# Math 5B: Supplemental notes on lecture and Chapters 1, 2, and 4

*You should be able to read these notes along with your book (and refer to the pictures in the book when needed). This is a brief summary of the topics we've covered in the first half of the course; you can use this to help you review for the midterm and to figure out what material and examples to read in the book!*

## Chapter 1

**Vectors**

We usually describe vectors by their *coordinates*. For example, in $\mathbb{R}^3$, the vector $\mathbf{v} = (a, b, c)$ means that if we move the vector so that it starts at the origin $(0, 0, 0)$, it will point to the point $(a, b, c)$.

In $\mathbb{R}^2$, we sometimes describe points $(x, y)$ by their *polar coordinates* $(r, \theta)$. See Figure 1.2 in the book.

We can add (or subtract) vectors. Algebraically, this is done component-wise: for instance, if $\mathbf{v} = (2, -1, 0)$ and $\mathbf{w} = (-1, 2, 1)$, then $\mathbf{v} - \mathbf{w} = (3, -3, -1)$. Geometrically, we can understand the addition of vectors by the picture found in Figure 1.4.

Lengths of vectors are computed by applying Pythagoras' Theorem. In $\mathbb{R}^n$, if the coordinates of $\mathbf{v}$ are $(v_1, v_2, ..., v_n)$, then

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + ... + v_n^2}.$$

Vectors of length one are called *unit vectors*. Notice that multiplying a vector by a number $t$ just scales it, so $t\mathbf{v}$ points in the same direction as $\mathbf{v}$, but has a different length. Given a vector $\mathbf{v}$, we can always find a unit vector that points in the same direction by dividing $\mathbf{v}$ by its length $\|\mathbf{v}\|$.

The *standard unit vectors* are the unit vectors $\mathbf{i}, \mathbf{j}$, and $\mathbf{k}$ that point in the direction of the $x$, $y$ and $z$-axes, respectively. A vector with coordinates $(v_1, v_2, v_3)$ can also be written as $v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$.

There are two special products of vectors that we define. The *dot product* of two vectors returns a number, found by multiplying the corresponding coordinates together and adding up the results. (Also, we know that $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta$, where $\theta$ is the angle between the vectors.) The *cross product* $\mathbf{v} \times \mathbf{w}$ returns a vector that is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$. The calculation of the coordinates can be remembered as a determinant (see Chapter 1.5, and Chapter 1.4 if you need to review determinants.)

**Lines**

A vector is a convenient way to describe a line in $\mathbb{R}^n$. Since a vector defines a direction, one vector and one point specify a line. If $\mathbf{v} = (v_1, v_2, ..., v_n)$ points in the direction of a line, and $P = (a_1, a_2, ..., a_n)$ is a point on the line, how can we find all the other points $Q$ on the line? Notice that if we can scale $\mathbf{v}$ by a number $t$; the vector $t\mathbf{v}$ will point from $P$ to another point $Q$ on the line! See Figure 1.8 in the book. Every point on the line is found this way, so we can describe the coordinates of the points on the line $\mathbf{l}(t)$ by

$$\mathbf{l}(t) = (x_1(t), x_2(t), ..., x_n(t)) = (a_1, a_2, ..., a_n) + t(v_1, v_2, ..., v_n) = (a_1 + tv_1, a_2 + tv_2, ..., a_n + tv_n).$$

These equations for each of the coordinates $x_1(t) = a_1 + tv_1$, ..., $x_n(t) = a_n + tv_n$ are called the *parametric equations* of the line. Notice that each coordinate is a linear function of the *parameter* $t$.

For example, in $\mathbb{R}^2$, the equations $x(t) = 1 + t$ and $y(t) = -2 + 3t$ describe a line going through the point $(1, -2)$ in the direction of the vector $\mathbf{v} = \mathbf{i} + 3\mathbf{j}$. We are more used to seeing this line given as $y = 3x - 5$ (notice that if we solve $t = x - 1$ and plug that into the equation for $y$, we get this equation.)

