Lecture 2

Calculus and Aggregation

History of Data Science, Spring 2022 @ UC San Diego

Suraj Rampure

Announcements

- Homework 2 is released, and will be due on **Sunday, April 10th at 11:59PM**.
- We will start taking attendance from today (won't tell you when).
- Remember that there are office hours from 12-1PM on Fridays!

SDSC or Zoom

Agenda

- Recap: Archimedes.
- The development of calculus during the Scientific Revolution.
  - Fermat.
  - Newton.
  - Leibniz.
- Aggregation (e.g. means, medians, modes).

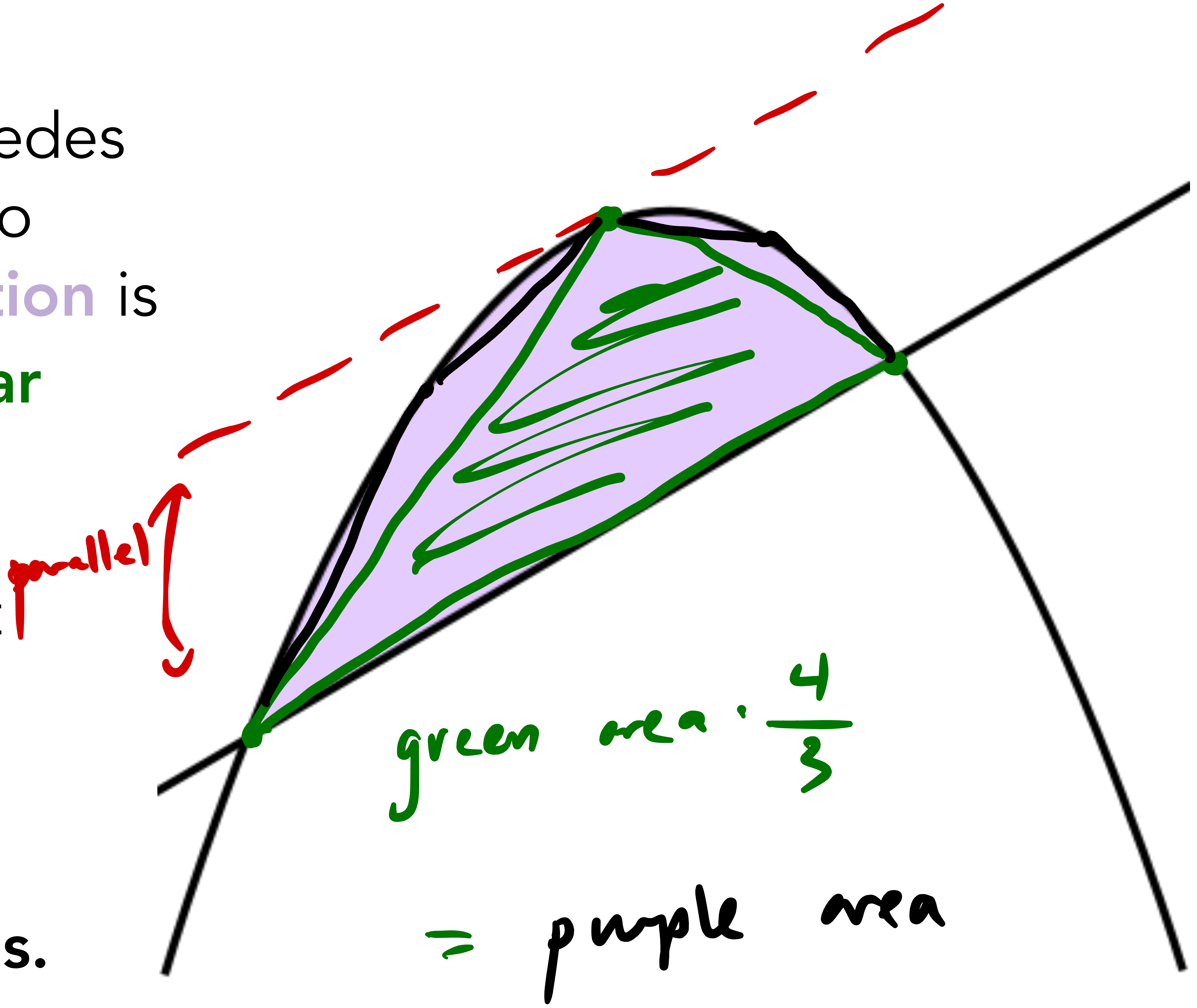
Recap of Archimedes

Quadrature of the parabola

- In *Quadrature of the Parabola*, Archimedes also used the **method of exhaustion** to show that the **area of a parabolic section** is equal to $\frac{4}{3}$ times the area of a **particular triangle**.
  - The argument also uses the fact that

