Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:05

classical ML (next week)

Calculus & optimization - 2

Monday: Maxima/Minima
MVC
GD ← code

→ Wed: Classfn problem
Const. optmzn

→ Fri: PCA

Section (chat)

Session [Ques/Audio]

1 / 2

Big-picture

Classification (Binary) → Linear Algebra & CG Sep. Πd → Max/Min problem

↓

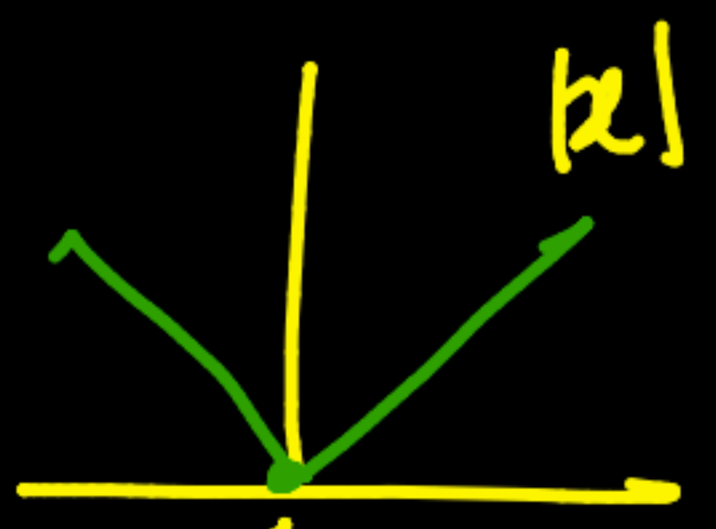
Calculus

↓

Limits; Derivative

DIFFERENTIABILITY