**Planes**

See Example 1.26 and Figure 1.17 in the book. One simple way to find the equation of a plane is to know a *normal vector* for the plane as well as a point on the plane. The normal vector $\mathbf{n}$ of a plane is one that is orthogonal to every vector in the plane. If $\mathbf{n} = (a, b, c)$ and the point $(x_o, y_o, z_o)$ is on the plane, the equation of the plane is given by

$$a(x - x_o) + b(y - y_o) + c(z - z_o) = 0.$$

Of course, we can geometrically define a plane by two vectors and a point, or by three points, or by two lines, or by a line and a point. In each of these cases, how can you go about finding a normal vector to the plane?

# Chapter 2

**Functions and Level Curves**

We spent time discussing functions that take points in $\mathbb{R}^n$ and return points (or vectors) in $\mathbb{R}^m$. The values that can be plugged into a function are called its *domain* and all the possible outputs of a function are called its *range*. If $m = 1$, that is, if a function returns numbers, it is called a *scalar function*.

A scalar function of two variables $f(x, y)$ can be graphed using the $z$-axis to represent the function values. For each point $(a, b)$ in the domain of the function, the point $(a, b, f(a, b))$ is plotted in $\mathbb{R}^3$. Drawing all of these points will sketch out a surface in three dimensions.

We also discussed some *vector-valued functions*: functions that return vectors. For example, if the range of $\mathbf{F}(x, y)$ is vectors in $\mathbb{R}^2$, it is often useful to think of $\mathbf{F}(x, y)$ as defining two different functions: one to describe the first coordinate and one to describe the second.

$$\mathbf{F}(x, y) = (F_1(x, y), F_2(x, y))$$

For example, consider $\mathbf{F}(x,y) = -y\mathbf{i} + x\mathbf{j}$. We can rewrite this as $\mathbf{F}(x,y) = (-y,x)$. We can picture this as a *vector field* in two dimensions. At each point $(x,y)$ in the domain, plot the vector $\mathbf{F}(x,y)$. (We did this in class; you can see a similar function in Figure 2.7.)

For a function $f(x,y)$, the *level curve of value* $c$ is the points in the $x-y$ plane where $f(x,y) = c$. To find the level curves, you need to consider all the '$c$' that are possible outputs of $f(x,y)$ and solve algebraically for where $f(x,y) = c$. Plotting several of these curves in the domain of the function $f$ is known as a *contour diagram* of the function. This is a useful 2-dimensional way to picture the function (similar to the way a flat map of a country can describe the elevation above sea level). For functions $f(x,y,z)$ of three variables, considering the *level surfaces* – where $f(x,y,z) = c$ – is a useful 3-dimensional way to visualize the function. (See Figure 2.11.)

**Limits and Partial Derivatives**

Limits of continuous functions are found by plugging in the point. For example, if $F(x,y,z) = x^2 - 2y + z$:

$$\lim_{(x,y,z)\to(0,1,3)} F(x,y,z) = F(0,1,3) = 0 - 2 + 3 = 1.$$

Limits at a point that cannot be plugged into the function (for example, you may have a fraction where both the numerator and denominator are approaching 0) are more complicated. For functions of two variables, there are many, many ways to approach a point $(a,b)$. It is often easy to show that limits do not exist by considering one-dimensional limits along lines going through $(a,b)$. If there are two lines along which the limit is different, the limit cannot exist. (See examples from class and the book.)

Aside: Interestingly, it can sometimes be even harder to show that a limit does not exist. For example, the limit of the function $f(x,y) = \frac{x^2y}{x^4+y^2}$ does not exist as $(x,y) \to (0,0)$. Can you compute the limit along any line $y = mx$ that goes through the origin?

$$\lim_{\substack{(x,y)\to(0,0)\\ \text{along the line } y=mx}} \frac{x^2y}{x^4+y^2} = \lim_{x\to 0} \frac{mx^3}{x^4+m^2x^2} = 0.$$

The limit is 0 along any straight line! But what is the limit along the curve $y = x^2$?

Derivatives are defined as limits. Partial derivatives are just the idea that we can fix all the variables except one and take a regular derivative. For example, the partial derivative of $f(x,y)$ with respect to $x$ (at a point $(a,b)$) is defined as

$$\frac{\partial f}{\partial x}(a,b) = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h}.$$

Notice $y = b$ is fixed inside the limit. Graphically, we can think of fixing $y = b$ as restricting our domain to the points $(x,b)$ on that line. Then, the graph of $z = f(x,b)$ is a curve on the surface $z = f(x,y)$. The partial derivative is really giving the slope of the tangent line to this curve at the point $(a,b)$. (See Figure 2.46.)