$$1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \ldots = \frac{4}{3}$$

  - This is an **infinite geometric series.**



Infinite geometric series

Let's try proving that

$$\frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \ldots = \frac{1}{3}$$

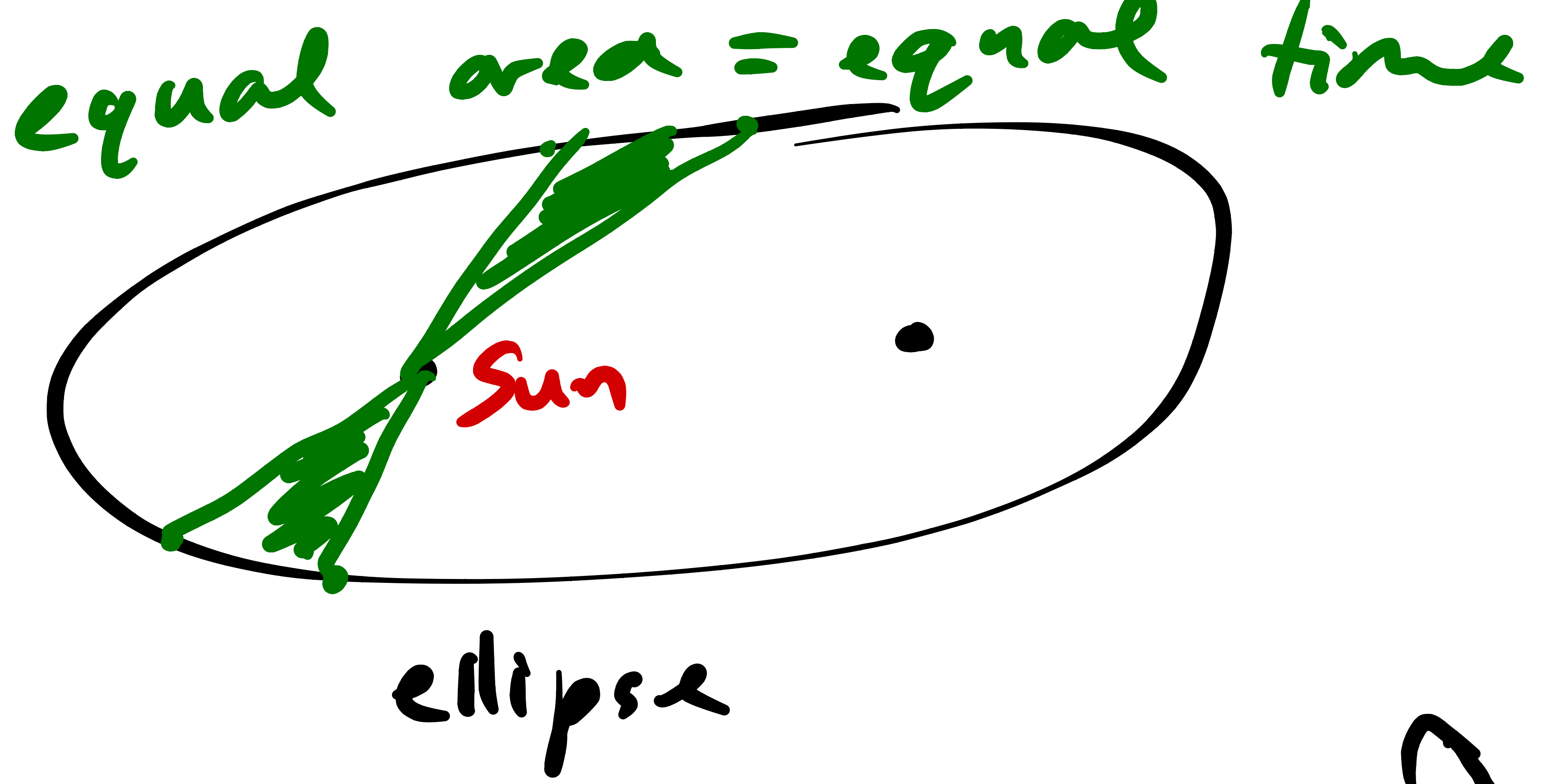
without using the geometric series formula.

$$\frac{a}{1-r} = \frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

the size of each partition is equal

$$3 \cdot A = 1 \Rightarrow A = \frac{1}{3}$$

The Scientific Revolution



The Scientific Revolution

- The Scientific Revolution, which roughly took place from 1550 to 1700, was a period of significant scientific development in Europe[1].

→ we orbit the sun

- **Copernicus** played a key role in establishing the heliocentric model of the Solar System, over the well-established geocentric model.

sun orbits us

- **Kepler** established his three laws of planetary motion, using **data** collected by **Tycho Brahe**.
- **Galileo** used a telescope to discover that the Moon's surface isn't smooth, and that Jupiter has as least 4 moons orbiting it.
- **Newton**, in addition to helping establish calculus, developed his more general laws of motion.
- The **Royal Society** was established in London in 1660.