$|x|$ is not diff. @ $x=0$



except @ $x=0$

$$\lim_{\Delta x \to 0} \frac{|x+\Delta x| - |x|}{\Delta x}$$ was not defined

→ $|x|$ is not diff @ $x=0$

→ x^2 is diff... @ everywhere in its Domain

$\boxed{\text{Math}}$

Defn

$f(x)$ is diff.... @ if

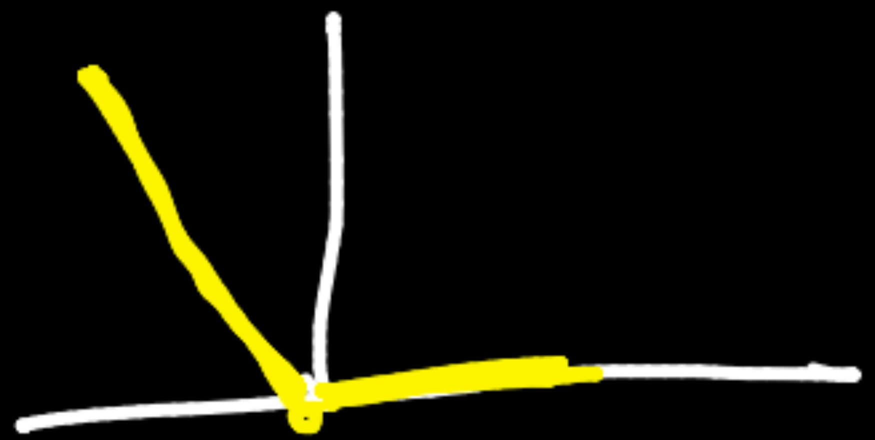
Domain

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

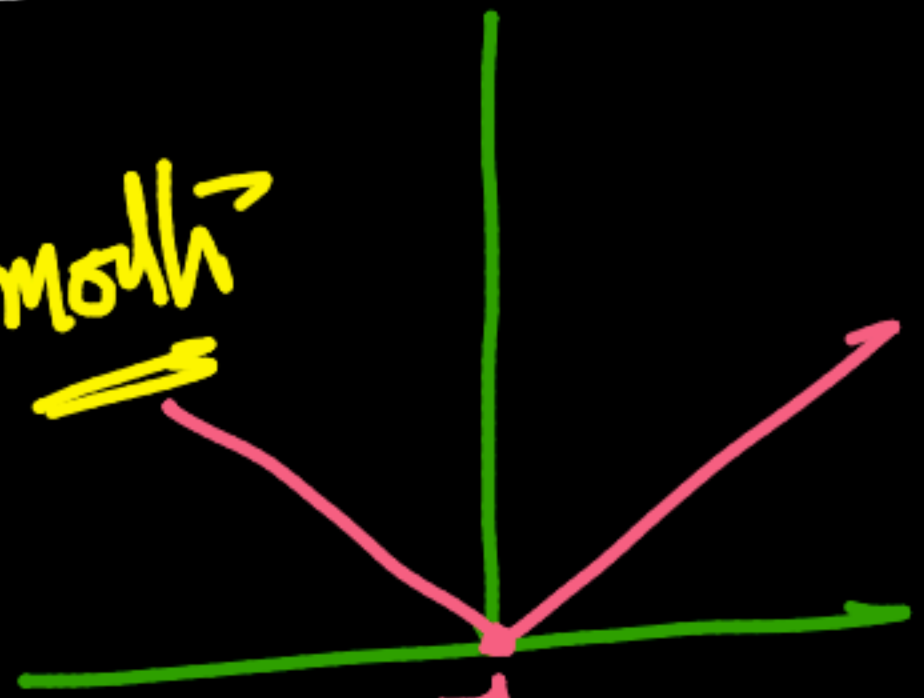
exists @ all x

LSL RSL