We often organize all the partial derivatives of the function into a vector called the *gradient* of $f$. For example, if the function $f(x,y,z)$ is a function of three variables,

$$\nabla f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right).$$

**Chain Rule**

We have been using the usual chain rule when computing partial derivatives whenever it applies: for example, $\frac{\partial}{\partial x}(ye^{x^2}) = ye^{x^2}2x$, by holding $y$ constant and taking "the derivative of the exponential function times the derivative of the inside function." Remember in calculus, before you internalized this rule, you rewrote $f(u) = e^u$ where $u(x) = x^2$ and used the chain rule $\dfrac{\mathrm{d}f}{\mathrm{d}x} = \dfrac{\mathrm{d}f}{\mathrm{d}u}\dfrac{\mathrm{d}u}{\mathrm{d}x}$ to calculate the derivative of $e^{x^2}$.

For functions of more than one variable, we have a more general chain rule since a function of more than one variable can have more than one "inside function." For example if $f(u, v, w)$ has three functions $u(x, y)$, $v(x, y)$, and $w(x, y)$ plugged into it, we can find $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ using the general chain rule. The right way to compute, e.g., $\frac{\partial f}{\partial x}$ is to have three terms in the chain rule since $f$ depends on $u, v$, and $w$, each of which depends on $x$:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x}.$$

**Directional Derivatives**

For a function of two variables, instead of fixing the line $x = a$ or $y = b$ and taking a partial derivative of $f(x, y)$, we should be able to think of fixing *any* line $\mathbf{l}(t) = \mathbf{p} + t\mathbf{u}$ that goes through the point $p$. (Notice this is the parametric equation for the line we discussed in Chapter 1.) For the purposes of finding directional derivatives, we will make sure $\mathbf{u}$ is always a unit vector. We want to find $D_{\mathbf{u}}f(\mathbf{p})$, "the directional derivative of $f$ at the point $\mathbf{p}$ in the direction $\mathbf{u}$". (See Figure 2.61.) We only look at the values of $f$ on this line $\mathbf{l}(t)$, and take the derivative with respect to $t$; the following is the definition of *directional derivative*:

$$D_{\mathbf{u}}f(\mathbf{p}) = \left[\frac{d}{dt}\ f(\mathbf{p} + t\mathbf{u})\right]\Bigg|_{t=0}$$

Evaluating at the point $t = 0$ just makes sure that at the end, we're looking at the point $\mathbf{l}(0) = \mathbf{p}$ on the line. In practice, we rarely compute the directional derivative this way; from the chain rule, it is easy to see that

$$D_{\mathbf{u}}f(\mathbf{p}) = \nabla f(\mathbf{p}) \cdot \mathbf{u} = u_1\frac{\partial f}{\partial x}(\mathbf{p}) + u_2\frac{\partial f}{\partial y}(\mathbf{p})$$

This makes some intuitive sense: when the unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$, $\mathbf{u}$ is pointing partly in the $x$ direction and partly in the $y$ direction, the directional derivative is a linear combination of the partial derivatives in each direction.

An important fact about the directional derivative is that, at a point $\mathbf{p}$, the directional derivative is the greatest in the direction of the vector $\nabla f(\mathbf{p})$ (that is, in the direction of the gradient evaluated at the point). This means the gradient always points in the direction the function increases the most. Knowing this, think about how looking at a contour diagram tells you something about the gradient of the function! (See Figures 2.66, 2.67, 2.68 and the examples from lecture.)

### Tangent Planes and Linear Approximations

For a function $f(x, y)$ of two variables, the *tangent plane* at a point $(a, b)$ touches the surface at the point $(a, b, f(a, b))$ and contains the two tangent lines in the $y = b$ plane and in the $x = a$ plane whose slopes are given by $\frac{\partial f}{\partial x}(a, b)$ and $\frac{\partial f}{\partial y}(a, b)$. Therefore, the equation of this plane is

$$z = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b).$$

Notice that this is indeed a linear equation (since all of the partial derivatives are evaluated at a point)! Also, notice that a normal to the tangent plane at the point $(a, b)$ is given by $(\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b), -1)$.