1: https://www.britannica.com/science/Scientific-Revolution https://health.ucsd.edu › news › releases › pages › 2012...

UC San Diego Professor Named to the Prestigious Royal ...

Apr 25, 2012 — Founded in 1660, **Royal Society** Fellows have included Isaac Newton, Charles Darwin, Ernest Rutherford, Albert Einstein, Dorothy Hodgkin, Francis ...

https://jacobsschool.ucsd.edu › news › release

Engineering Professor Elected to Royal Society

Paul Linden, a professor in **UC San Diego's** Jacobs School of Engineering, has been elected as a fellow to the United Kingdom's National **Academy** of Science in ...

https://library.ucsd.edu › object

Margaret Burbidge elected a Fellow of the Royal Society of ...

Oct 27, 2020 — Margaret Burbidge elected a Fellow of the **Royal Society** of London ... **University of California**, **San Diego**--History.

Disclaimer

- We don’t have enough time to cover every detail in the development of calculus, even if we just restrict ourselves to the Scientific Revolution.
  - There exist 4-unit upper-division courses dedicated to that, e.g. Math 163 (History of Math).
- Instead, we’ll look at the work of a select few mathematicians – in particular, Fermat, Newton, and Leibniz.

Fermat

$a=4,\ p=5$

4^5-4

$=(2^2)^5-4=2^{10}-4=1024-4$

$=1020$ multiple of 5

- Pierre de Fermat (1601-1665) was a French **lawyer**, who is responsible for several results in number theory.
  - **Fermat's little theorem**: if a is an integer and p is a prime number, then $a^p - a$ is an integer multiple of p. cryptography / $\boxed{\text{RSA}}$
  - **Fermat's Last Theorem**: there are no integer solutions to $a^n + b^n = c^n$ when $n > 2$.

    $a^4+b^4=c^4$
  - Also known as one of the creators of **probability** – we'll hear more about him in a few weeks.

latin

OBSERVATIO DOMINI PETRI DE FERMAT.

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos & generaliter nullam in infinitum vltra quadratum poteſtatem in duos eiuſdem nominis fas eſt diuidere cuius rei demonſtrationem mirabilem ſane detexi. Hanc marginis exiguitas non caperet.

"A cube into two cubes, a fourth power into two fourth powers, and in general to infinity, none other than square powers can be divided into two squares (literally, "two of its own name"), a fact of which I have discovered a truly remarkable proof. This margin is too small to contain it."[1]

1: https://www.maa.org/press/periodicals/convergence/mathematical-treasure-bachets-arithmetic-of-diophantus

Fermat and adequality

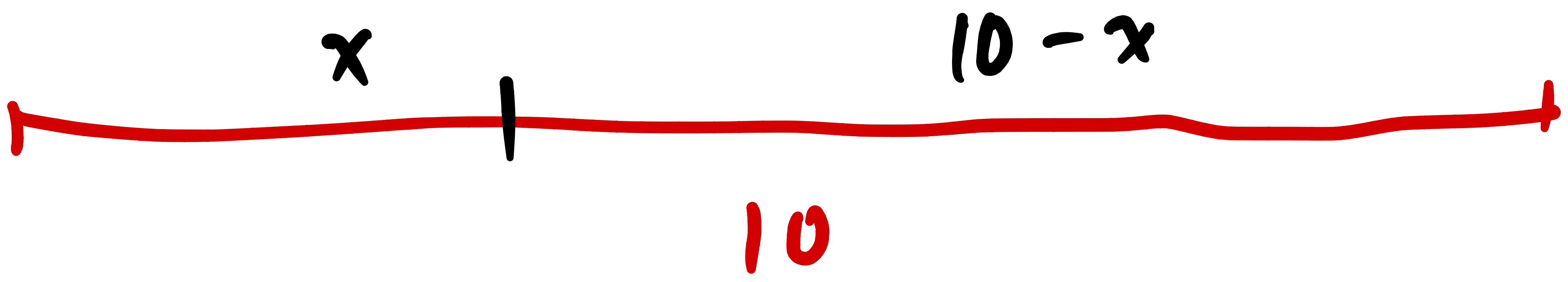
kind of like limits

- Fermat used a method called **adequality** to minimize/maximize functions and to find the slopes of tangent lines.
    - "Adequality" translates to "approximately equal". The word was first used by **Diophantus**, a Greek mathematician who lived in 200s AD in Alexandria.
    - His method resembles the concept of a limit.

