINITION we require both of these improper integrals to converge.)

Absolute versus conditional convergence (an optional preview): After all this, it may seem sensible to define the improper integral of a function $f(x)$ defined on $(-\infty, \infty)$ as convergent if the total positive area is finite and the total negative area is finite. This is in fact a nice notion, but it is not quite the notion we've adopted, the reason being that in our very first example, $\int_a^\infty f(x)dx$ could converge even though the total positive and negative areas are both infinite: there could be a phenomenon of "alternation" between the positive and negative areas, so that each positive piece cancels out most of the negative piece that has come before it.

Extra credit: Show that $\int_\pi^\infty \frac{\sin x}{x}dx$ is convergent, despite the fact that the total positive area and the total negative area are both infinite.

Warning: At this point in the course, this problem would be very difficult to do. If you are interested you should at least graph the function in question (ok to use a calculator, it's extra credit) and get a sense of why there might be cancellation between the positive area and the negative area. Then, in about two weeks when we've covered the Alternating Series Test, try the problem again.

Aside: There's also some local history behind this problem. It was in fact posed on the qualifying exam for Harvard math (graduate!) students in the Fall of 1999 (the year after I took the exam); the exact wording began with the phrase "Using only freshman calculus...". But, as proof of how not easy it is, my friend David Dumas didn't get this problem (he had some complicated idea which, when the smoke cleared, amounted to comparing the integral to $\int_\pi^\infty dx/x$, which, as you all know, is divergent). His solution was graded as follows: "This proves nothing; 8 out of 10," which shows that Harvard grade inflation is not limited to the college. (Lest you think

I am being mercilessly instead of mischievously mean, I hasten to add that David is an excellent math student; indeed he holds the honor of being the only person I've ever met who could solve the notorious last problem on the Fall 2000 first Math 1a exam right off the top of his head; so far I know two Fields Medalists who could not.)

Anyway, if we have an improper integral such that the positive and the negative areas are *both* finite, this is the best possible convergence and could be called *absolute convergence* (we will later meet this notion in the context of infinite series.) An integral like in the previous examples, where *as we approach* ∞ the total positive areas and negative areas are both infinite, but somehow they are mixed together in such a way that the limit is still finite, is called *conditionally convergent.* No way will this material be on the exam.

10. PROBABILITY

A probability density function $f(x)$ is by definition a function defined on $(-\infty,\infty)$ which is continuous, non-negative and such that $\int_{-\infty}^{\infty} f(x)dx$ (converges and) equals 1. There are several ways to view $f(x)$ intuitively – perhaps the best, in line with other course material, is to view $f(x)$ as giving the density function of a mass spread out along the entire real line, such that the total mass is 1. (This also suggests that if we had a non-negative function f such that $\int_{-\infty}^{\infty} f(x)dx = C$ (finite!), then we can "renormalize" and get $\frac{1}{C}f(x)$ as a probability density function.)

We then define the probability that $a \leq x \leq b$ to be $\int_a^b f(x)dx$. So, what it really means to say that the probability density function $f(x)$ has value y_0 at a point x_0 is to say that in a small interval containing x_0 of length l, the probability that the value of the function lies in that interval is $l \times y_0$.

Expected value: If $f(x)$ is a probability density function, then its expected (or mean) value is by definition

$$E = \int_{-\infty}^{\infty} xf(x)dx$$

if it converges. One can show that this expected value is a limit of expected values of variables on *finite* sample spaces, but I feel this is too intricate to discuss in class (or here). Intuitively, we "weight" the integral by x so that, on some small interval $[x_i, x_{i+1}]$ where the probability density function is approximately constantly equal to $p_i = f(x_i)$, then the probability that the value lies on that subinterval is $(\Delta x)f(x_i)$, whereas the value itself is approximately equal to x_i (to pick a point on the subinterval). Therefore, in comparison to the finite expected value, the expected value of this piece of the probability density function is about $(\Delta x)f(x_i)x_i$; this Riemann sum leads to the indicated integral.

Remark: Interestingly, the expected value can be interpreted also as the center of mass of the distribution (see pages 483-484 of your text) – given that the total mass is 1, the formula is verbatim the same. This means that the expected value is the value at which the probability distribution would balance if supported by a pin.

A final strange example: Find the constant C such that $f(x) = \frac{1}{x^2+1}$ is a probability density function. What is its expected value?

Solution: $\int_{-\infty}^{\infty} \frac{dx}{x^2+1} = \arctan(\infty) - \arctan(-\infty) = \frac{\pi}{2} - \frac{-\pi}{2} = \pi$. So $\frac{1}{\pi}\frac{1}{1+x^2}$ is a probability density function. If we consider its graph, it is symmetric about $x = 0$ – it looks sort of like a fake normal distribution. In particular, we expect that the expected value (sorry) should be 0. But, by definition the expected value is $\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x}{x^2+1}dx$, and this improper integral is *divergent* at both $-\infty$ and $+\infty$. (It is again a case of having infinite positive area at ∞ and infinite negative area at $-\infty$.) Isn't that weird?
Good luck on the exam!

DEPT. OF MATHEMATICS, 1 OXFORD ST., CAMBRIDGE, MA 02138
E-mail address: `pete@math.harvard.edu`