Differentiation

Barsam Rahimi

November 2023

Note: A pre-requisite to this handout is the Limits and Continuity handout

1 Average and Instantaneous Rate Of Change

In the very beginning of our high school journey, we are introduced to the idea of a slope. We discuss the general equation of a linear function, $y = mx + b$ and how the m term is the slope of the graph of the function, meaning how steep it is. However, when discussing higher degree polynomials and other types of functions such as trigonometric, logarithmic, rational, exponential, etc. such quantity must be defined over some interval $[x_1, x_2]$ as the slope constantly changes across the function. If we draw a line that would connect the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$, it would be called a secant line (as by definition, a secant line touches the function at 2 different points). By determining the slope of this secant line, we can determine the Average Rate Of Change (AROC) of the function over that interval. The Average Rate Of Change can be determined using the following formula:

$$AROC \ or \ m = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Note that by multiplying both sides of the equation by the term $x_2 - x_1$, we obtain the Point-Slope form of a linear function which is of the form:

$$f(x_2) - f(x_1) = m(x_2 - x_1)$$

Whenever we are asked to determine the Average Rate of Change over some interval, we may simply obtain the coordinates of the two points and substitute into the AROC formula. For example, let us determine the average rate of change of the function $f(x) = sin(x)$ over the interval: $[\frac{\pi}{6}, \frac{3\pi}{2}]$. By substituting in the values, we find that the coordinates of the two points lying on the secant line are $(\frac{\pi}{6}, \frac{1}{3})$ and $(\frac{3\pi}{2}, -1)$. Now we substitute the coordinates into the AROC formula:

$$AROC = \frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{-1 - \frac{1}{3}}{\frac{3\pi}{2} - \frac{\pi}{6}} = \frac{\frac{-4}{3}}{\frac{4\pi}{3}} = \frac{1}{\pi}$$

Therefore, the Average Rate Of Change or the slope of the secant line over the given interval is $\frac{1}{\pi}$

In the Advanced Functions course, we are also introduced to the idea of Instantaneous Rate Of Change. This is the rate at which the function changes in exactly one point. For linear functions, this quantity is always equal to the Average Rate Of Change as they do not have any curvature. However, for higher degree polynomials such as quadratics and other continuous functions, due to the existence of curvature, this quantity is usually not equal to the AROC. The Instantaneous Rate Of Change describes the slope of a line that only touches the function at one point. This is called the tangent line. Now, this may seem impossible at a first glance, you may ask, well even if it was possible to find a line that only touches the function at one point, how can we possibly find the slope if we need 2 points but only have 1 ??
The answer lies within the idea of limits that was covered in the previous handout. To determine the Instantaneous Rate Of Change, we simply choose a point that is infinitely close to the point of tangency (the point in which the tangent line touches the function) and use the coordinates of that point in our slope formula. Hence, we define h, an infinitely small quantity which is near zero and so we define the coordinates of our points using h: The point of tangency which has the coordinates $(x, f(x))$ and the point that is infinitely close to it $(x + h, f(x + h))$. Using this approach, we may now substitute these two coordinates in the slope formula to determine the Instantaneous Rate Of Change of the function over the interval $[x, x + h]$:

$$IROC\ or\ m_t = \frac{f(x+h) - f(x)}{x + h - x} = \frac{f(x+h) - f(x)}{h}$$

Here m_t is read as "m sub t" and means the slope of the tangent. There is, however, one small problem with how this deduction came through, what is the value of h ? If we want the coordinates, we need numbers not a hypothetical variable. So how is it possible to use this formula if we cannot determine the value of h ? Depending on which school you have attended, you may be told - in Advanced Functions - by your teacher to set h equal to 0.01 or 0.001 or 0.0001 and in the end round your answer to the nearest whole number. The truth, is that the smaller we take h, the more accurate of a result we are going to get. This intuitively makes sense as the more we make h close to zero, the more close our two points are and hence we will obtain a more accurate value of the slope which brings us to the idea of the derivative.