Let a be an arbitrarily chosen unknown of the question (whether it has one, two, or three dimensions, as follows from the statement). We will express the maximum or minimum quantity in terms of a, by means of terms of any degree. We will then substitute $a + e$ for the primitive unknown a, and express the maximum or minimum quantity in terms containing a and e to any degree. We will *ad-equate,* to speak like Diophantus,[3] the two expressions of the maximum and minimum quantity, and we will remove from them the terms common to both sides. Having done this, it will be found that on both sides, all the terms will involve e or a power of e. We will divide all the terms by e, or by a higher power of e, such that on at least one of the sides, e will disappear entirely. We will then eliminate all the terms where e (or one of its powers) still exists, and we will consider the others equal, or if nothing remains on one of the sides, we will equate the added terms with the subtracted terms, which comes to be the same. Solving this last equation will give the value of a, which will lead to the maximum or the minimum, in the original expression.

1: https://digitalcommons.ursinus.edu/cgi/viewcontent.cgi?article=1011&context=triumphs_calculus

Maximization via adequality



- Consider the task of maximizing the function $f(x) = bx - x^2$, for some constant b. (For example, let $b = 10$.)
  - We can think of it as the task of finding a number x such that the product of x and $10 - x$ is as large as possible.
  - Observation: the product is large when x and $10 - x$ are close together.
  - Let e be some small positive number. Then…

$$f(x) = 10x - x^2$$
$$= x(10 - x)$$

$$f(a) \sim f(a+e)$$
$$\cancel{10a} - \cancel{a^2} \sim \cancel{10a} + 10e - \cancel{a^2} - 2ae - e^2$$
$$\frac{0}{e} \sim \frac{10e - 2ae - e^2}{e}$$
$$0 \sim 10 - 2a - \cancel{e}$$
$$10 \sim 2a$$
$$\boxed{a \sim 5}$$

Maximization via adequality

$$\boxed{f(x) = bx^2 - x^3}$$

General idea: to find the a that maximizes $f(x)$,

adequate

1. Set $f(a) = f(a + e)$.

2. Divide both sides by e, and cancel any remaining terms involving e.

3. Solve for a.

① $\cancel{ba^2} - \cancel{a^3} = \cancel{ba^2} + 2bae + be^2 - \cancel{a^3} - 3a^2e - 3ae^2 - e^3$

