Week 2: Derivative

AVERAGE RATE OF CHANGE

1.1. The **average rate of change** of a function f on the interval $[a, b]$ is

$$\frac{f(b) - f(a)}{b - a}$$

if b is larger than a. Unlike the **instantaneous rate of change** which is a limit and the slope of the function at a point, the average rate of change is defined even for non-smooth functions. The name average rate of change will be justified later because it is the average value over all derivatives $f'(x)$ on the interval $[a, b]$ for nice functions: the relation $f(b) - f(a) = \int_a^b f'(x)\, dx$ will be a consequence of the **fundamental theorem of calculus** which is one of the high lights of single variable calculus. The **derivative** $f'(x)$ will be defined as a limit of $[f(x + h) - f(x)]/h$ when h goes to zero and is also called the **instantaneous rate of change**.
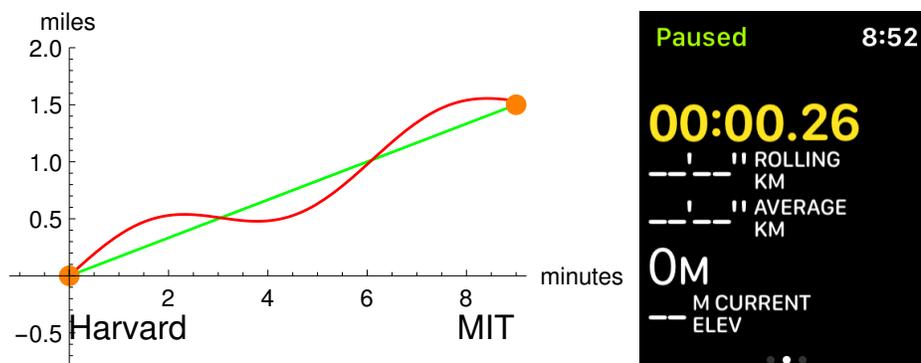

FIGURE 1. A possible Harvard MIT run can be quantified by the function giving the distance from Harvard as a function of time. The picture to the right is a screen shot of a smart watch app. It displays either the **rolling pace** or the **average pace**. This translates into **instantaneous rate of change** or the **average rate of change**. In the Harvard-MIT case, the average rate of change would show 6 minutes per mile, which means 10 miles per hour or 16km per hour.

1.2. The average rate of change also is defined if the function is known only at a few points. Assume you run the 1.5 miles from **Harvard to MIT** along the Charles and you clock yourself with 9 minutes. Then your average speed is $1.5/9 = 1/6$ miles per minute. This means you run with an average speed of 10 Miles per hour, which is 16 km/h. The instantaneous rate of change varies over the run. You might run slower before the Boston university bridge as it goes up a bit but then run faster near the de Wolfe Boathouse, where it goes down. If you have a smart watch, it measures average rates of changes by looking up your location using **GPS satellites** every once in a while. It then displays the rolling pace, or the average pace, which determines the average rate of change over the entire run so far.