2 The Definition of the Derivative

Differentiation is one of the most important tools to ever be invented by humanity. Its applications range from evaluating square roots by hand all the way to building the most complex of structures and inventions, predicting stock market values, describing the behavior of electric and magnetic fields, modelling harmonic motion and research in medicine for diseases such as cancer.

The derivative is a name that sounds fancy, difficult to understand and simply too abstract of an idea to grasp. However, if you have done advanced functions, you already know what the derivative is! The derivative is the same thing as the Instantaneous Rate Of Change which is the same thing as the slope of the tangent line. The slope of the tangent line at some point $(x_0, f(x_0))$ is defined as the derivative of the function f(x) at the point $(x_0, f(x_0))$. There is, however, one small difference. In the end of the previous section we discussed how the smaller h gets, the more accurate of a value we can obtain for the slope of the tangent line. Now that we are familiarized with the idea of limits, we can see that if we set h equal to zero, we get the most accurate result possible, one that does not need any rounding or estimation. However, if we look at the formula, we see that it is not possible to divide the whole expression by zero. Therefore, instead of setting h directly equal to zero, we take the limit such that h approaches to zero. Consequently, the derivative of a function at some point $(x, f(x))$ is defined as:

$$f'(x) \ or \ \frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Here, there are two notations that we see which both represent the derivative. The notation $f'(x)$ which is read as "f prime of x" was invented by Joseph-Louis Lagrange and is referred to as Lagrange notation. The other notation which may be easier to use in the later stages of Calculus, $\frac{df}{dx}$ is read as "df by dx" or simply "df, dx" which is another notation invented by Gottfried Willhelm Leibniz, one of the founders of Calculus, and it is referred to as Leibniz notation. Both of these notations have the same meaning and are equally valid so you may choose to use whichever one you prefer. Let us use this formula to determine the derivative of the function $f(x) = x^2$ at the point $x = 3$:

$$\frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h}$$

$$= \lim_{h \to 0} \frac{x^2 + 2hx + h^2 - x^2}{h} = \lim_{h \to 0} \frac{2hx + h^2}{h} = \lim_{h \to 0} 2x + h = 2x$$

$$\therefore \frac{df}{dx} = 2x \implies if \ x = 3, \ \frac{df}{dx} = 6$$

This means that the derivative of the function $f(x)$ at the point x = 3 is 6. Thus, the slope of the tangent or the Instantaneous Rate Of Change of the function at the point $(3, 9)$ is 6. Note that if

there is no specific point to substitute in the equation, we simply have a function that can give us the slope of the tangent line at any point we wish of the function $f(x)$.
Now, let us use the same formula to determine the derivative of the function $f(x) = x^3$

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{f(x+h)^3 - x^3}{h}$$

$$= \lim_{h\to 0} \frac{x^3 + 3hx^2 + 3h^2x + h^3 - x^3}{h} = \lim_{h\to 0} \frac{3hx^2 + 3h^2x + h^3}{h}$$

$$= \lim_{h\to 0} 3x^2 + 3hx + h^2 = 3x^2$$

$$\therefore f'(x) = 3x^2$$

After careful examination of these two examples you may be able to witness a pattern where every time we take the derivative of a polynomial function, we bring down the power of the x as a coefficient and we deduct one from the power of x. However, does this pattern also work with polynomials of different form ? To answer, let us take the derivative of another cubic function that also contains a quadratic, a linear and a constant term. Let $f(x) = x^3 - 2x^2 + 5x - 3$:

$$\frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{[(x+h)^3 - 2(x+h)^2 + 5(x+h) - 3] - [x^3 - 2x^2 + 5x - 3]}{h}$$

$$= \lim_{h\to 0} \frac{x^3 + 3hx^2 + 3h^2x + h^3 - 2(x^2 + 2hx + h^2) + 5x + 5h - 3 - x^3 + 2x^2 - 5x + 3}{h}$$