② $0 = 2ba + \cancel{be} - 3a^2 - \cancel{3ae} - \cancel{e^2}$

$\Rightarrow 0 = 2ba - 3a^2$

adequality

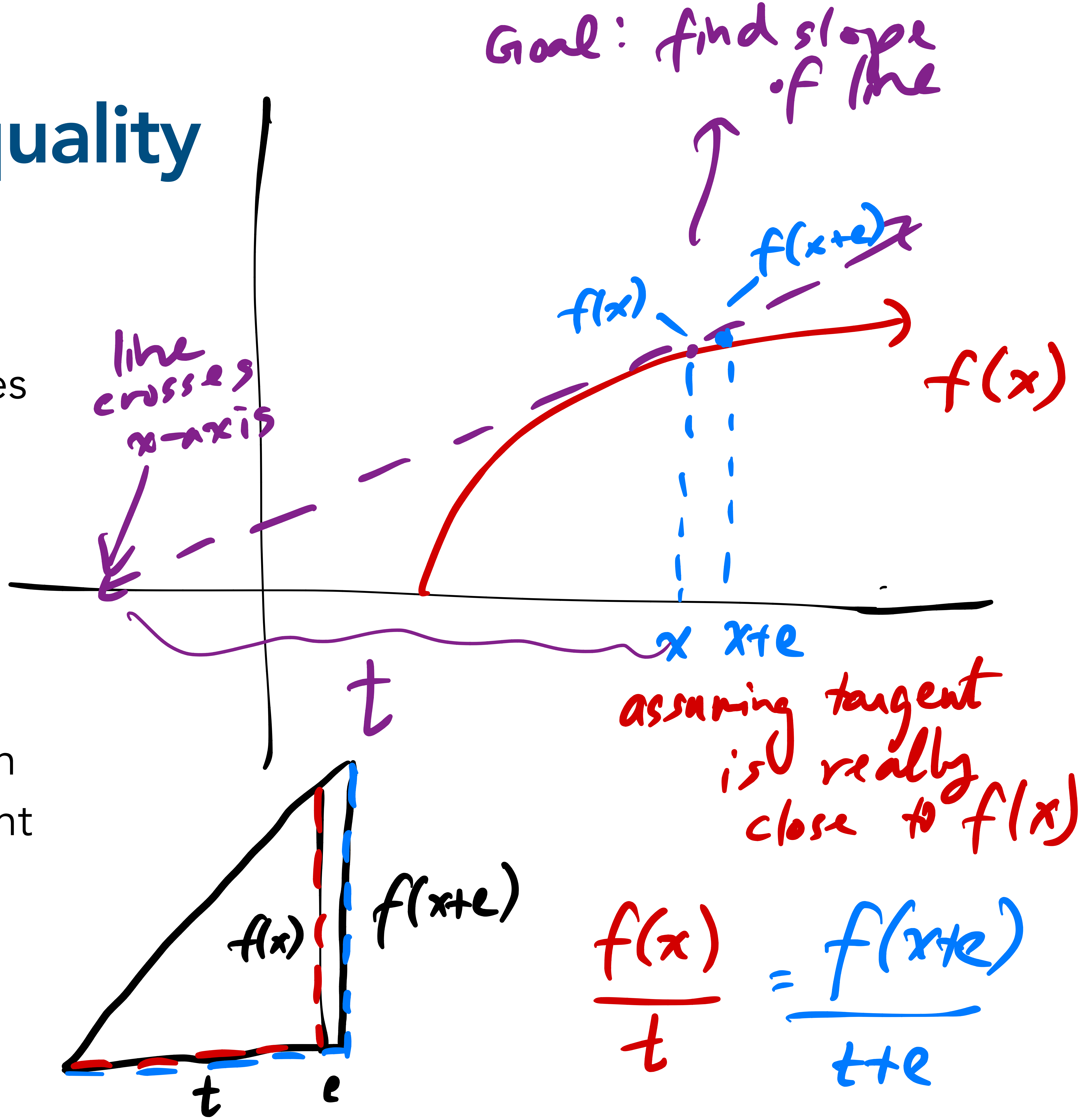
$0 = a(2b - 3a)$

$a = 0$ or $2b = 3a$

$\Rightarrow \boxed{a = \frac{2}{3}b}$

Tangent lines via adequality

- Fermat was also able to find the slopes of tangent lines using adequality.
- Idea: use adequality to find the **subtangent**, t, of a curve at a point $(a, f(a))$.
    - t is defined as the distance between a and the $x-$intercept of the tangent line of f at a.
- How? **Similar triangles**.



$$\frac{f(x)}{t} = \frac{f(x+e)}{t+e}$$

adequality

Example: $f(x) = x^2$

slope of tangent when $x = 3$

$$\frac{f(3)}{t} = \frac{f(3+e)}{t+e} \Rightarrow \frac{9}{t} = \frac{9+6e+e^2}{t+e}$$