INSTANTANEOUS RATE OF CHANGE

1.3. The **instantaneous rate of change** can be defined as a limit of average rates of changes when the interval $[x, x + h]$ gets smaller and smaller, but it can also be understood intuitively and geometrically as the **slope of the tangent** at the graph. You all have already computed slopes and seen for example that for the function $f(x) = x^2$, the slope at a point x is $2x$. This can be justified when looking at the average rate of change $[f(x + h) - f(x)]/h = [(x + h)^2 - x^2]/h$. If we foil out this expression and simplify, it becomes $[x^2 + 2hx + h^2 - x^2]/h = 2x + h$. One can see that if h goes to zero, the limit $2x$ appears. This value $2x$ is called the **derivative** of f at the point x and denoted $f'(x)$.
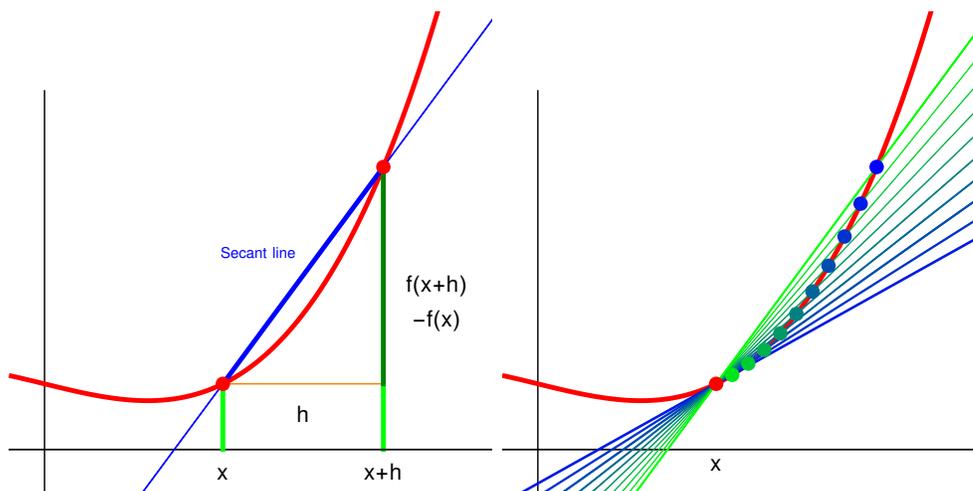

FIGURE 2. The average rate of change is the slope of the secant connecting two points on the graph. The instantaneous rate of change is a limiting value when the intervals get smaller and smaller. It is the slope of the tangent at x.

1.4. Historically, the notion of derivative needed a lot of time to develop. One of the first who investigated the notion seriously was **Zeno of Elea** who was born around 490 BC, just around the time, when Pythagoras (570-495 BC) died. Already Aristotle objected to the paradoxa that Zeno relied on the false supposition that time is composed of indivisible "nows" or "instants". Still, one must state that the question **"what is time?"** is still not answered in a definite way. When we talk about space and

time, we assume to have a nice calculus set-up, while in reality, we know that space and time looks on the **Planck scale** is described by a different kind of mathematics. The questions of Zeno remain of interest today. It might well be that the notion of **instantaneous rate of change** is an idealization and that in reality, time is tangled up in a complicated way with space. [1]

FIGURE 3. Part of the renaissance fresco **"School of Athens"** by Raphael painted around 1510 shows Zeno of Elea to the very right. He escapes in terror after realizing that the arrow paradox, that forbids motion does not work. He might get hit by an arrow.

1.5. The nice thing about the average rate of change is that it does not use any notion of limit. Calculus without limits is also called **quantum calculus**. [2] In this frame work, we have a good notion of average rate of change for any function, it does not have to be smooth. The function $f(x) = |\sin(x)|$ is far from smooth but we can compute the average rate of change for any interval $[a, b]$. The average rate of change also makes sense in discrete settings, when we deal with **data**. For example, if we look at the **primes** $(2, 3, 5, 7, 11, 13, 17, ...)$ and consider the prime function $p(n)$ which gives the n'th prime then the rate of change is naturally be given by taking neighboring points and look at $p(n+1) - p(n)$. This is a difficult function. The **prime twin problem** for example asks whether there are infinitely many n for which $p(n+1) - p(n)$ is equal to 2.

THE DERIVATIVE

1.6. The **derivative** of a function $f(x)$ at a point x is the limit of the expression

$$\frac{[f(x+h) - f(x)]}{h}$$

when h goes to zero. It is denoted by $f'(x)$ and geometrically measures the **slope** of the tangent line at the point. The function f is called **differentiable at** x, if this limit exists. If the derivative is positive, the function is **increasing** at x, if the derivative is negative, the function is **decreasing** at x. If the derivative is zero, the function is

---

[1] See the book "Zeno's paradox'" (2008) by Joseph Mazur or watch "Ant man and the Wasp"

[2] My blog www.quantumcalculus.org is dedicated to this kind of calculus.

stationary there. As we will see later these are candidates for maxima or minima. The average rate of change on an interval $[x, x+h]$ approximates the derivative at x if h is small. If he **secant line** connecting $f(x)$ and $f(y)$ is always above the graph for x, y in $[a, b]$ the function is **concave up** on $[a, b]$. If the secant line connecting $(x, f(x))$ and $(y, f(y))$ is below the graph for any x, y in the interval $[a, b]$, the function is called **concave down** on $[a, b]$.