We think of the tangent plane as the best plane we could use to approximate the surface near the point $(a, b)$ (Figure 2.49). In other words, the equation above gives the best *linear approximation* to the function $f(x, y)$ near the point $(a, b)$:

$$f(x, y) \approx f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$$

We can rewrite this equation as $\Delta f \approx \mathrm{d}f$ where $\Delta f = f(x, y) - f(a, b)$ and $\mathrm{d}f$ is the *differential* of the function at $(a, b)$: $\mathrm{d}f = f_x(a, b)(x - a) + f_y(a, b)(y - b)$ (or $\mathrm{d}f = f_x(a, b)\Delta x + f_y(a, b)\Delta y$).

# Chapter 4

### Higher Derivatives and Taylor Approximations

When we take a partial derivative of a function, we get another function back. Differentiating again, with respect to one of the variables, gives a *second-order derivative of the original function*. For example, if $f(x, y)$ is a function of two variables, there are four second-order derivatives.

$$f_{xx}(x, y) = \frac{\partial^2 f}{\partial x^2}; \quad f_{xy}(x, y) = \frac{\partial^2 f}{\partial y \partial x}; \quad f_{yx}(x, y) = \frac{\partial^2 f}{\partial x \partial y}; \quad \text{and} f_{yy}(x, y) = \frac{\partial^2 f}{\partial y^2}$$

This notation specifies the order in which the partial derivatives are taken; the conventions are defined on page 220 of the book. However, we don't usually worry about which order to take the partial derivatives in, since the main theorem of this section is that mixed partial derivatives of nice functions (e.g., $\frac{\partial^2 f}{\partial y \partial x}$ and $\frac{\partial^2 f}{\partial x \partial y}$) are always equal.

We also defined the Laplacian in this section: $\Delta u = u_{xx} + u_{yy}$ (for functions of two variables $u(x, y)$) Any function $u(x, y)$ that satisfies $\Delta u = 0$ is called *harmonic*.

We know how to find the best linear approximation of a function by using the equation of the tangent plane. *Taylor's formula* gives us a way to find the best quadratic (and higher) approximation of a function. For a function of two variables $f(x, y)$, the best *quadratic approximation* near a point $(a, b)$ is

$$\begin{aligned} f(x, y) \approx f(a, b) &+ f_x(a, b)(x - a) + f_y(a, b)(y - b) \\ &+ \frac{1}{2}\left[f_{xx}(a, b)(x - a)^2 + 2f_{xy}(a, b)(x - a)(y - b) + f_{yy}(a, b)(y - b)^2\right] \end{aligned}$$

Notice that all the partials are evaluated at the point, so the above is a polynomial (with numbers as the coefficients) in $x$ and $y$. Also notice that the linear part is just the equation of the tangent plane. The nonlinear part should remind you of Taylor's theorem for functions of one variable: for $g(x)$, the best quadratic approximation is $g(x) \approx g(a) + g'(a)(x-a) + \frac{g''(a)}{2}(x-a)^2$.

If we want to find the approximation to $f(a+\Delta x, b+\Delta y)$, we can use the above equation, which can also be rewritten in a more compact form, using vectors and matrices. Define $\mathbf{x_o} = (a,b)$, $\mathbf{x} = (x,y)$, and $\Delta\mathbf{x} = (\Delta x, \Delta y)$. Also, let $Hf$ be the matrix of second-order derivatives of $f$:

$$Hf = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

$Hf$ is called the *Hessian matrix* of $f$. Then, the quadratic equation can be written as

$$f(\mathbf{x}+\Delta\mathbf{x}) \approx f(\mathbf{x_o}) + \nabla f(\mathbf{x_o}) \cdot \Delta\mathbf{x} + \tfrac{1}{2}\Delta\mathbf{x}^T Hf(\mathbf{x_o})\Delta\mathbf{x}$$

If you work out all of the matrix multiplication in the above equation you will end up with exactly the quadratic approximation you saw in the previous equation! The most important thing introduced here is the Hessian matrix, which will reappear in the next section.