$$\Rightarrow \frac{\cancel{9t}+9e}{e} = \frac{\cancel{9t}+6et+e^2t}{e}$$

$$9 = 6t + \cancel{et}$$

$$\Rightarrow 6t = 9$$

$$t = \frac{3}{2}$$

$$\text{slope} = \frac{f(x)}{t} = \frac{9}{\frac{3}{2}}$$

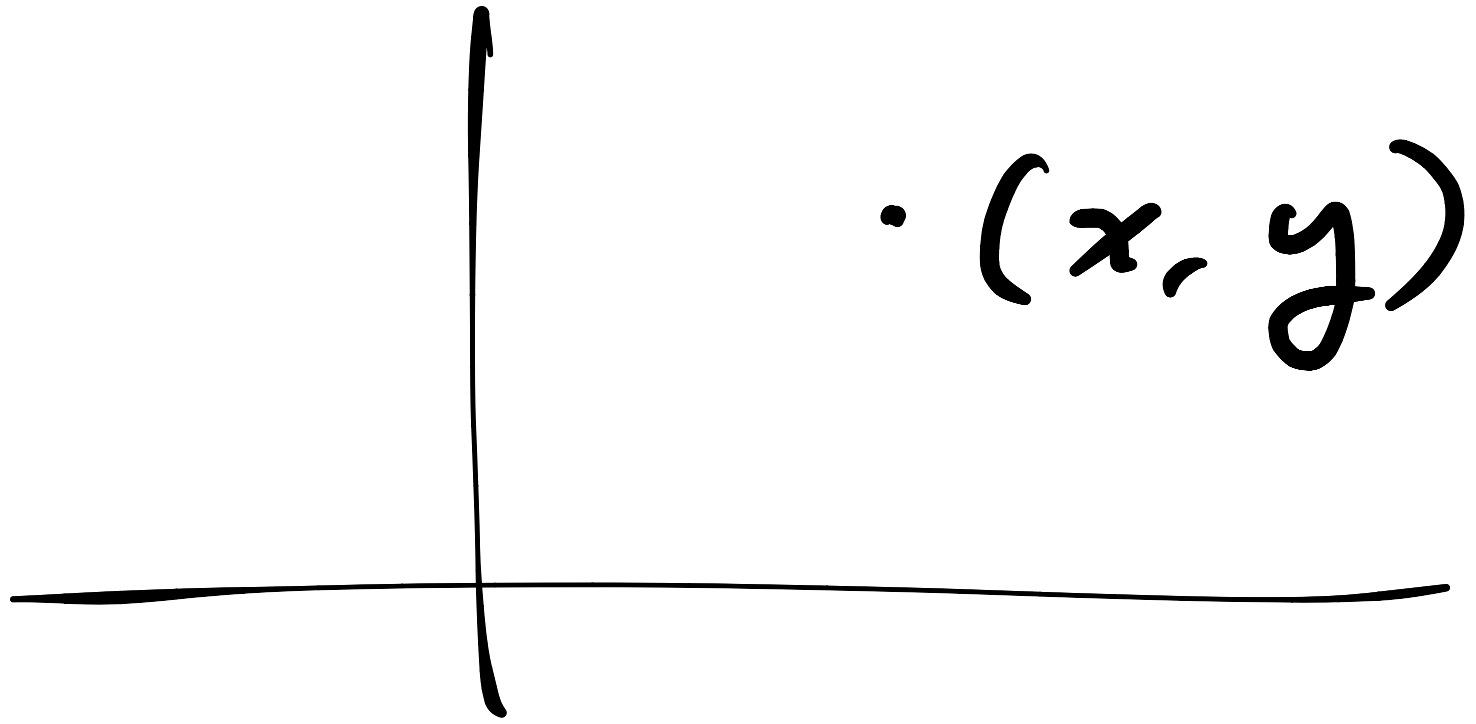
$$= 9 \cdot \frac{2}{3} = \boxed{6}$$

Tangent lines via adequality

General idea: to find the slope of $f(x)$ at point a,

1. Set $\dfrac{f(a)}{t} = \dfrac{f(a+e)}{t+e}$.
2. Divide both sides by e, and cancel any remaining terms involving e.
3. Solve for t.
4. The slope of the tangent line is then $\dfrac{f(a)}{t}$.

