Math 208H

Topics for the second exam

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

Local Extrema

The partial derivatives of f measuire the rate of change of f in each of the coordinate directions. So they are giving us partial information (no pun intended) about how thew function f is rising and falling. And just as in one-variable calculus, we ought to be able to turn this into a procedure for findong out when a function is at its maximum or minimum. The basic idea is that at a max or min for f, then, thinking of f just as a function of x, we would *still* think we were at a max or min, so the derivative, as a function of x, will be 0 (if it is defined). In other words, $f_x = 0$. similarly, we would find that $f_y = 0$, as well. following one-variable theory, therefore, we say that

A point (a, b) is a **critical point** for the function f if $f_x(a, b)$ and $f_y(a, b)$ are *each* either 0 or undefined. (A similar notion would hold for functions of more than two variables.)

Just as with the one-variable theory, then, if we wish to find the max or min of a function, what we first do is find the critical points; *if* the function has a max or min, it will occur at a critical point.

And just as before, we have a 'Second Derivative Test' for figuring out the difference between a (local) max and a (local) min (or *neither*, which we will call a *saddle point*). The point is that at a critical point, f looks like its quadratic approximation, which (simplifying things somewhat) is described as $Q(x, y) = Dx^2 + Exy + Fy^2$ (since the first derivatives are 0). By completing the square, we can see that the actual shape of the graph of Q is basically described by *one number*, called the descriminant, which (in terms of partial derivatives) is given by

$$D = f_{xx}(\text{a,b})f_{yy}(\text{a,b}) - (f_{xy}(\text{a,b}))^2$$

(Basically, Q looks like one of x^2+y^2 (local min), $-x^2-y^2$ (local max), or x^2-y^2 (saddle), and D tells you if the signs are the same ($D > 0$) or opposite ($D < 0$) . More specifically, if, at a critical point (a, b),

$D > 0$ and $f_{xx} > 0$ then (a, b) is a local min; if
$D > 0$ and $f_{xx} < 0$ then (a, b) is a local max; and if
$D < 0$, then (a, b) is a saddle point
(We get no information if $D = 0$.)

Global Extrema: Unconstrained Optimization

Critical points help us find local extrema. To find *global* extrema, we take our cue from one-variable land, where the procedure was (1) Identify the domain, (2) find critical points *inside* the domain, (3) plug critical points and *endpoints* into f, (4) biggest is the max, smallest is the min.

For two variables, we do (essentially) *exactly the same thing*:

(1) Identify the domain
(2) Find critical points in the *interior* of the domain
(3) Identify the (potential) max and min values on the *boundary* of the domain (more about this later!)
(4) Plug the critical points, and your potential points on the boundary
(5) biggest is max, smnallest is min

This works if the domain is *closed* and *bounded* (think, e.g., of a closed interval in the x direction and a closed interval in the y direction, or the inside of a circle in the plane (including the circle)). Usually, in practice, we don't have such nice domains; but we usually know from physical considerations that our function *has* a max or min (e.g., find the maximum volume you can enclose in a box made from 300 square inches of cardboard...), and so we *still* know that it has to occur at a critical point of our function.

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:

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

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

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 $\displaystyle\int\int_R f\ dA$, 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!

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 $0 \leq r \leq R_0$ and $0 \leq \theta \leq 2\pi$. 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).

Triple integrals with spherical and cylindrical coordinates

It turns out that we can impose *two* new coordinate systems on 3-space, analogous to polar coordinates in 2-space; each can sometimes be used render an integration problem more tractible, usually by making the *region* we integrate over more ‘routine’.

With cylindrical coordinates, we simply replace (x, y, z) with (r, θ, z), i.e., use polar coordinates in the xy-plane. In the new coordinate system, $dV = (r\ dr\ d\theta)\ dz$, since that will be the volume of a small ‘cylinder’ of height dz lying over the small sector in the xy-plane that we use to compute dA above.

Usually, we will actually integrate in cylindrical coordinates in the order $dz\ dr\ d\theta$, since this coordinate system is most useful when the shadow of our region - the points in the (x, y)-plane that our region sits over - is a *disk* (or other polar rectangle).

Spherical coordinates are much like polar coordinates; we describe a point (x, y, z) by distance (which we call ρ and direction, except we need to use *two* angles to completely specify the direction; first, the angle θ that $(x, y, 0)$ makes with the x-axis in the xy-plane, and then the angle ϕ that the line through our point makes with the (positive) z-axis (which we can always assume lies between 0 and π). A little trigonometry leads us to the formulas

$$(x, y, z) = (\rho \cos\theta \sin\phi, \rho \sin\theta \sin\phi, \rho \cos\phi)$$

Again, the idea is that regions troublesome to describe in rectangular coordinates can be far more routine to describe spherically; for example, the inside of a sphere of radius R_0 can be described as the *rectangle* $0 \leq \rho \leq R_0$, $o \leq \theta \leq 2\pi$, and $0 \leq \phi \leq \pi$.

