Math 208H
Topics for the second exam

(Technically, everything covered on the first exam, *plus*)

Constrained Optimization: Lagrange Multipliers

Most optimization problems that arise naturally are not unconstrained; we are usually trying to maximize one function while *satisfying* another. We can use the one-variable calculus trick of solving the constraint for one variable, and plugging this into the function we wish to maximize, **or** we can take a completely different (and often better) approach: The basic idea is that if we think of our constraint as describing a level curve (or surface) of a function g, then we are trying to maximize or minimize f among all the points of the level curve. If the level curves of f are cutting *across* our level curve of g, it's easy to see that we can increase or decrease f while still staying on the level curve of g. So at a max or min, the level curve of f has to be *tangent* to our constraining level curve of g. This in turn means:

At a max or min of f subject to the constraint g, $\nabla f = \lambda \nabla g$ (for some real number λ)
We must also satisfy the constraint : $g(x, y) = c$.

So to solve a constrained optimization problem (max/min of f subject to the constraint $g(x, y) = c$) we solve

$\nabla f = \lambda \nabla g$ **and** $g(x, y) = c$ for x, y, and λ. All of the pairs (x, y) that arise are candidates for the max/min; and the max and min must occur at some of these points. [Technically, as before, we must also include points along $g(x, y) = c$ where ∇f is undefined; we won't run into this possibility in practice, however.]

This also works for functions of more than two variables; the procedure is exactly the same. In all of these cases, the real work is in solving the resulting equations! A basic technique that often works is to solve each of the coordinate equations in $\nabla f = \lambda \nabla g$ for λ ; the other halves of the equations are then all equal to one another (since they all equal λ).

This in turn allows us to finish our procedure for finding global extrema, since step (3) can be interpreted as a constrained optimization problem (max or min on the *boundary*). In these terms,

To optimize f subject to the condition $g(x, y) \leq c$, we

(1) solve $\nabla f = 0$ and $g(x, y) < c$,
(2) solve $\nabla f = \lambda \nabla g$ and $g(x, y) = c$,
(3) plug all of these points into f,
(4) the biggest is the max, the smallest is the min.

[This works fine, unless the region $g(x, y) \leq c$ runs off to infinity; but often, physical considerations will still tell us that one of our critical points is an optimum.]

The Definite Integral of a Function of Two Variables

In an entirely formal sense, the intergal of a function of one variable is a great big huge sum of little tiny numbers; we add up things of the form $f(c_i)\Delta x_i$, where we cut the interval $[a, b]$ we are integrating over into little intervals of length Δx_i, and pick points c_i in each interval. In esssence, the integral is the sum of areas of very thin rectangles, which leads us to iterpret the integral as the area under the graph of f.

For functions of two variables, we do the exact same thing. To integrate a function f over a rectangle in the plane, we cut the rectangle into lots of tiny rectangles, with side lengths Δx_i and Δy_j, pick a point in each rectangle, and then add up $f(x_i, y_j)\Delta x_i \Delta y_j$. This gives an *approximation* to the actual integral; letting the little side lengths go to zero, we arrive at what we would call the integral of f over the rectangle R, which we denote by

$\int\int_R f \, dA$ (where dA denotes the 'differential of area' $dxdy$ (or $dydx$)

The idea is that if we think of f as measuring height above the rectangle, then $f(x_i, y_j)\Delta x_i \Delta y_j$ is the volume of a thin rectangular box; letting the Δ's go to zero, the integral would then measure the *volume* under the graph of f, lying over the rectangle R.

If the region R *isn't* a rectangle, we can still use this method of defining an integral; we simply *cover* R with tiny rectangles, take the same sum, and let the Δ's go to 0.

Of course, we have no reason to believe that as the Δ's go to 0, this collection of sums will *converge* to a single number. But it is a basic fact that if the function f is *continuous*, and the region R isn't too ugly, then these sums always will converge.

Iterated Integrals

Of course, the preceding approach is no way to *compute* a double integral! Instead, we (as usual) steal an idea from one-variable calculus.

The idea is that we already *know* how to compute volumes, and so we *implicitly* know how to compute double integrals! We can compute the volume of a region by integrating the *area of a slice*. You can do this two ways; (thinking in terms of the region R in the plane) you can slice R into horizontal lines, and integrate the area of the slices dy, or you can slice R into vertical lines, and integrate the slices dx.

But each *slice* can be interpreted as an integral; the area of a horizontal slice is the integral of f, thought of as *just a function of* x, and the area of a vertical slice is the integral of f, thought of as *just a function of* y. This leads to *two* ways to compute our integral:

$$\int\int_R f\ dA = \int_c^d (\int_a^b f(x,y)\ dx)\ dy \text{ (for horiz slices)} = \int_a^b (\int_c^d f(x,y)\ dy)\ dx \text{ (for vert slices)}$$

In each case, the *inner* integral is thought of as the integral of a function of *one variable*. It just happens to be a *different* variable in each case. In the case of a rectangle, the limits of integration are just numbers, as we have written it. In the case of a more complicated region R, the inner limits of integration might depend on where we cut. The idea is that a slice along a horizontal line is a slice along $y =$ constant, and the endpoints of the integral might *depend on* y; for a slice along a vertical line ($x =$ constant), the endpoints might depend on x .