1.7. Before we have differentiation rules, we need to compute a limiting procedure. From the Binomial formula, one sees for $f(x) = x^n$, that $f'(x) = nx^{n-1}$. For example, for $f(x) = x^4$ we have $f(x + h) = x^4 + 4x^3h + 6x^2h^2 + 4xh^3 + h^4$ so that $[f(x + h) - f(x)]/h = 4x^3 + 6xh^2 + 4xh^2 + h^3$ which is $4x^3$ when h is zero. This can then be extended to all polynomials like $f(x) = x^4 + 3x^2 - 6x + 1$. The derivative is then $f'(x) = 4x^3 + 6 - 6$. This is justified by the fact that the average rate of change of a sum of functions $f + g$ on $[a, b]$ is the sum of the average rate of changes of f and g.

1.8. For root functions like $f(x) = \sqrt{x}$, the computation of the limit $[\sqrt{x+h} - \sqrt{x}]/h$ needs a **trick**: multiply both the nominator and denominator with $[\sqrt{x+h} + \sqrt{x}]$ to get the average rate of change on $[0, h]$ as

$$\frac{[\sqrt{x+h} - \sqrt{x}][\sqrt{x+h} + \sqrt{x}]}{h[\sqrt{x+h} + \sqrt{x}]}$$

which simplifies with the formula $(a - b)(a + b) = a^2 - b^2$ to

$$\frac{x + h - x}{h[\sqrt{x+h} + \sqrt{x}]} = \frac{1}{[\sqrt{x+h} + \sqrt{x}]} .$$

Now, we can pass to the limit $h \to 0$ and get $f'(x) = \frac{1}{2\sqrt{x}}$. We see that the limit does not exist for $x = 0$. Indeed, the slope of the tangent at that point is infinite.

1.9. For the **exponential function** $f(x) = e^x$, one can also compute the derivative

$$\frac{e^{x+h} - e^x}{h} = \frac{[e^x e^h - e^x]}{h} = e^x \frac{e^h - 1}{h} .$$

The reasoning why $(e^h - 1)/h$ goes to 1 for $h \to 0$ depends on how you have defined the exponential function. In a calculator for example, the exponential function is implemented as a sum $e^h = 1 + h + h^2/2! + h^3/3! + \cdots$ in which case, the fact $(e^h - 1)/h \to 1$ can be directly seen. One can also define e^h as the limit $(1 + h/n)^n$ for n going to infinity. Also here, the derivative is $n(1 + h/n)^{n-1}/n = (1 + h/n)^{n-1}$ is at $h = 0$ equal to 1 for any n so that in the limit $n \to \infty$ we have also 1. This justifies $(e^x)' = e^x$. In the same way you can verify that $(e^{bx})' = be^{bx}$ and especially $(a^x)' = \log(a)a^x$, the reason being that $a^x = e^{x \log(a)}$ with the natural log $\log = \ln$. Similarly as $e^x = \exp(x)$ are notations for the exponential both $\ln(x) = \log(x)$ are notations for the logarithm.
3

---

[3] We would write $\log_{10}(x)$ for the logarithm with base 10.

1.10. An other important function is the **reciprocal function** $f(x) = 1/x$. Its graph is called a **hyperbola**. Lets see whether we can compute the derivative limit. Lets start with

$$f(x+h) - f(x) = \frac{1}{x+h} - \frac{1}{x} = \frac{-h}{x(x+h)}$$

which was done by making a common denominator. Now we can divide by h and get $-1/(x(x+h))$ which has a nice limit for $h \to 0$. The result is $-1/x^2$. One can do that also for higher powers like $f(x) = 1/x^2$. Lets try:

$$(f(x+h) - f(x) = \frac{1}{(x+h)^2} - \frac{1}{x^2} = \frac{-h(h+2x)}{x^2(x+h)^2} \ .$$

Again, we can divide by h and take the limit to see that the result is $-2/x^3$.