It is a bit more work to compute what dV is in spherical coordinates; computing the volume of a small ‘spherical box’, we find that it is

$$dv = \rho^2 \sin\phi\ d\rho\ d\theta\ d\phi$$

.

So the ‘change of variables formula’ for spherical coordinates reads:

$$\int\int\int_W f(x, y, z)\ dV = \int\int\int_R f(\rho \cos\theta \sin\phi, \rho \sin\theta \sin\phi, \rho \cos\phi)\ \rho^2 \sin\phi\ d\rho\ d\theta\ d\phi$$

So, for example, the integral of the function $f(x, y, z) = xz$ over the top half of a ball of radius 5 could be computed as

$$\int_0^5 \int_0^{2\pi} \int_0^{\pi/2} (\rho\cos\theta\sin\phi)(\rho\cos\phi)\ (\rho^2\sin\phi)\ d\phi\ d\theta\ d\rho$$

Vector-valued functions / Parametrized curves

So far, we have talked about functions of several variables; functions which need several inputs in order to get a single output. Our next topic is *parametrized curves*; functions which have one input but several outputs. We will focus on functions of the form

$\vec{r}(t) = (x(t), y(t))$ or $\vec{r}(t) = (x(t), y(t), z(t))$

i.e., curves in the plane or 3-space. If we think of t as time, then what $\vec{r}$ does is give us a point in the plane or 3-space at each moment of time. Thinking of $\vec{r}$ as the position of a particle, the particle sweeps out a path or curve, C, in the plane or 3-space as time passes; we think of $\vec{r}$ as *parametrizing* this curve C.

We therefore make a distinction between a *curve* (= a collection of point laid out in a string) and a *parametrized curve* (= a function which traces out a curve). A single curve can have many different parametrizations; for example,

$\vec{r_1}(t) = (\cos t, \sin t)$, $0 \leq t \leq 2\pi$
$\vec{r_1}(t) = (\cos 2t, \sin 2t)$, $0 \leq t \leq \pi$
$\vec{r_1}(t) = (\sin t, \cos t)$, $0 \leq t \leq 2\pi$
$\vec{r_1}(t) = (\cos t^2, \sin t^2)$, $0 \leq t \leq \sqrt{2\pi}$

all parametrize the (unit) circle in the plane. Their diffences with the first are that they go twice as fast, or travel in the opposite direction, or starts slowly and then moves faster and faster, respectively.

Of special interest are *lines*; they can be described as having a starting place and a direction they travel, and so can be parametrized by $\vec{r}(t) = P + t\vec{v}$, where P is the starting point and $\vec{v}$ is the direction (for example, the difference of two points lying *along* the line).

As with ordinary functions, we can build new parametrized curves from old ones by, for example, adding constants to each coordinate (which translates the curve by those amounts), or multiplying coordinates by constants (which streches the curve in those directions).

Velocity and acceleration

When we think of t as time, we can imagine ourselves as travelling along the parametrized curve $\vec{r}(t)$, and so at each point we can make sense of both *velocity* and *acceleration*. Velocity, which is the instantaneous rate of change of position, can be easily calculated from our parametrization $\vec{r}(t) = x(t), y(t), z(t)$ as $\vec{v}(t) = \vec{r}'(t) = x'(t), y'(t), z'(t)$

Similarly, acceleration can be computed as $\vec{a}(t) = \vec{r}''(t) = x''(t), y''(t), z''(t)$

On useful fact: if the length of the velocity (i.e., its *speed*), $||\vec{v}(t)||$ is constant, then $\vec{a}(t)$ is always perpendicular to $\vec{v}(t)$

And speaking of length, we can compute the *length* of a parametrized curve can be computed by integrating its speed: the length of the parametrized curve $\vec{r}(t)$, $a \leq t \leq b$, is

$$\text{Length} = \int_a^b ||\vec{v}(t)||\ dt$$

Vector fields

A vector field is a field of vectors, i.e., a choice of vector $F(x, y)$ (or $F(x, y, z)$) in the plane for every point in some part of the plane (the domain of F), and similarly in 3-space. We can think of F as $F(x, y) = (F_1(x, y), F_2(x, y))$; each coordinate of F is a function of several variables. We can represent a vector field pictorially by placing the vector $F(x, y)$ in the plane with its *tail* at the point (x, y) . A vector field is therefore a choice of a direction (and magmitude) at each point in the plane (or 3-space...). Such objects naturally occur in many disciplines, e.g., a vector field may represent the wind velocity at each point in the plane, or the direction and magnitude of the current in a river. Another important

class of examples are *force fields*: the vectors describe the direction and magnitude of a force (gravity, friction, magnetism) acting at each point.
One of the most important classes of vector fields that we will encounter are the *gradient vector fields*. If we have an (ordinary) function $f(x,y,z)$ of several variables, then for each point (x,y,z), $\nabla(f)$ can be thought of as a *vector*, which we have in fact already taken to drawing with its tail at the point (x,y,z) (so that, for example, we can use it as a normal vector for the tangent plane to the graph of f). Many vector fields are gradient vector fields, e.g., $(y,x)=\nabla(xy)$; one of the questions we will need to answer is 'How do you tell when a vector field is a gradient vector field?'. We shall see several answers to this question later on.