So, e.g., to integrate a function f over the region lying between the graphs of $y = 4x$ and $y = x^3$, we would compute either

$$\int_0^2 (\int_{x^3}^{4x} f(x,y)\ dy)\ dx \qquad \text{or} \qquad \int_0^8 (\int_{y/4}^{y^{1/3}} f(x,y)\ dx)\ dy$$

Which should we compute? Whichever one appears to require the least effort! They give the *same* number!

Area and average value as integration

We can think of a double integral over a region R as the volume of the region lying under the graph of f; this helped us to formulate iterated integrals as an approach to their computation. A different point of view comes when we think of the rectangles we cut R into as having the same size; then a Riemann sum is (a constant times) the sum of values of f at evenly distriibuted points around the region R. In symbols,

$$\int\int_R f\ dA \ \approx\ \Delta A \sum f(P_i) = (n\Delta A)(\frac{1}{n}\sum f(P_i)) = (\sum_i \Delta A)(\frac{1}{n}\sum f(P_i))$$

But $\sum_i \Delta A$ approximates $\int\int_R 1\ dA$, which we interpret as the volume of a cylinder with base R and height 1, which has volume (Area of R)(1) = Area of R. And $(\frac{1}{n}\sum f(P_i))$

can be thought of as an approximation for the *average value of the function* f, since the n points P_i are evenly distributed around R. Putting this all together, we end up with

$$\int\int_R f\ dA = (\text{Average value of } f)(\text{Area of } R).$$

In this sense, the average value of f over R is the height of cylinder over the base R which will have the same volume as the region lying under the graph of f.

We can use this either as a <u>definition</u> of the average value of f (and then use the integrals to compute it!) <u>or</u> as a method for approximating $\int\int_R f\ dA$, by picking a large number of points in R at random and computing the average value of f at those points (as a way of approximating this 'ideal' notion of the average value of f) and then multiplying by the area of R. This second interpretation is probably the one that sees actual use more often; as we have seen in one variable calculus, there are a lot of integrals that resist our abilities to compute the 'old-fashioned' way (by the fundamental theorem of calculus), and multiple integrals are no different! But applications in physics, economics, biology, engineering, and a whole host of other disciplines still need to know the value of those integrals!

Change of variables

Polar coordinates illustrate the benefit of decribing points in the plane differently - it can simplify the description of some regions in the plane. This, in turn, can aid us in the computation of some double integrals. The idea is to *change variables*; it's basically u-substitution for function of two variables.

The general idea is that if a region R can be described more conveniently using a different sort of coordinates, this means that we are describing x and y as *functions* of different variables s and t. For example, a circle of radius 4 is better described as

$$x = r\cos\theta \text{ and } y = r\sin\theta, \text{ for } 0 \le r \le 4 \text{ and } 0 \le \theta \le 2\pi$$

(i.e., polar coordinates). In general, changing coordinates means describing the region R by

$$x = x(s,t) \text{ and } y = y(s,t), \text{ for } s \text{ and } t \text{ in some region } S$$

Then we write the integral of the function f over R as the integral of *something else* (written in terms of s and t) over the region S. The question is, the integral of *what*? The answer comes from thinking of cutting up S into little rectangles S_{ij}, and looking at the little regions R_{ij} the change of variables carries each to. The integral of f over R can be approximated by adding up values in each region R_{ij}, times the area of R_{ij}. By choosing (s_i, t_j) in S_{ij}, we can use the point $(x(s_i,t_j), y(s_i,t_j))$ in R_{ij}; the question is, what is the area of R_{ij}?.

If we think of the rectangles S_{ij} as having sides of length ds and dt, then using linear approximations to $x(s,t)$ and $y(s,t)$, R_{ij} can be approximated by a parallelogram with sides the vectors

$$(\frac{\partial x}{\partial s}, \frac{\partial y}{\partial s})ds \text{ and } (\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t})dt$$

Luckily, we know how to compute the area of such a parallelogram; it's given by the length of the cross product of the two sides (add 0's to the vectors, so they are in 3-space!), which turns out to be:

$$\Delta A_{ij} = |x_s y_t - x_t y_s|\ ds\ dt$$

Taking limits as the size of the S_{ij} goes to zero, we obtain:

$$\int\int_R f(x,y)\ dx\ dy = \int\int_S f(x(s,t), y(s,t))\ |x_s y_t - x_t y_s|\ ds\ dt$$