**Maximums and Minimums**

*Critical points* are all the points where $\nabla f(\mathbf{x}) = \mathbf{0}$. This means that to find critical points, we set each partial derivative of the function to 0, and be careful with the algebra when solving this system of equations!

We can think of critical points as points where the tangent plane is flat. (Remember that when $\nabla f = \mathbf{0}$, the directional derivative is 0 in every direction!) These points *may* be local maximums or local minimums of the function.

At a critical point $\mathbf{x_o}$ of a function $f(\mathbf{x})$, we can try to use the *second derivative test* to find out if $\mathbf{x_o}$ is a local max, local min, or a saddle point. It depends on the determinant of the Hessian matrix, **evaluated at the critical point**:

1. If $\det Hf(\mathbf{x_o}) > 0$, then $f$ is either a local maximum or minimum. To decide which, look at whether $f_{xx}(\mathbf{x_o}) > 0$ (local minimum) or $f_{xx}(\mathbf{x_o}) < 0$ (local maximum).
2. If $\det Hf(\mathbf{x_o}) < 0$, then the point $\mathbf{x_o}$ is a *saddle point*.
3. If $\det Hf(\mathbf{x_o}) = 0$, this tells us nothing about the point $\mathbf{x_o}$

To see why the entire Hessian matters (and not just the signs of $f_{xx}(\mathbf{x_o})$ and $f_{yy}(\mathbf{x_o})$, consider the function $f(x,y) = x^2 + y^2 + 3xy$. This function has a critical point at $(0,0)$, and $f_{xx}(0,0) = 2 > 0$ and $f_{yy}(0,0) = 2 > 0$; however, $\det Hf(0,0) = -5 < 0$ and $(0,0)$ is a saddle point. Although the curve $f(x,0)$ (above the x-axis) is concave up and the curve $f(0,y)$ (above the $y$-axis) is concave up, the curve above the line $y = -x$ bends downwards (since $f(x,-x) = -x^2$).

The *Extreme Value Theorem* states that on a closed and bounded set a continuous function must attain its maximum and minimum values. For example, the continuous function $f(x) = \sin x$ on the closed interval $[0,\pi]$ attains its minimum value 0 at both $x = 0$ and $x = \pi$ and attains its maximum value 1 at $x = \frac{\pi}{2}$.

Given a continuous function $f(x, y)$ and a (closed and bounded) set $D$, we can find the absolute maximum and minimum values by the following procedure:

1. Find all the critical points of $f(x, y)$ **inside the set** $D$. These points may be local minimums or local maximums, so check the value of the function at each of these points.

2. Check the values of $f(x, y)$ at every point on the boundary since the function might be biggest or smallest at a point on the edge of the set $D$.

Please see the handout on critical points and absolute extrema for some examples of finding absolute extrema.

**Divergence and Curl**

We define the divergence and curl of vector fields $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ (or $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$). Intuitively, the divergence is a measure of how much the vector field is spreading out at a point, and the curl is a measure of how much (and in what direction) it is rotating. If we believe this, we expect divergence to be a scalar and curl to be a vector. These interpretations of divergence and curl will be justified using integral theorems later.

Both div $\mathbf{F}$ and curl $\mathbf{F}$ are defined by thinking of the vector of partial derivative operators:

$$\nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

Notice that this gives us the usual gradient $\nabla f$, where $f$ is a scalar function: each component of this vector is the correct partial derivative operator acting on the function $f$. For vector fields $\mathbf{F}$ as above, we can take either the dot product of $\nabla$ with $\mathbf{F}$ or the curl of $\nabla$ and $\mathbf{F}$. These are our definitions of divergence and curl:

$$\text{div } \mathbf{F} = \nabla \cdot \mathbf{F}$$
$$\text{curl } \mathbf{F} = \nabla \times \mathbf{F}$$

If we write $\mathbf{F}$ in terms of its components: $\mathbf{F}(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z))$, then

$$\text{div } \mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$$

$$\text{curl } \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \left( \frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) \mathbf{i} + \left( \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) \mathbf{j} + \left( \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \mathbf{k}.$$

When computing curl $\mathbf{F}$ for a two-dimensional vector field$\mathbf{F}(x, y) = (F_1(x, y), F_2(x, y)$, we simply use the above formula, letting $F_3 = 0$. Since $F_3 = 0$ and since neither $F_1$ nor $F_2$ depends on $z$, the first two components of curl $\mathbf{F}$ will be zero, and the curl $\mathbf{F}$ will point in the $\mathbf{k}$-direction (orthogonal to any rotation of the vector field in the $x - y$ plane).