$$= \lim_{h\to 0} \frac{x^3 + 3hx^2 + 3h^2x + h^3 - 2x^2 - 4hx - 2h^2 + 5x + 5h - 3 - x^3 + 2x^2 - 5x + 3}{h}$$

$$= \lim_{h\to 0} \frac{3hx^2 + 3h^2x + h^3 - 4hx - 2h^2 + 5h}{h} = \lim_{h\to 0} 3x^2 + 3hx + h^2 - 4x - 2h + 5$$

$$= 3x^2 - 4x + 5$$

$$\therefore \frac{df}{dx} = 3x^2 - 4x + 5$$

Now, what can we see from this result ? All that tedious calculation in order to get the derivative of such a simple function is not very efficient and useful. Let us analyse and compare the derivative of the function with the original. We can see that the cubic term in the original function has turned into a quadratic term. We have simply brought down the exponent and turned it into the coefficient. Then, we deduct one from the exponent of x which turns it into a quadratic function. The same thing happens with the linear term, except, since there was already a coefficient of -2, it gets multiplied by the exponent that we bring down so instead of it being $2x$, it is multiplied by -2 which gives us $-4x$. Regarding the linear term, since the power x is 1, we bring it down which makes no difference and then we deduct one from the power of x. So it is now raised to the power of 0. However, since we know that anything to the power of 0 is 1, we can simply replace x^0 with 1 which gives us the term 5 in the derivative. This rule, however, is not applicable to the term -3. Here we cannot assume that the power of x is 0 and so we have to deduct one from it which turns it into -1. Instead, I ask you to think of a the graph of the line y = -3. If we look at the graph of this function, we see that the slope across all points of this function is 0 because it does not change and is a horizontal line. Therefore, this term cancels out to 0.

3 Power Rule

The pattern that we examined so carefully in the previous section shows us that there is some sort of rule that is correct for differentiation of all polynomial functions. Hence, let us examine the general case by examining the derivative of the function $f(x) = kx^n$, where k is some constant and n is a natural number or in mathematical notation, $n \in \mathbb{N}$. Note that you do not need to understand this proof, however, it would be helpful to try to get a sense of it:

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{k(x+h)^n - kx^n}{h}$$

$$= \lim_{h\to 0} \frac{k\binom{n}{0}x^n h^0 + k\binom{n}{1}x^{n-1}h^1 + k\binom{n}{2}x^{n-2}h^2 + k\binom{n}{3}x^{n-3}h^3 + \cdots + \binom{n}{n-1}kx^1k^{n-1} + k\binom{n}{n}x^0h^n - kx^n}{h}$$

$$= \lim_{h\to 0} \frac{k\binom{n}{1}x^{n-1}h^1 + k\binom{n}{2}x^{n-2}h^2 + k\binom{n}{3}x^{n-3}h^3 + \cdots + \binom{n}{n-1}kx^1k^{n-1} + k\binom{n}{n}x^0h^n}{h}$$

$$= \lim_{h\to 0} \left[k\binom{n}{1}x^{n-1} + k\binom{n}{2}x^{n-2}h^1 + k\binom{n}{3}x^{n-3}h^2 + \cdots + \binom{n}{n-1}kx^1h^{n-2} + k\binom{n}{n}x^0h^{n-1}\right]$$

$$= k\binom{n}{1}x^{n-1} = \frac{kn!}{1!(n-1)!}x^{n-1} = \frac{kn(n-1)!}{(n-1)!}x^{n-1} = knx^{n-1}$$

$$\therefore\ f(x) = kx^n, \implies f'(x) = knx^{n-1}$$

This rule is commonly referred to as the power rule and is true for all polynomials of degree 1 or higher. Let us apply this rule to find the derivative of several polynomial functions:

$$f(x) = 4x^5 - 3x^2 + \frac{1}{4}x - \frac{3^\pi}{e} \implies f'(x) = 20x^4 - 6x + \frac{1}{4}$$

$$f(x) = \frac{\sqrt{\pi}}{2\varphi}x^{\frac{\varphi}{\sqrt{\pi}}} - 3\sqrt{2}e^{i\frac{\pi}{\sqrt{e}}} + \frac{1}{2}x^2 + \varphi \implies \frac{df}{dx} = \frac{1}{2}x^{\frac{\varphi}{\sqrt{\pi}}-1} + x$$