Folium of Descartes

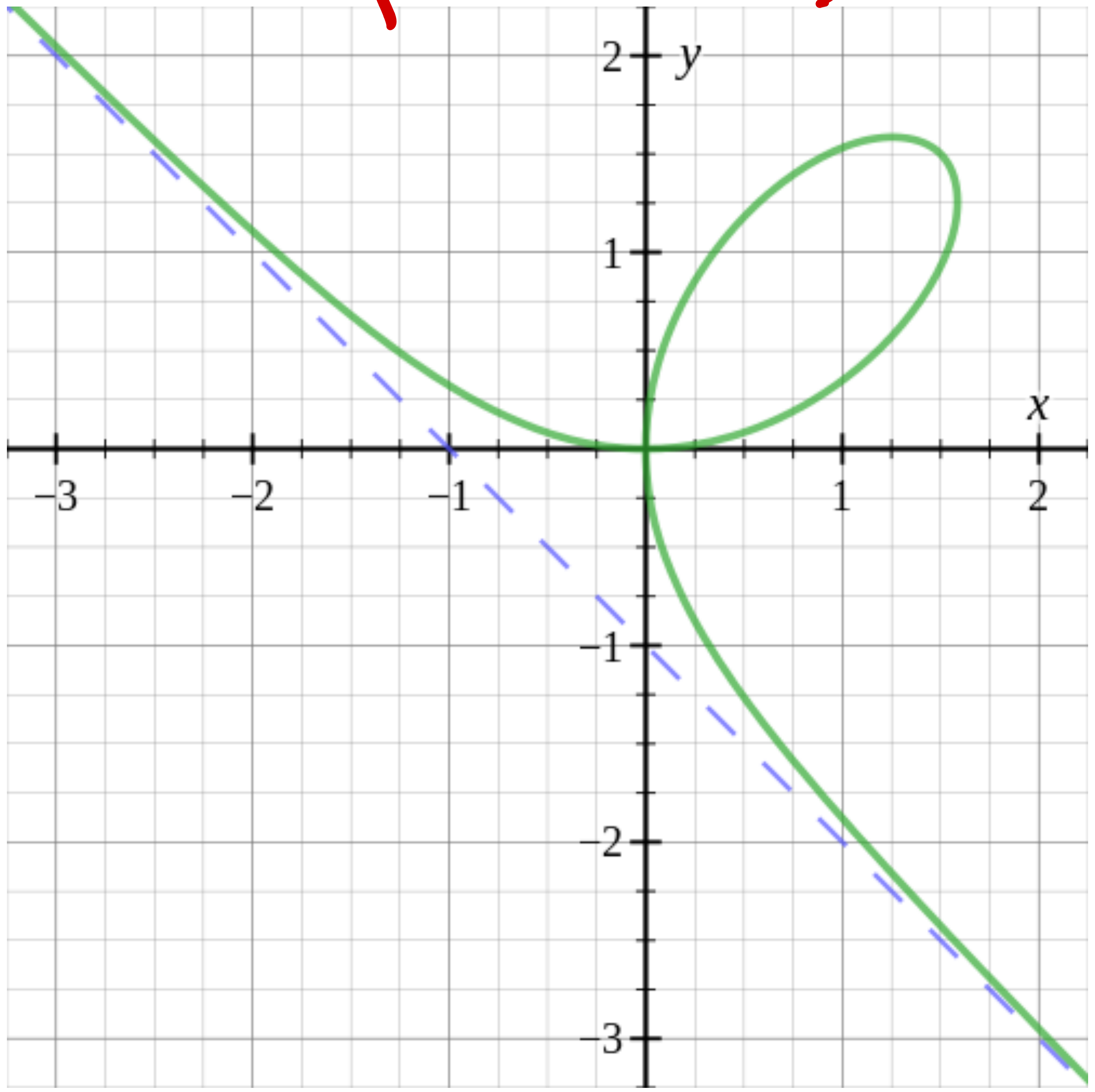


- René Descartes is a famous mathematician in his own right, and is also known as the founder of modern philosophy.
  - Also known for the Cartesian plane.
- Around the same time as Fermat, he developed his own method of finding tangents to curves, using **normals**.
- Fermat told Descartes about his subtangent technique, and Descartes was skeptical. Descartes challenged Fermat to find the tangent line to the **implicit** relationship $\underbrace{x^3 + y^3 - 3xy = 0}_{\text{implicit}}$, known as the Folium of Descartes.

Folium of Descartes

leaf

looks like a leaf!



$x^3 + y^3 - 3xy = 0$

Fermat was able to find the tangent line easily, with a small modification to his approach.

$$H(x,y) = x^3 + y^3 - 3xy = 0$$

$$\Rightarrow \boxed{H(x,y) = H\left(x+e,\ y + \underbrace{\frac{ey}{t}}\right)}$$

adequality

Hint: look at the similar triangles

Fermat and integration

FTC: $\int_a^b f(x)\,dx = F(b) - F(a)$

$\frac{d}{dx} F(x) = f(x)$

- Fermat was also able to find the areas under some curves.
  - For instance, he proved that the area under $y = x^2$ between $x = 0$ and $x = a$ is $\frac{a^3}{3}$.
  - In other words, he showed that $\int_0^a x^2 = \frac{a^3}{3}$ (not with this notation, though).
- Fermat, then, developed methods for both finding slopes and finding areas. **Why is he not credited with the discovery of calculus?**
  - Answer: he didn't find (or document) any **connection** between areas and slopes.

Newton and Leibniz

Newton

$F = ma$

$F = G\frac{m_1 m_2}{r^2}$

- Isaac Newton (1642-1727) was an English mathematician and physicist. He is credited with discovering:
  - Newton's 3 laws of motion and the law of universal gravitation.
  - The fact that light is made up of a seven color rainbow 🌈.
  - The binomial theorem.
  - Calculus*.
- He went to Cambridge University, and supposedly produced some of his best work while school was cancelled due to "The Great Plague of London" in 1665.

bubonic plague

Fluents and fluxions

fluents : $x(t)$

fluxions : $x'(t) = \frac{d}{dt} x(t)$

- Newton worked under Isaac Barrow, another English mathematician.
  - Barrow was the first to publish the **fundamental theorem of calculus**, though Newton (supposedly) proved it on his own later.
- In 1665, Newton wrote notes on "fluxions", but did not publish them until in the 1700s.
  - A fluent is a quantity whose value changes with time, and a fluxion is the **instantaneous** rate of change of a quantity a given moment in time.
  - If x is a fluent, $\dot{x}$ represents the corresponding fluxion.
- In his notes, he posed both the problem of finding a fluxion given a fluent (differentiation) and the problem of finding a fluent given a fluxion (integration).

$\dot{x} = \frac{dx}{dt}$

$\frac{\dot{y}}{\dot{x}} = \frac{dy}{dx}$

PROBLEM 1

GIVEN THE RELATION OF THE FLOWING QUANTITIES TO ONE ANOTHER, TO DETERMINE THE RELATION OF THE FLUXIONS.

SOLUTION

Arrange the equation by which the given relation is expressed according to the dimensions of some fluent quantity, say x, and multiply its terms by any arithmetical progression and then by $\dot{x}/x$. Carry out this operation separately for each one of the fluent quantities and then put the sum of all the products equal to nothing, and you have the desired equation.[87]

EXAMPLE 1. If the relation of the quantities x and y be $x^3-ax^2+axy-y^3=0$, I multiply the terms arranged first according to x and then to y in this way.