1.11. In general, we have the rule

$$\boxed{\frac{d}{dx}x^n = nx^{n-1}}$$

And this works for any real number n. We have verified above explicitly in the case when $n = 2, 4, 1/2, -1, -2$. For positive integers n, it follows from a general Binomial formula for $(x+h)^n$, or then from the product rule, which tells how to differentiate a product of functions. In general, the limit computation $(f(x+h) - f(x))/h$ can become tedious, when done directly. [Remark about the future: We will learn methods to do this in general. An elegant way to apply the chain rule to $x^n = e^{n \log(x)}$ and get the derivative $e^{n \log(x)}(n/x)$ which again simplifies to nx^{n-1}. But it will still take some time until we get to the chain rule and to the derivative $\log'(x) = 1/x$ for the log which also follows from the chain rule as $\exp(\log(x))) = x$ implies $\exp(\log(x)) \log'(x) = 1$ so that $\log'(x) = 1/x$.]

2. SPRINTERS

2.1. In the video you have seen the data of the run of **Usain Bolt** breaking the world record at the Beijing 2008 Olympics. Here are some more illustrations

2.2. Humans are not that fast if we compare with other animals. A **domestic cat** can run 30 miles per hour, this is better than the fastest speed around 28 miles per hour which a human can run. A lion can run 50 miles/hour and a cheetah up to 70 miles per hour. If we look at all animals including birds speeds can get even bigger. A **peregrine falcon** can reach more than 200 miles per hour. There is some consolation: we normal mortals still can compete with **squirrels**. They can run 12 miles per hour. You can try on a tread mill to see whether you can top that.
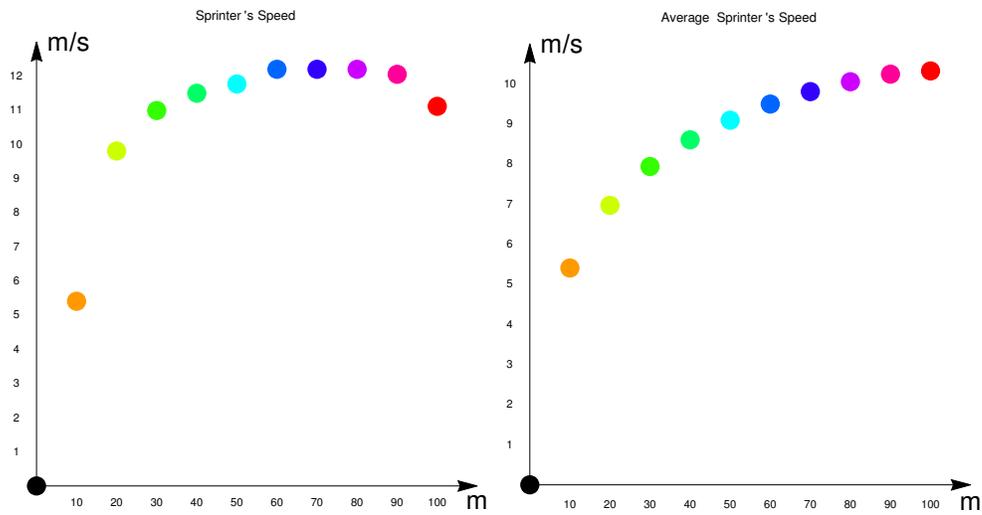4

---

[4]https://www.infoplease.com/math-science/biology/plants-animals/speed-of-animals

FIGURE 4. The first graph shows the average speed of Bolt on the intervals $[0, 10], [10, 20], \ldots, [90, 100]$. This is the **rolling speed** and an approximation for the **instantaneous rate of change**. The second graph shows the **average speed** on the already transversed part $[0, x]$, where x is the number of meters. The last point shows the average rate of change of the run over the entire interval from 0 to 100 m. A speed of 12 m/s corresponds to 43km/h. One year later, in Berlin in 2009, Bolt ran a 9.58 second 100 meter final.