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1x^1 + a_0$$

$$\implies f'(x) = a_n n x^{n-1} + a_{n-1}(n-1)x^{n-2} + a_{n-2}(n-2)x^{n-3} + \cdots + a_2 x^1 + a_1$$

The very last example states the general case for polynomials of degree 1 or higher with consideration that the derivative of any constant is zero (since it will be a horizontal line of slope 0).

Now, let us do an examination of applying the power rule to functions where the power of x is not just a natural but a real number such as $\frac{1}{3}$ or -2, etc. The proof to this theorem follows from the axioms and the same process of the proof of power rule.

$$f(x) = \sqrt{x}$$

$$f'(x) = \lim_{h\to 0} \frac{\sqrt{x+h}-\sqrt{x}}{h} = \lim_{h\to 0} \frac{\sqrt{x+h}-\sqrt{x}}{h} \cdot \frac{\sqrt{x+h}+\sqrt{x}}{\sqrt{x+h}+\sqrt{x}}$$

$$= \lim_{h\to 0} \frac{(x+h)-x}{h(\sqrt{x+h}+\sqrt{x})} = \lim_{h\to 0} \frac{h}{h(\sqrt{x+h}+\sqrt{x})} = \lim_{h\to 0} \frac{1}{\sqrt{x+h}+\sqrt{x}}$$

$$= \frac{1}{\sqrt{x}+\sqrt{x}} = \frac{1}{2\sqrt{x}} = \frac{1}{2}x^{\frac{-1}{2}}$$

$$\therefore f(x) = \sqrt{x} \implies f'(x) = \frac{1}{2\sqrt{x}}$$

Since we have proven that this exists for $\frac{1}{2}$, applying the same formula to $f(x) = \sqrt[3]{x}$ we find that the derivative, $f'(x) = \frac{1}{3}x^{\frac{-2}{3}}$. Proving this theorem for $f(x) = \sqrt[n]{x}$ will be left as an exercise for the reader. Since this follows from the previous case, we can conclude that:

$$f(x) = kx^n, n \in \Re \implies \frac{df}{dx} = knx^{n-1}$$

4 Sums and Difference Rule

The sum and difference rule is a rule that we previously used as a matter of fact. This rule states that if $f(x)$ is the sum or difference of two other functions, $P(x)$ and $Q(x)$, then the derivative of f is equal to the sum or difference of the derivatives of the two functions, or in algebraic notation:

$$f(x) = P(x) \pm Q(x) \implies f'(x) = P'(x) \pm Q'(x)$$

Proof:

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h\to 0} \frac{P(x+h) \pm Q(x+h) - P(x) \mp Q(x)}{h}$$

$$= \lim_{h\to 0} \frac{P(x+h)-P(x)}{h} \pm \lim_{h\to 0} \frac{Q(x+h)-Q(x)}{h} = P'(x) \pm Q'(x)$$

$$\therefore f(x) = P(x) \pm Q(x) \implies f'(x) = P'(x) \pm Q'(x)$$

5 Product and Quotient Rule

The Product and Quotient Rule, similar to the case of the sums and difference rule are used when a function is a product or quotient of two or more functions. It states that if $f(x) = g(x) \cdot k(x)$, then $f'(x) = g'(x) \cdot k(x) + g(x)k'(x)$

Proof:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{g(x+h) \cdot k(x+h) - g(x) \cdot k(x)}{h}$$