The expression $|x_s y_t - x_t y_s|$ is called the *Jacobian* associated to the change of variables, and is sometimes written

$$|x_s y_t - x_t y_s| = \frac{\partial(x,y)}{\partial(s,t)}$$

For example, to integrate a function f over the triangle with vertices (1,1), (2,3), and (3,8), we can instead integrate over the triangle with vertices (0,0), (1,0), and (0,1), by changing coordinates. It turns out we can always do this by writing

$$x = as + bt + c \text{ and } y = ds + et + f$$

for appropriate choices of a, b, c, d, e and f. All you need to do is solve the equations $1 = a0 + b0 + c, 1 = d0 + e0 + f, 2 = a1 + b0 + c, 3 = d1 + e0 + f, 3 = a0 + b1 + c$, and $8 = d0 + e1 + f$ which, in this case, gives a=1,b=2,c=1,d=2,e=7,f=1. So $x = s + 2t + 1$ and $y = 2s + 7t + 1$, giving Jacobian $1 \cdot 7 - 2 \cdot 2 = 3$. So under this change of coordinates,

$$\int\int_R f(x,y)\ dA = \int_0^1 \int_0^{1-t} f(s + 2t + 1, 2s + 7t + 1) \cdot 3\ ds\ dt$$

Double integrals with polar coordinates

Polar coordinates describe a point in the plane by *distance* and *direction*, r and θ. We can translate from rectangular to polar coordinates by

$$(x, y) = (r\cos\theta, r\sin\theta)$$

We can use this new coordinate system to simplify some integration problems, in part because a circular disk is a *polar rectangle*, defined by $0 \le r \le R_0$ and $0 \le \theta \le 2\pi$. Similarly, circular sectors can be described as 'polar rectangles'.

But in so doing, we must interpret dA in terms of dr and $d\theta$; this is completely analogous to what we must do with u-substitution. If we have a small circular sector, made between the circles of radius r and $r + \Delta r$, and between the lines making angles θ and $\theta + \Delta\theta$, it has area approximately $r\Delta r\Delta\theta$; so

$$dA = r\ dr\ d\theta$$

and so $\int\int_R f(x,y)\ dA = \int\int_D\ f(r\cos\theta, r\sin\theta)\ r\ dr\ d\theta$, where D is how we describe the region R *in polar coordinates.*

For example, the integral of the function $f(x, y) = xy$ on the semicircle lying between the x-axis and y=$\sqrt{9 - x^2}$ can be computed as

$$\int_0^\pi \int_0^3 (r\cos\theta)(r\sin\theta) r\ dr\ d\theta$$

Triple Integrals

Triple integrals are just like double integrals, only more so. We can define them as a limit of a huge sum; here the terms in the sum would be the value of the function f time the volume of a tiny rectangular box. The usual interpretation of a triple integral arises by thinking of the function f as giving the density of the matter at each point of a solid region W in 3-space. Since density times volume is mass, the integral of f over the region W would compute the mass of the solid object occupying the region W. In the special case that f is the function 1, the integral will compute the volume of the region W.

Again, as with double integrals, the way we really comupute a triple integral is as a (triply) iterated integral. You pick a direction to slice (x=constant, y=constant, or z=constant) W up, and compute the integral of f over each slice. Each of these is a double integral (computed as an iterated integral), whose value *depends* on the variable you sliced along. To compute the integral over W, you integrate these double integrals over the last variable, getting three iterated integrals.

Put slightly differently, you can evaluate a triple integral by integrating out each variable, one at a time. Typically, we start with z, since our region W is usually described as the region lying between the graphs of two functions, given as z=blah and z=bleh . The idea is to first, for each fixed value of x and y, integrate the function f, dz, from blah to bleh. (Ther resulting values depend on x and y, i.e., are a *function* of x and y.) Then we integrate over the region, R, in the plane consisting of the points (x, y) such that the vertical line *hits* the region W. We usually call this region R the *shadow* of W in the x-y plane. In symbols

$$\int\int\int_W f\ dV = \int\int_R (\int_{a(x,y)}^{e(x,y)} f(x,y,z)\ dz)\ dA$$

For example, the integral of a function over the region lying above the x-y plane and inside the sphere of radius 2, centered at the origin, would be computed as

$$\int\int_R (\int_0^{\sqrt{4-x^2-y^2}} f(x,y,z)\ dz)\ dA = \int_{-2}^{2} (\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} (\int_0^{\sqrt{4-x^2-y^2}} f(x,y,z)\ dz)\ dy)\ dx$$

where R is the shadow of W (in this case, the disk of radius 2, centered at the origin, in the x-y plane).

Change of variables for Triple Integrals

As with double integrals, we can carry out a change of coordinates for 3 variables; we then write

$$x = x(s,t,u),\ y = y(s,t,u), \text{ and } z = z(s,t,u)$$

A little box with sides of length ds, dt, and du gets carried to a little parallelopiped, with sides the vectors

$$(x_s, y_s, z_s)\ ds,\ (x_t, y_t, z_t)\ dt, \text{ and } (x_u, y_u, z_u)\ du$$

(call these V_s, V_t, and V_u). This has volume $|V_s \bullet (V_t \times V_u)|$, which is the Jacobian of this change of variables, and serves as the necessary "fudge factor" to express an integral in terms of s, t, and u.