Flow lines / integral curves

Thinking of a vector field as a velocity field, like the current in a river, an object dropped into the river will tendd to be carried along in the direction of the velocity field. That is , the path that the object traces out will, at each point, have velocity equal to the value of the vector field at that point. Such a curve we call an *integral curve*; it is the curve $\gamma(t)=(x(t),y(t)$ whose velocity $\gamma'(t)$ equals the value of the vector field $\vec{F}$ at the point $\gamma(t)$; that is,

$$\gamma'(t)=\vec{F}(\gamma(t) \quad \text{for all } t$$

.

This equation is a *system of differential equations* $x'(t)=F_1(x(t),y(t))$, $y'(t)=F_2(x(t),y(t))$, where the F_i are the components of the vector field $\vec{F}(x,y)=(F_1(x,y),F_2(x,y))$. Finding solutions to such a system of equations in somewhat outside of the scope of our class! We can solve some such equations, e.g., if $\vec{F}(x,y)=(1,x^2+1)$ [or some other function of x in the second coordinate], then we wish to solve $x'(t)=1$ [so $x(t)=t+c$] and $y'(t)=(x(t))^2+1$ [so $y'(t)=(t+c)^2+1$, which we can solve by integration].

A common technique for <u>approximating</u> flow lines is Euler's method; starting at a point, the vector field tells you which direction to go, so go <u>in</u> <u>a</u> <u>straight</u> <u>line</u> for a short amount of time, then use the vector field at the new point to determine a new direction to go, and repeat. Doing this several times, letting the time interrval decrease to 0, you will find the path you traverse will converge to an integral curve.

Line Integrals

We introduced vector fields $F(x,y)$ in large part because these are the objects that we can most naturally integrate over a (parametrized) curve. The reason for this is that along a curve we have the notion of a velocity vector $\vec{v}$ at each point, and we can *compare* these two vectors, by taking their dot product. This tells us the extent to which F points in the direction of $\vec{v}$. Integration is all about taking averages, and so we can think if the integral of F over the curved C as measuring the *average* extent to which F points in the same direction as C.
We can set this up as we have all other integrals, as a limit of sums. Picking points $\vec{c_i}$ strung along the curve C, we can add together the dot products $F(\vec{c_i})\bullet(\vec{c_{i+1}}-\vec{c_i})$, and then take a limit as the lengths of the vectors $\vec{c_{i+1}}-\vec{c_i}$ between consecutive points along the curve goes to 0. We denote this number by

$$\int_C F\bullet d\vec{r}$$

Such a quantity can be interpreted in several ways; we will mostly focus on the notion of *work*. If we interpret F as measuring the amount of force being applied to an object at each point (e.g., the pull due to gravity), then $\int_C F\bullet d\vec{r}$ measures the amount of work done by F as we move along C. In other words, it measures the amount that the force field F *helped* us move along C (since moving in the same direction, it helps push us along, while when moving opposite to it, it would slow us down).

In the case that F measures the current in a river or lake or ocean, and C is a *closed* curve (meaning it begins and ends at the same point), we interpret the integral of F along C as the *circulation* around C, since it measures the extent to which the current would *push* you around the curve C.

Computing using parametrized curves

Of course, as usual, we would never want to *compute* a line integral by taking a limit! But if we use a parametrization of C, we can interpret $\int_C F \bullet d\vec{r}$ as an 'ordinary' integral. The idea is that if we use a parametrization $\vec{r}(t)$ for C then $F(\vec{c_i}) \bullet (\vec{c_{i+1}} - \vec{c_i})$ becomes

$$F(\vec{r}(t_i)) \bullet (\vec{r}(t_{i+1}) - \vec{r}(t_i))$$

But using tangent lines, we can approximate $\vec{r}(t_{i+1}) - \vec{r}(t_i)$ by $\vec{r}'(t_i)(t_{i+1} - t_i) = \vec{r}'(t_i)\Delta y$

.

so we can instead compute our line integral as

$$\int_C F \bullet d\vec{r} = \int_a^b F(\vec{r}(t)) \bullet \vec{r}'(t)\ dt$$

where $\vec{r}$ parametrizes C with $a \le t \le b$.

Some notation that we will occasionally use: If the vector field $F = (P, Q, R)$ and $\vec{r}(t) = (x(t), y(t), z(t))$, then $d\vec{r} = (dx, dy, dz)$, so $F \bullet d\vec{r} = Pdx + Qdy + Rdz$. So we can write

$\int_C F \bullet d\vec{r} = \int_a^b Pdx + Qdy + Rdz$