Now, we use a "trick" of adding "0" which we will use to simplify the expression into something where we can group terms. Hence, we add the term $g(x) \cdot k(x+h) - g(x) \cdot k(x+h)$ to the numerator to get the following:

$$= \lim_{h \to 0} \frac{g(x+h) \cdot k(x+h) - g(x) \cdot k(x) + g(x) \cdot k(x+h) - g(x) \cdot k(x+h)}{h}$$

$$= \lim_{h \to 0} \left[g(x) \cdot \frac{k(x+h) - k(x)}{h} + k(x+h) \cdot \frac{g(x+h) - g(x)}{h} \right] = g'(x) \cdot k(x) + g(x)k'(x)$$

The same can be applied to prove quotient rule which states that for a function, $f(x) = \frac{u(x)}{v(x)}$ where f is the quotient of u and v, the derivative of the function, $\frac{df}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$.

Proof:

$$f(x) = \frac{u}{v} \implies v \cdot f = u$$

$$\frac{du}{dx} = \frac{dv}{dx} f + \frac{df}{dx} v$$

$$\therefore \frac{\frac{du}{dx} - \frac{dv}{dx} f}{v} = \frac{df}{dx}, \; f = \frac{u}{v}$$

$$\implies \frac{df}{dx} = \frac{\frac{du}{dx} - \frac{dv}{dx} \cdot \frac{u}{v}}{v} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$$

6 Chain Rule

Perhaps, the most important rule of differentiation, the chain rule is the key to differentiating every function we see. Using the rules that we have learned so far, we are able to differentiated many functions. However, we will not be able to differentiate composite functions or more complicated ones such as $f(x) = sin(x^2) - \frac{1}{e^{cos(x)}}$ where f is a mix of multiple functions. To do so, we must know the chain rule. The chain rule states that if a function, f, is a composite of another function, g, then the derivative of f would be equal to the derivative of f with respect to g times the derivative of g. Or in derivative notation; $u(x) = f(g(x))$, then $u'(x) = f'(g(x)) \cdot g'(x)$. Here is the proof for this theorem:

Proof:

$$u(x) = f(g(x))$$

$$\implies \frac{du}{dx} = \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{h}$$

$$= \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{h} \cdot \frac{g(x+h) - g(x)}{g(x+h) - g(x)} = \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \cdot \frac{g(x+h) - g(x)}{h}$$

$$= \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \cdot \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} = f'(g(x)) \cdot g'(x)$$

$$\therefore \frac{du}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$$

In other words, when taking the derivative of a composite function, take the derivative of the function with respect to g(x), then go back and mulitply by the derivative of the inside. Combining this, extremely important, rule with all the other differentiation rules that we have learned, we can find the derivative of almost every function we can come up with.

7 Derivatives of Trigonometric Functions

In order to find the derivative of the trigonometric functions, we will be using something called a "Taylor Expansion", a concept taught in Calculus II of university. Therefore, do not worry if you do not understand the next few proofs. However, we will only use Taylor Expansion for finding the derivative of sin(x) and cos(x). The rest we will use the already established rules of differentiation. In order for you to understand this proof, we will also do this proof using compound angle formulas.
If $f(x) = \sin(x)$, then the derivative, $f'(x) = \cos(x)$.
Proof:

$$f(x) = \sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

Applying the power rule to each term of this series, we will get the following series:

$$f'(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = \cos(x)$$

An alternate approach is using compound angle formulas:

$$f'(x) = \lim_{h\to 0} \frac{\sin(x+h) - \sin(x)}{h} = \lim_{h\to 0} \frac{\sin(x)\cos(h) + \sin(h)\cos(x) - \sin(x)}{h}$$

$$\lim_{h\to 0} \left[\sin(x) \cdot \frac{\cos(h)}{h} + \cos(x) \cdot \frac{\sin(h)}{h}\right] = 0 \cdot \sin(x) + 1 \cdot \cos(x) = \cos(x)$$