Multiply	x^3	$-ax^2$	$+axy$	$-y^3$	Mult.	$-y^3$	$+axy$	$\begin{matrix}-ax^2\\+x^3\end{matrix}$
by	$\frac{3\dot{x}}{x}$.	$\frac{2\dot{x}}{x}$.	$\frac{\dot{x}}{x}$.	0,	by	$\frac{3\dot{y}}{y}$.	$\frac{\dot{y}}{y}$.	0.
there comes	$3\dot{x}x^2$	$-2\dot{x}ax$	$+\dot{x}ay$	*.	comes	$-3\dot{y}y^2$	$+a\dot{y}x$	*.

And the sum of the products is $3\dot{x}x^2-2a\dot{x}x+a\dot{x}y-3\dot{y}y^2+a\dot{y}x=0$, an equation which gives the relation between the fluxions $\dot{x}$ and $\dot{y}$. Precisely, should you assume x arbitrarily the equation $x^3-ax^2+axy-y^3=0$ will give y, and with these determined it will be $\dot{x}:\dot{y}=(3y^2-ax):(3x^2-2ax+ay)$.

Descartes' folium: $x^3+y^3-3xy=0$ → implicit diff

$$x^3+0x^2-3yx+y^3$$

$$3\frac{\dot{x}}{x} \quad 2\frac{\dot{x}}{x} \quad 1\frac{\dot{x}}{x} \quad 0\frac{\dot{x}}{x}$$

$$3x^2\dot{x} \qquad -3y\dot{x}$$

$$y^3+0y^2-3xy+x^3$$

$$3\frac{\dot{y}}{y} \quad 2\frac{\dot{y}}{y} \quad 1\frac{\dot{y}}{y} \quad 0\frac{\dot{y}}{y}$$

$$3y^2\dot{y} \qquad -3x\dot{y}$$

$$3x^2\dot{x}-3y\dot{x}+3y^2\dot{y}-3x\dot{y}=0$$

$$\dot{x}(x^2-y)+\dot{y}(y^2-x)=0$$

$$\boxed{\frac{\dot{y}}{\dot{x}}=\frac{x^2-y}{x-y^2}}$$

Leibniz

1665: Newton "finds" calc
1684: Leibniz "finds and publishes"
1700: Newton "publishes"

- Gottfried Wilhelm Leibniz (1646-1716) was a German polymath.
  - Leibniz a mathematician, philosopher, and politician, and was educated as a lawyer.
  - He is credited with developing the binary number system – more on this later in the quarter.
- Independently from Newton, Leibniz developed the field of calculus.
  - He published his results in 1684, before Newton did, even though Newton supposedly developed the theory of fluents and fluxions first. This sparked the "**priority dispute**."

Notation

$\underbrace{f'(x)}_{\text{Euler}}$ $\int x^2 dx$ $\frac{dy}{dx}$

- The modern notation we use in calculus largely comes from Leibniz, who was much more deliberate with the notation he used.
  - To Leibniz, dx represented an infinitesimally small change in x.
  - The integral symbol, $\int$, came from stretching out an s, for "sum".
- Newton's notation is occasionally used in physics.
- The "prime" notation, e.g. $f'(x)$, was only developed by Euler, many years later.
  - Euler was the first to establish the notion of a function, as well.

Further development

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

- The development of calculus didn't stop with Newton and Leibniz.
  - George Berkeley, a philosopher (who the California city and university is named after), questioned the idea of fluxions and called them "the ghost of departed quantities".
  - Cauchy and Weierstrass developed the theory of limits well after both Newton and Leibniz died.
  - Riemann, who created the Riemann sum that the modern formal definition of an integral is based on, was only born in 1826.

Aggregation

Mean, median, and mode

- We're often interested in **summarizing** a set of numbers using just a single number.
  - **Mean**: Add up all the numbers and divide by the number of numbers.
  - **Median**: Sort the numbers and pick the middle one.
  - **Mode**: Pick the most common number.
- While the idea of summarizing data using a single number is common now, it wasn't always so ubiquitous.

Accounts of Thucydides

- Thucydides (~400 BC) was an early historian of science, and provided accounts of how the Greeks used methods of aggregation in times of war.
  - One example: In order to construct a ladder to reach the top of an enemy's wall, they needed to estimate the height of the wall. To do this, they would have many different people count the number of bricks in the wall, and hoped that the majority would answer correctly.
  - What method does this resemble?

mode

1. http://jse.amstat.org/v11n1/bakker.html

Summary, next time

Summary

yellkey.com/occur

brick

- Fermat developed techniques for finding the area under certain curves, finding tangent lines, and for maximizing/minimizing functions, but he did not develop a unifying theory of calculus.

- Newton and Leibniz independently developed the field of calculus, unifying the ideas of slopes and areas.

    - Newton thought in terms of "fluents and fluxions".

    - Leibniz established the notation we use today.

- **Next time:** more on the history of the mean and least squares.