Geom - Intuition

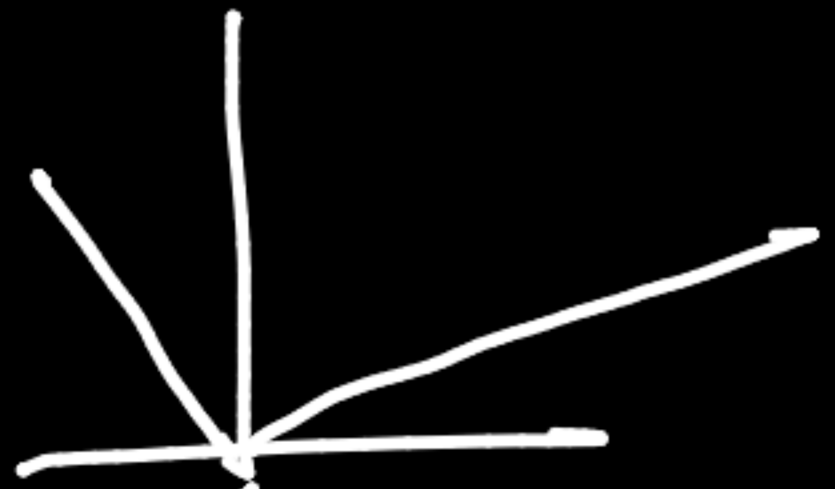


not smooth

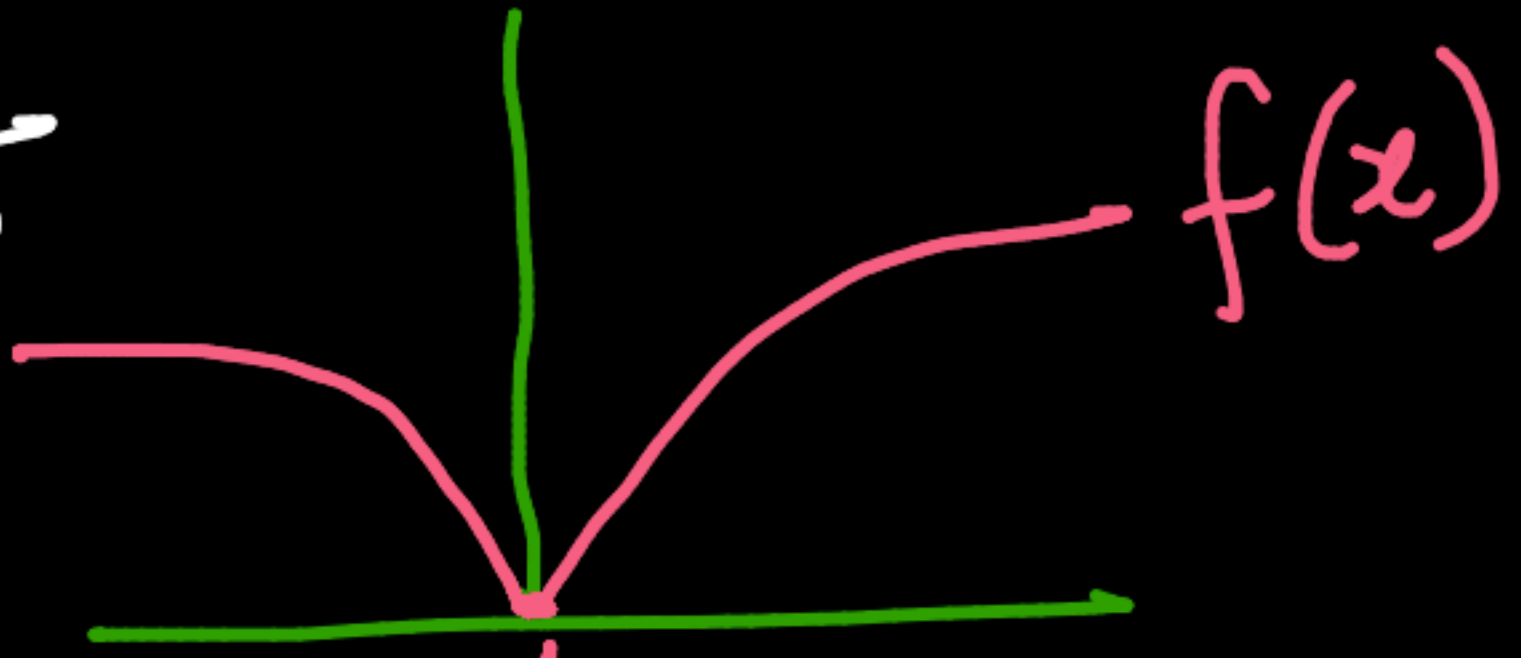


not smooth

Continous ✓



not smooth



not smooth

✓ { DIFFerentiability = continous & 'smooth'

Notation:

$$\frac{d}{dx} f(x) = \frac{df}{dx} = f' = y' = \frac{dy}{dx}$$

$$\dot{y}$$

$$\boxed{f(x) = y}$$

Rules:

1

$$\frac{d}{dx}\left[f(x)+g(x)\right] = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$$