The reason why the two fractions were evaluated to a 0 and a 1 will be discussed in the next handout which is the Applications of Derivatives handout. Let us use the same technique for $f(x) = \cos(x)$:

$$f(x) = \cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$$

$$f'(x) = -x + \frac{x^3}{3!} - \frac{x^5}{5!} + \frac{x^7}{7!} - \dots = -\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots\right)$$

$$= -\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = -\sin(x)$$

Or applying the first principle definition of the derivative:

$$f'(x) = \lim_{h\to 0} \frac{\cos(x+h) - \cos(x)}{h} = \lim_{h\to 0} \frac{\cos(x)\cos(h) - \sin(x)\sin(h) - \cos(x)}{h}$$

$$= \lim_{h\to 0} \frac{\cos(x)\cos(h) - \sin(x)\sin(h) - \cos(x)}{h} = \lim_{h\to 0} \cos(x)\frac{\cos(h)}{h} - \lim_{h\to 0} \sin(x)\frac{\sin(h)}{h}$$

$$= 0 \cdot \cos(x) - 1 \cdot \sin(x) = -\sin(x)$$

For now, ignore why the limits of the expressions evaluated to these numbers, we will explore this in the next chapter when we learn about L'Hôpital's rule (Yes, the name of the rule is actually the hospital rule, named after Guillaume de l'Hôpital, the french mathematician). Now let us apply this knowledge to determine the derivatives of the functions $\tan(x), \sec(x), \cot(x)$ and $\csc(x)$:

$$\frac{d}{dx}\tan(x) = \frac{d}{dx}\frac{\sin(x)}{\cos(x)} = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} = \sec^2(x)$$

$$\frac{d}{dx}\sec(x) = \frac{d}{dx}(\cos(x))^{-1} = -(\cos(x))^{-2} \cdot (-\sin(x)) = \frac{\sin(x)}{\cos^2(x)} = \sec(x)\tan(x)$$

$$\frac{d}{dx}\cot(x) = \frac{d}{dx}\frac{1}{\tan(x)} = -(tan(x))^{-2} \cdot \sec^2(x) = -\frac{\cos^2(x)}{\sin^2(x)} \cdot \frac{1}{\cos^2(x)} = \frac{-1}{\sin^2(x)} = -\csc^2(x)$$

$$\frac{d}{dx}\csc(x) = \frac{d}{dx}(\sin(x))^{-1} = -(\sin(x))^{-2} \cdot (\cos(x)) = -\frac{\cos(x)}{\sin^2(x)} = -\csc(x)\cot(x)$$

8 Advanced: Implicit Differentiation

Now that we have learned how to differentiate most differentiable functions, we can learn implicit differentiation. Let us look back at one of the first lessons of grade 11: the vertical line test. We learned that if a graph does not pass the vertical line test, it is not a function but a relation between variables x and y. Although it may seem counter-intuitive, relations are also differentiable. We may differentiate relations to determine rates of change of variables with respect to each other. This takes us into the many applications of derivatives such as related rates. More on this in the next handout. Here is how implicit differentiation works: we simply treat the relation like a function that we are taking the derivative of with respect to some variable x. The thing to keep in mind is that we are always trying to isolate the equation, not for y, but for the derivative of y. For example, let us implicitly differentiate the following relation:

$$y^2 + x^2 = 1$$

The first step is to take the derivative of both sides with respect to x, as we want to see the relationship between the rate of change of y, and x.

$$\frac{d}{dx}\left(y^2 + x^2\right) = \frac{d}{dx}1$$

$$\frac{d}{dx}y^2 + \frac{d}{dx}x^2 = 0 \implies 2y \cdot \frac{dy}{dx} + 2x \cdot \frac{dx}{dx} = 0$$