**Implicit Differentiation**

*Implicit surfaces* in $\mathbb{R}^3$ are surfaces described by an equation $F(x, y, z) = 0$ that cannot necessarily be solved for $z$ as a function of $x$ and $y$. Any point that satisfies the equation is a point on the surface.

Our main example is $x^2 + y^2 + z^2 = 1$. Near any point $(x_o, y_o, z_o)$ on the surface, (except for one with $z_o = 0$) we can *locally* think of the surface as a function of $x$ and $y$. If $z_o > 0$, we can consider $z = \sqrt{1 - x^2 - y^2}$; this surface contains the original point, and also all the points nearby. If $z_o < 0$, the point $(x_o, y_o, z_o)$, and all the points close to it, are on the surface described by $z = -\sqrt{1 - x^2 - y^2}$. However, if $z_o = 0$, no single function can describe all the points on the surface close to $(x_o, y_o, 0)$.

The *Implicit Function Theorem* states that for a surface $F(x, y, z) = 0$, we can locally solve for $z$ as a function of $x$ and $y$ near a point $(x_o, y_o, z_o)$ on the surface as long as

$$\frac{\partial F}{\partial z}(x_o, y_o, z_o) \neq 0.$$

For our example of the sphere, $\frac{\partial F}{\partial z} = 2z$, so this theorem tells us that we can locally solve for $z$ near any point $(x_o, y_o, z_o)$ with $z_o \neq 0$ (just as we already knew!)

When we want to be able to solve for two variables in terms of two others, we need a more general implicit function theorem. Our main example for this is polar coordinates, given by the equations $x = r\cos\theta$ and $y = r\sin\theta$. We'd like to be able to solve for $r$ and $\theta$ in terms of $x$ and $y$: From the implicit function theorem below, we can do this (locally) whenever $(x_o, y_o, r_o, \theta_o)$ satisfy the two equations and $r \neq 0$. (See lecture notes.)

A more general *Implicit Function Theorem* states that if we have a system of equations $F_1(x, y, u, v) = 0$ and $F_2(x, y, u, v) = 0$, we can locally solve (near a point $(x_o, y_o, u_o, v_o)$ that satisfies both equations) for $u$ and $v$ as functions of $x$ and $y$ whenever the determinant of the matrix of partial derivatives of $F_1$ and $F_2$ with respect to $u$ and $v$ at the given point does not equal 0; that is, when

$$\begin{vmatrix} \frac{\partial F_1}{\partial u}(x_o, y_o, u_o, v_o) & \frac{\partial F_1}{\partial v}(x_o, y_o, u_o, v_o) \\ \frac{\partial F_2}{\partial u}(x_o, y_o, u_o, v_o) & \frac{\partial F_2}{\partial v}(x_o, y_o, u_o, v_o) \end{vmatrix} \neq 0$$

Read the examples from lecture and in the textbook (Examples 4.68, 4.69). The implicit function theorem generalizes to three equations with six variables when you'd like to solve for three in terms of the other three (see Example 4.70), and so on for higher numbers of variables.

Whenever we think of $u$ as a function of $x$ and $y$ (even if we don't explicitly solve for $u(x, y)$), we should be able to compute the partial derivatives. If we want to find, for example, $\frac{\partial u}{\partial x}$, the strategy is to differentiate *both sides* of the equation(s) we know with respect to $x$ and then algebraically solve for $\frac{\partial u}{\partial x}$. The important thing to remember is the chain rule when we differentiate $u$ (for example $\frac{\partial}{\partial x}\left[xu^2\right] = u^2 + 2u\frac{\partial u}{\partial x}$) since we are thinking of $u$ as a function that depends on $x$!! Again, see the computation examples from lecture and the book.