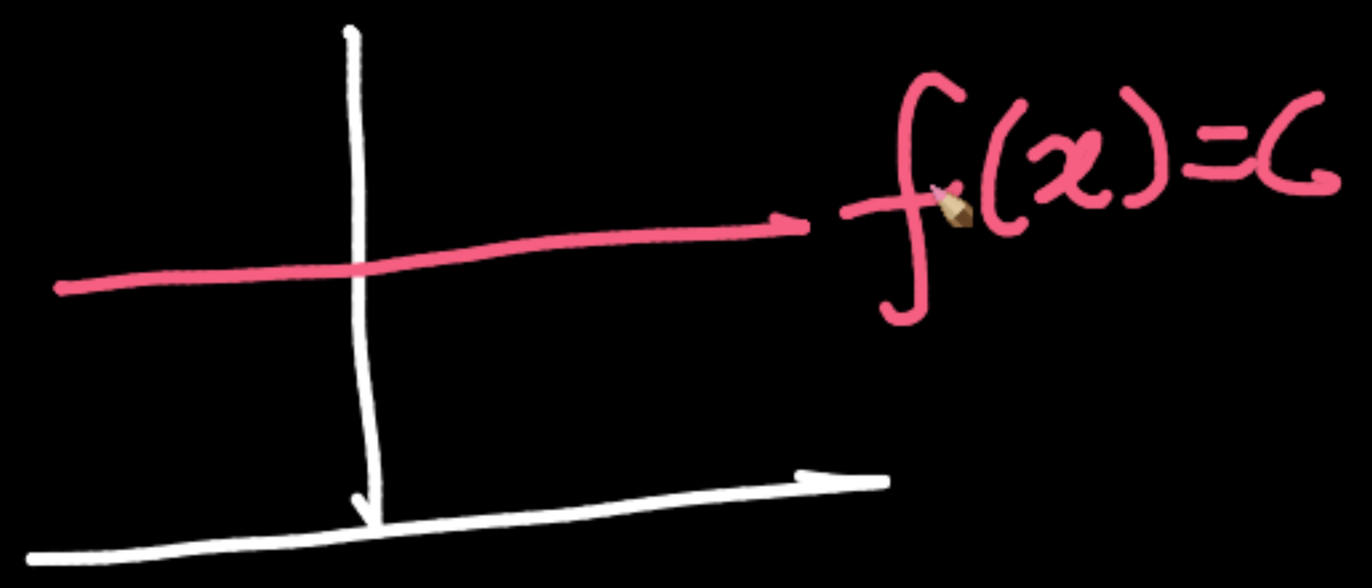
$$(f+g)' = f' + g'$$

7 / 8

2

C : Constant

$$\frac{dc}{dx} = 0$$



3

product rule:

$$\frac{d}{dx}\left[f(x)\cdot g(x)\right] = f(x)\frac{dg}{dx} + g(x)\frac{df}{dx}$$

$$\frac{d}{dx}(x^2) = \frac{d}{dx}(\underset{f(x)}{x}\cdot\underset{g(x)}{x}) = x\cdot 1 + x\cdot 1$$

$$= 2x$$

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

$n \neq 0$

$\boxed{Q}$

$$\frac{d}{dx}(\underset{f(x)}{\underset{\downarrow}{x}}\cdot\underset{g(x)}{\underset{\downarrow}{\log x}}) = x\cdot\frac{1}{x} + \log x\cdot 1$$

$$= \underline{\underline{1+\log(x)}}$$

CHAIN-RULE

[most widely used]

$$\left\{ \frac{d}{dx} f(g(x)) = f'(g(x)) \cdot g'(x) \right.$$

e.g.:

u

$$\frac{d}{dx}\left(\log \boxed{x^2}\right) = \frac{1}{x^2} \cdot 2x = \frac{2}{x}$$

$g(x) = x^2$

$f = \log$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:22

$\rightarrow$ natural-log

$$\frac{d}{dx}\left(\log x^2\right)$$

Let

$x^2 = u$

$\downarrow$

$\boxed{\ln}$

$$\frac{d}{dx}\log(x^2) = \frac{d}{dx}(\log u) = \frac{d\log(u)}{du}\cdot\frac{du}{dx}$$

x^2

$$= \frac{1}{u}\cdot 2x$$

$$= \frac{2x}{x^2} = \frac{2}{x}$$

12 / 12

outer inner

$$\frac{d}{dx}\log(x^2) = \frac{1}{x^2}\cdot 2x = \frac{2}{x} \Big\} \checkmark$$

✓ **Jeff Dean:**

Sr V.P

NeurIPS

DL

Tensorflow

Distb. systems

"Chain rule"

& Matrix Mul

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:26

Quoitent rule:

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)\,g(x) - g'(x)\,f(x)}{[g(x)]^2}$$

product!

$$\hookrightarrow f(x)\cdot\frac{1}{g(x)}$$

15 / 15

Q

$$\frac{d}{dx}\frac{\log(x)}{x} = \frac{\frac{1}{x}\cdot x - 1\log(x)}{x^2} = \frac{1-\log(x)}{x^2}$$

$\boxed{Q}$

$$\frac{d}{dx}\left(\frac{\overbrace{x}^{f(x)}}{\underbrace{\log x}_{g(x)}}\right) = \frac{1\cdot\log(x) - \frac{1}{\cancel{x}}\cdot\cancel{x}}{(\log x)^2}$$

$$= \frac{\log - 1}{(\log x)^2}$$

Chain rule:

DL

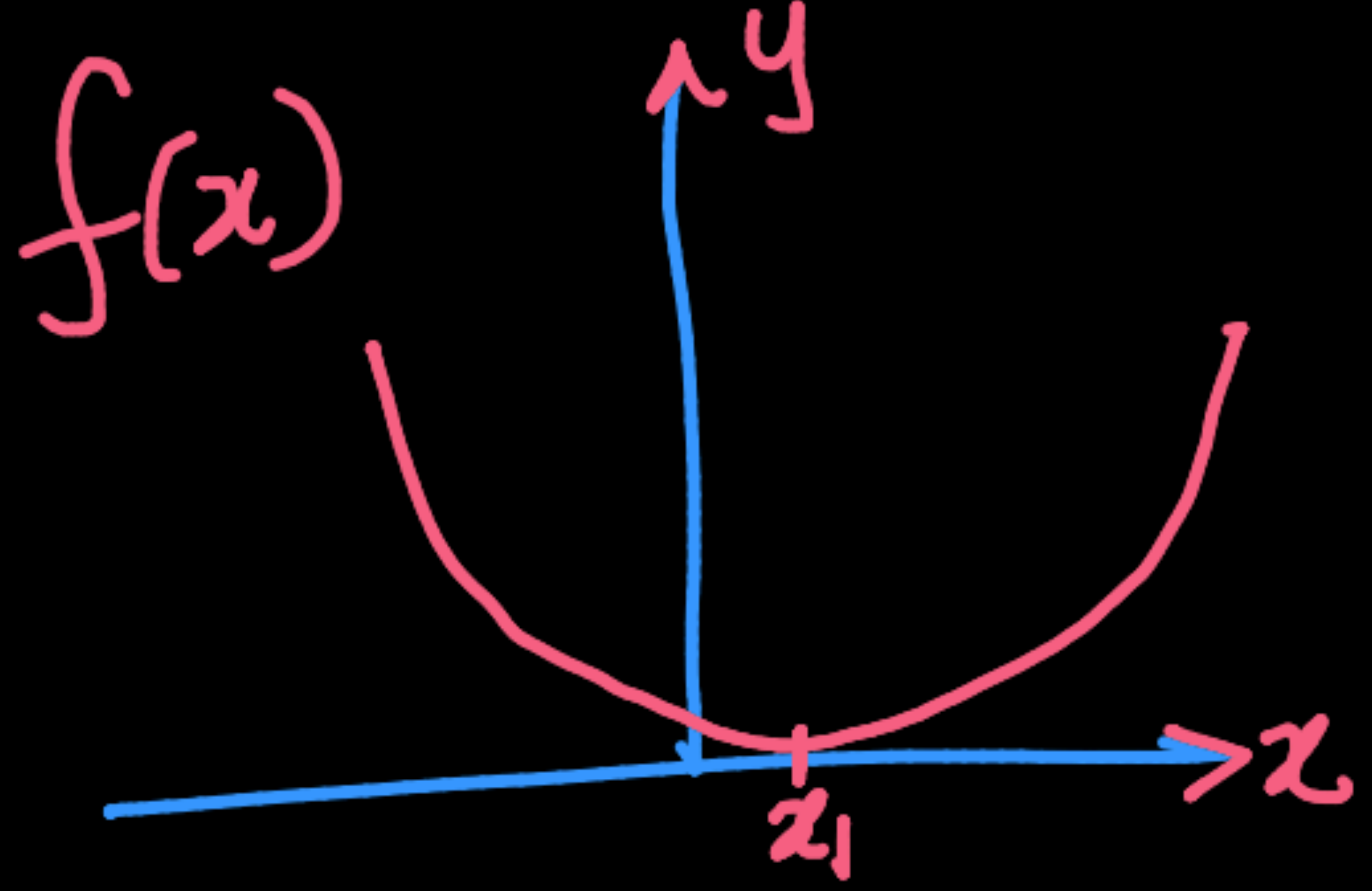
DL'