Now we have what we are solving for, $\frac{dy}{dx}$ and the value of $\frac{dx}{dx}$ is 1 since the derivative of some variable with respect to itself is 1. Therefore we get the following expression:

$$2y\frac{dy}{dx} + 2x = 0 \implies \frac{dy}{dx} = \frac{x}{y}$$

Note that it is OK to have y as a part of our answer when implicitly differentiating. Now, let us geometrically interpret what this result tells us. As you may have noticed, we have differentiated the equation of a circle with radius of length 1. The derivative of this equation, tells us that if we draw a tangent line to the circle at some point (x, y), the relationship $\frac{dy}{dx} = \frac{x}{y}$ tells the slope of this tangent line or the instantaneous rate of change of the circle at this point. Let us do this to another relation:

$$y^2 + x^2 = \sin(xy)$$

$$\implies 2y\frac{dy}{dx} + 2x\frac{dx}{dx} = \cos(xy) \cdot \frac{d}{dx}(xy)$$

$$2y\frac{dy}{dx} + 2x = \cos(xy) \cdot (y + x\frac{dy}{dx})$$

$$2y\frac{dy}{dx} - x\cos(xy)\frac{dy}{dx} = y\cos(xy) - 2x$$

$$\frac{dy}{dx} \cdot (2y - x\cos(xy)) = y\cos(xy) - 2x \implies \frac{dy}{dx} = \frac{y\cos(xy) - 2x}{2y - x\cos(xy)}$$

9 Derivatives of Exponential Functions

Let us now use the first principle definition of the derivative to differentiate an exponential function, $f(x) = a^x$ where a is some real number:

$$f'(x) = \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} = \lim_{h \to 0} \frac{a^x \cdot (a^h - 1)}{h} = a^x \cdot \lim_{h \to 0} \frac{a^h - 1}{h} = a^x \cdot \ln a$$

Again, how the second term turns into ln a is a concept that we will discuss in future chapters where we properly define the logarithm using integrals. For now, just accept this as a convention and "go with it". Using this definition, we can determine the derivative of one of the most useful functions which will keep re-occurring in all branches of mathematics, $f(x) = e^x$ where e is Euler's number.

$$f'(x) = e^x \cdot \ln e = e^x$$

This discovery is truly a marvel; A function whose derivative is itself! Later on, when we explore the concept of Integration By Parts, you will see that this is an extremely useful property when integrating extremely complicated functions that include the exponential function with base e. And it is worthwhile mentioning that when the exponent is a function itself, then we must multiply the entire expression by the derivative of that function due to chain rule. Or in mathematical notation:

$$f(x) = e^u \implies \frac{df}{dx} = e^u \frac{du}{dx}$$

If you wish to not read the next chapter which is advanced, you will need to know that the derivative of the natural logarithm is the following, if u is some function of x, then:

$$f(x) = ln(u) \implies f'(x) = \frac{u'}{u}$$

10 Advanced: Derivative of Inverse Functions

The derivative of inverse functions is one of the most important and useful formulas to know for integration and later stages of calculus involving partial derivatives and integrals of rational functions. However, for now we stick to determining the derivative of the inverse of some function $f(x)$ whose inverse is $f^{-1}(x)$. In order to achieve this, we must make use of the fundamental properties of inverse functions: that any function operating on its inverse is equal to the variable of the function, now matter what type of function it is we are dealing with. Hence:

$$(fof^{-1})(x) = x \; or \; f(f^{-1}(x)) = x$$

Now we use implicit differentiation:

$$\frac{d}{dx}f(f^{-1}(x)) = 1 \implies f'(f^{-1}(x)) \cdot \frac{d}{dx}f^{-1}(x) = 1$$

$$\therefore \frac{d}{dx}f^{-1}(x) = \frac{1}{f'(f^{-1}(x))}$$

Using this we can find the derivative of functions like the logarithm. Let us use this technique to find the derivative of a logarithmic function with base b where b is some real number such that $f^{-1}(x) = \log_b x$:

$$f(x) = b^x \implies f^{-1}(x) = \log_b x$$

$$\frac{d}{dx}f^{-1}(x) = \frac{1}{f'(\log_b x)} = \frac{1}{b^{\log_b x}\ln b} = \frac{1}{x \ln b}$$

