Notes for Section 3.1

Iterated or *higher (order) partial derivatives* are partial derivatives of partial derivatives of multivariable functions. They are the multivariable analogs of the second, third, etc derivatives you know from single variable calculus. If $f : \mathbb{R}^n \to \mathbb{R}$, then the partial derivatives of f are themselves functions from $\mathbb{R}^n$ to $\mathbb{R}$. So each $\frac{\partial f}{\partial x_i}$ can have its own partial derivatives with respect to any of its input variables:

$$\frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_i}\right), \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_i}\right), \ldots, \frac{\partial}{\partial x_n}\left(\frac{\partial f}{\partial x_i}\right).$$

So f would have n partial derivatives, n^2 second partial derivatives, n^3 third partial derivatives, etc. Common notations for these are

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j \partial x_i} = (f_{x_i})_{x_j} = f_{x_i x_j} = D_{x_j}\left(D_{x_i} f\right) = D_{x_j x_i} f.$$

Notice that when using the $\frac{\partial^2 f}{\partial x_j \partial x_i}$ and the $D_{x_j x_i} f$ notation, the variables are listed from right to left according to the order in which the partial derivatives were taken. This is the result of the prefix notation and similar to how you would interpret a composite function $f \circ g(x)$. The order is from left to right when using the $f_{x_i x_j}$ notation because of the postfix notation. The longer versions of these notations using parentheses should give you a hint about the correct order.

An iterated partial derivative in which the differentiation is done with respect to several different variables is called a *mixed partial derivative*. For example, if $f : \mathbb{R}^3 \to \mathbb{R}$,

$$\frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial z \partial x}, \frac{\partial^3 f}{\partial x^2 \partial z}$$

are all examples of mixed partial derivatives, while

$$\frac{\partial^2 f}{\partial x^2}, \frac{\partial^3 f}{\partial y^3}, \frac{\partial^4 f}{\partial z^4}$$

are not.

Calculating iterated partial derivatives is straightforward as all you are doing is repeating (iterating) the process of finding the partial derivative of a multivariable function. And as we noted when we first encountered partial derivatives, calculating a partial derivative is essentially single variable differentiation. I think you will find it easy to work your way through Examples 1-3 in the textbook and understand the process.

Higher order partial derivatives encode information about the finer shape of the graph of f much the same way as higher order derivatives in single variable calculus do. Of course, things are a little more complex since we have so many more of these.

One interesting property of iterated partial derivatives is that in many cases, the mixed partial derivatives only depend on which variables the differentiation is done with respect to and not on their order. For example, if $f : \mathbb{R}^3 \to \mathbb{R}$, it is often the case that

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^3 f}{\partial x^2 \partial y} &= \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x^2} \\ f_{xyz} &= f_{xzy} = f_{yxz} = f_{yzx} = f_{zxy} = f_{zyx}. \end{aligned}$$

The relevant theorem is often referred to as Clairaut's Theorem or Schwarz's Theorem. I will include the proof here because the textbook glosses over a number of details. Perhaps that is alright as the details can distract from the main ideas, which are not that difficult.

Theorem. *Let $f : \mathbb{R}^n \to \mathbb{R}$ such that the mixed partial derivatives $\frac{\partial^2 f}{\partial x_j \partial x_i}$ and $\frac{\partial^2 f}{\partial x_i \partial x_j}$ both exist and are both continuous in some neighborhood of $x_0 \in \mathbb{R}^n$. Then*

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0).$$

Proof: For the sake of keeping the notation simpler, we will prove this in the special case $f : \mathbb{R}^2 \to \mathbb{R}$. The same argument works when $n > 2$ too. So we want to prove that

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$$

assuming that both mixed partials exist and are continuous near (x_0, y_0). Let Δx and Δy be small enough that all of the points of the rectangle whose opposite vertices are are (x_0, y_0) and $(x_0+\Delta x, y_0+\Delta y)$ in the xy-plane are in the neighborhood where both $\frac{\partial^2 f}{\partial y \partial x}$ and $\frac{\partial^2 f}{\partial x \partial y}$ are continuous. Call this rectangular region E. See Figure 3.1.1 in your textbook for what E looks like. Let

$$S(\Delta x, \Delta y) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0) - f(x_0, y_0 + \Delta y) + f(x_0, y_0).$$

Let g be the single variable function

$$g(x) = f(x, y_0 + \Delta y) - f(x, y_0).$$

In other words, we are treating y_0 and Δy as constant parameters while we let x vary. Notice that

$$S(\Delta x, \Delta y) = g(x_0 + \Delta x) - g(x_0).$$

Similarly, let h be the single variable function

$$h(y) = f(x_0 + \Delta x, y) - f(x_0, y).$$

Now we are treating x_0 and Δx as constant parameters while we let y vary. Notice that

$$S(\Delta x, \Delta y) = h(y_0 + \Delta y) - h(y_0).$$

Since the x-partial of f exists at every point $(x, y) \in E$, we know that

$$g'(x) = \frac{\partial f}{\partial x}(x, y_0 + \Delta y) - \frac{\partial f}{\partial x}(x, y_0)$$

also exists at every x in the interval $[x_0, x_0 + \Delta x]$. (If $\Delta x < 0$, just switch x_0 and $x_0 + \Delta x$ as the endpoints of the interval.) In other words, g is differentiable on $[x_0, x_0 + \Delta x]$. Hence g is also continuous on $[x_0, x_0 + \Delta x]$. By the Mean Value Theorem from single variable calculus, there is some point $\overline{x} \in (x_0, x_0 + \Delta x)$ such that

$$g(x_0 + \Delta x) - g(x_0) = g'(\overline{x})\Delta x.$$

So

$$S(\Delta x, \Delta y) = g(x_0 + \Delta x) - g(x_0) = g'(\overline{x})\Delta x = \left[\frac{\partial f}{\partial x}(\overline{x}, y_0 + \Delta y) - \frac{\partial f}{\partial x}(\overline{x}, y_0)\right] \Delta x.$$