If the base of the logarithm is u, where u is some function of x, then:

$$\frac{d}{dx}\log_b u = \frac{u'}{u \cdot \ln b}$$

11 Advanced: Derivatives of Inverse Trigonometric Functions

Now that we have learned most differentiation techniques, there is one more important one to explore which we can use to determine the derivatives of inverse trigonometric functions. Let us see how implicit differentiation is used by going through an example. Let's try to find the derivative of the the inverse of the sine function, $y = \sin^{-1}(x)$:

$$y = \sin^{-1}(x) \implies x = \sin y$$

Now we apply implicit differentiation to determine the derivative of y:

$$\frac{d}{dx}\sin y = 1 \implies \cos y \cdot \frac{dy}{dx} = 1$$

$$\frac{dy}{dx} = \frac{1}{\cos y}$$

$$\cos^2 y + \sin^2 y = 1 \implies \cos y = \sqrt{1 - \sin^2 y}$$

$$\therefore \frac{dy}{dx} = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}}$$

Let us do the same for $y = \cos^{-1} x$. The rest will be left as an exercise. Note: In integration, the most useful ones tend to be inverse of cosine, sine, tangent and secant.

$$y = \cos^{-1} x \implies \cos y = x$$

$$-\sin y \cdot \frac{dy}{dx} = 1 \implies \frac{dy}{dx} = \frac{-1}{\sin y} = \frac{-1}{\sqrt{1-x^2}}$$

12 Higher Order Derivatives

As you may have thought of already, when we differentiate a function, we get a another function which tells us the slope of the tangent to the initial function at every point. However, derivatives themselves have derivatives. You may differentiate the derivative of some function and it will give you the slope of the tangent of that function at some point. The derivative of the derivative is called the "second derivative" of the function. The notations for the second derivative are the same:

$$\frac{d}{dx}\frac{df}{dx} = \frac{d^2 f}{dx^2} = f''(x)$$

The Leibniz notation is simply read as "the second derivative with respect to x" whereas the Lagrange notation is read as "f double prime of x". The same can be done for the "third derivative".

$$\frac{d}{dx}\frac{d^2 f}{dx^2} = \frac{d^3 f}{dx^3} = f'''(x)$$

Let us describe this for a general case of the nth derivative. Note that the zeroth derivative of f is equal to f itself.

$$\frac{d}{dx} \cdots \frac{d^{n-1} f}{dx^{n-1}} = \frac{d^n f}{dx^n} = f^{(n)}(x)$$

Note that for Lagrange notation, normally after the third or fourth, a number replaces the primes so the tenth derivative would be written as $f^{10}(x)$ in Lagrange notation.

13 Application to Physics

When studying the motion of particles along a path, we learn about quantities called position, velocity and speed, and acceleration. During the grade 11 physics course, we learn that these are connected through the area under the graphs or the slope of the graph in some interval or point. Now that we have learned about derivatives, we can extended these definitions to describe this relationship using Calculus. If we have some function $s(t)$ where t is the time and s is the position, then the derivative of this function will give us a function, $v(t)$ that models the speed of the particle along the same path.

$$\frac{d}{dt}s(t) = \frac{ds}{dt} = s'(t) = v(t)$$

Similarly, the acceleration of a particle along a path is given by differentiating the function describing its speed along that path.

$$\frac{d}{dt}v(t) = \frac{dv}{dt} = v'(t) = a(t)$$

Another, less known, quantity is Jerk. This quantity describes the rate of change of a particle's acceleration along some path in space. Although we do not see this very often in high school physics, it is very important in topics such as Quantum Mechanics and General Relativity.

$$\frac{d}{dt}a(t) = \frac{da}{dt} = a'(t) = j(t)$$

Relating these quantities to the second and third derivatives, we can see the following relationships:

$$j(t) = \frac{da}{dt} = \frac{d^2v}{dt^2} = \frac{d^3s}{dt^3}$$

$$s(t) = \int v(t)dt = \iint a(t)d^2t = \iiint j(t)d^3t$$

Do not worry about the integral sign here, all that you need to know is that the integral is the opposite of the derivative and therefore this relationship is true. We will discuss integrals in great detail in a later chapter.