We can now apply the Mean Value Theorem again to $\frac{\partial f}{\partial x}(\overline{x}, y)$ as a function of y. That it is a differentiable and continuous function of y on the interval $[y_0, y_0 + \Delta y]$ follows from the fact that the mixed partial $\frac{\partial^2 f}{\partial y \partial x}$ exists at $(\overline{x}, y) \in E$. So by the MVT, there exists some $\overline{y} \in (y_0, y_0 + \Delta y)$ such that

$$\frac{\partial f}{\partial x}(\overline{x}, y_0 + \Delta y) - \frac{\partial f}{\partial x}(\overline{x}, y_0) = \frac{\partial^2 f}{\partial y \partial x}(\overline{x}, \overline{y})\Delta y.$$

Hence

$$S(\Delta x, \Delta y) = \frac{\partial^2 f}{\partial y \partial x}(\overline{x}, \overline{y})\Delta y \Delta x.$$

So

$$\frac{\partial^2 f}{\partial y \partial x}(\overline{x}, \overline{y}) = \frac{S(\Delta x, \Delta y)}{\Delta x \Delta y}.$$

We can now do the same thing with h. There is some $\overline{\overline{y}} \in (y_0, y_0 + \Delta x)$ such that

$$S(\Delta x, \Delta y) = h(y_0 + \Delta y) - h(y_0) = h'(\overline{\overline{y}})\Delta y = \left[\frac{\partial f}{\partial y}(x_0 + \Delta x, \overline{\overline{y}}) - \frac{\partial f}{\partial y}(x_0, \overline{\overline{y}})\right] \Delta y.$$

And now there is some $\overline{\overline{x}} \in (x_0, x_0 + \Delta x)$ such that

$$\frac{\partial f}{\partial y}(x_0 + \Delta x, \overline{\overline{y}}) - \frac{\partial f}{\partial y}(x_0, \overline{\overline{y}}) = \frac{\partial^2 f}{\partial x \partial y}(\overline{\overline{x}}, \overline{\overline{y}})\Delta x.$$

Hence

$$S(\Delta x, \Delta y) = \frac{\partial^2 f}{\partial x \partial y}(\overline{\overline{x}}, \overline{\overline{y}})\Delta x \Delta y.$$

So

$$\frac{\partial^2 f}{\partial y \partial x}(\overline{\overline{x}}, \overline{\overline{y}}) = \frac{S(\Delta x, \Delta y)}{\Delta x \Delta y}.$$

Finally, we let $(\Delta x, \Delta y) \to (0, 0)$. This implies $\Delta x \to 0$ and $\Delta y \to 0$. Since $\overline{x}$ and $\overline{\overline{x}}$ are between x_0 and $x_0 + \Delta x$, both $\overline{x}$ and $\overline{\overline{x}}$ approach x_0 as $\Delta x \to 0$. Similarly, both $\overline{y}$ and $\overline{\overline{y}}$ approach y_0 as $\Delta y \to 0$. So as $(\Delta x, \Delta y) \to (0, 0)$, both $(\overline{x}, \overline{y})$ and $(\overline{\overline{x}}, \overline{\overline{y}})$ approach (x_0, y_0). Here is where we will use the continuity of the mixed partial derivatives:

$$\lim_{(\Delta x, \Delta y) \to (0,0)} \frac{S(\Delta x, \Delta y)}{\Delta x \Delta y} = \lim_{(\Delta x, \Delta y) \to (0,0)} \frac{\partial^2 f}{\partial y \partial x}(\overline{x}, \overline{y}) = \lim_{(\overline{x}, \overline{y}) \to (x_0, y_0)} \frac{\partial^2 f}{\partial y \partial x}(\overline{x}, \overline{y}) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

by the continuity of $\frac{\partial^2 f}{\partial y \partial x}$. Similarly,

$$\lim_{(\Delta x, \Delta y) \to (0,0)} \frac{S(\Delta x, \Delta y)}{\Delta x \Delta y} = \lim_{(\Delta x, \Delta y) \to (0,0)} \frac{\partial^2 f}{\partial x \partial y}(\overline{\overline{x}}, \overline{\overline{y}}) = \lim_{(\overline{\overline{x}}, \overline{\overline{y}}) \to (x_0, y_0)} \frac{\partial^2 f}{\partial x \partial y}(\overline{\overline{x}}, \overline{\overline{y}}) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$$

by the continuity of $\frac{\partial^2 f}{\partial x \partial y}$. We can now conclude that

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0).$$

Note that we can show that higher order mixed partial derivatives that are continuous do not depend on the order of the variables in the differentiation either by repeatedly applying the above theorem to switch the order of variables two at a time.

Examples 4 and 5 illustrate the equality of mixed partial derivatives on specific functions. The historical note at the end of the section tells you about how higher partial derivatives show up in some very prominent equations in physics. Solving such differential equations is definitely beyond the scope of our course, but you know enough about partial derivatives at this point to be able to verify if a given function satisfies such an equation. See Example 6 for how this can be done.

If you are curious to see a function whose mixed partial derivatives are not equal (because they are not continuous), exercise 3.1.32 has one for you. Working out the first and second partial derivatives of the function in this exercise takes some work because the function is a piecewise defined function. So its partial derivatives at $(0, 0)$ cannot be found simply by differentiating a formula using the usual rules of single variable differentiation. Limits must be used, although the limits, which are single variable limits, turn out to be quite easy to do.