input-vec

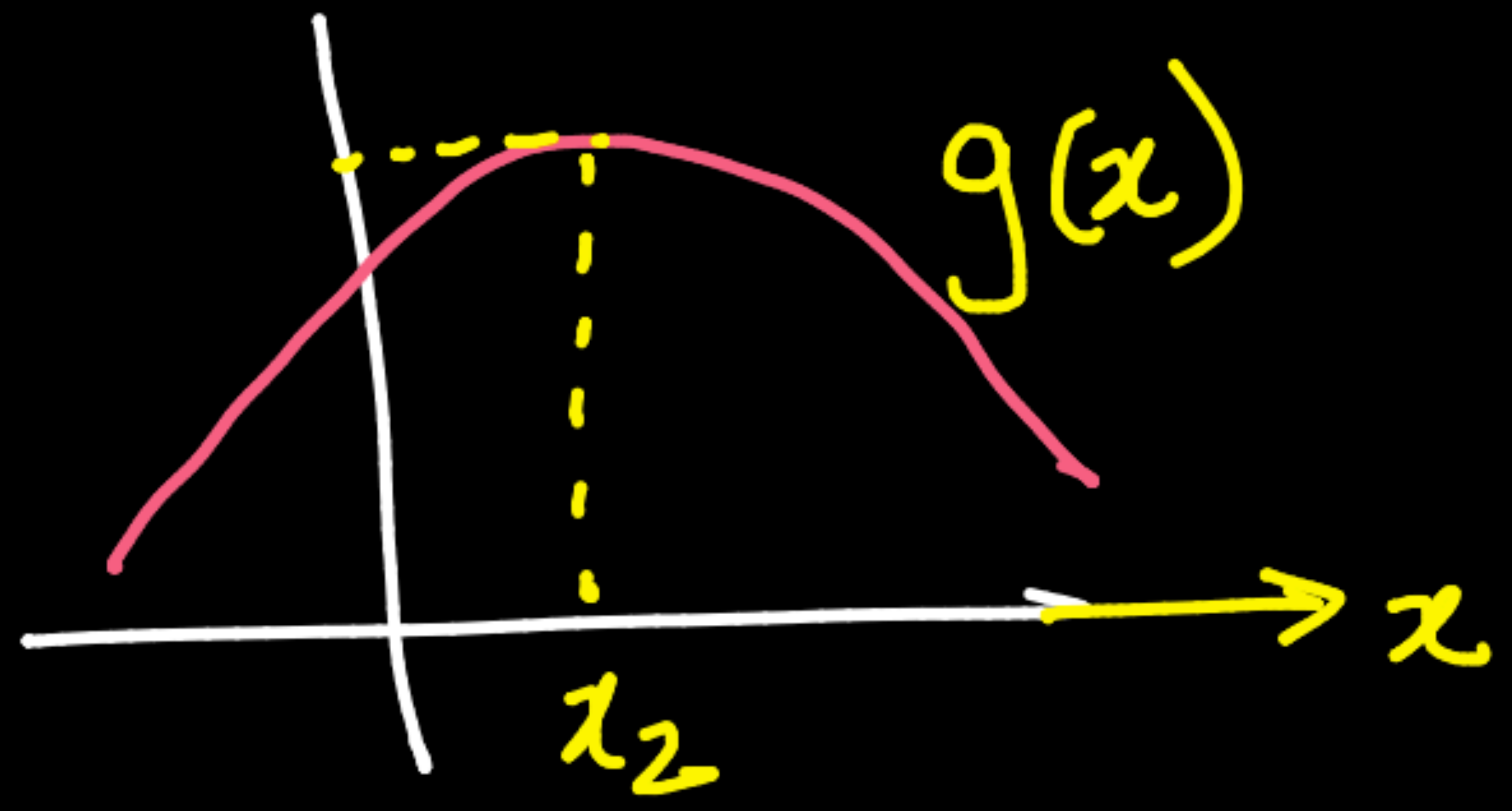
$$f_3(f_2(f_1(x)))$$

$$\frac{d}{dx} f(x) \longrightarrow \underline{\underline{\text{code}}}$$

MAXIMA & MINIMA



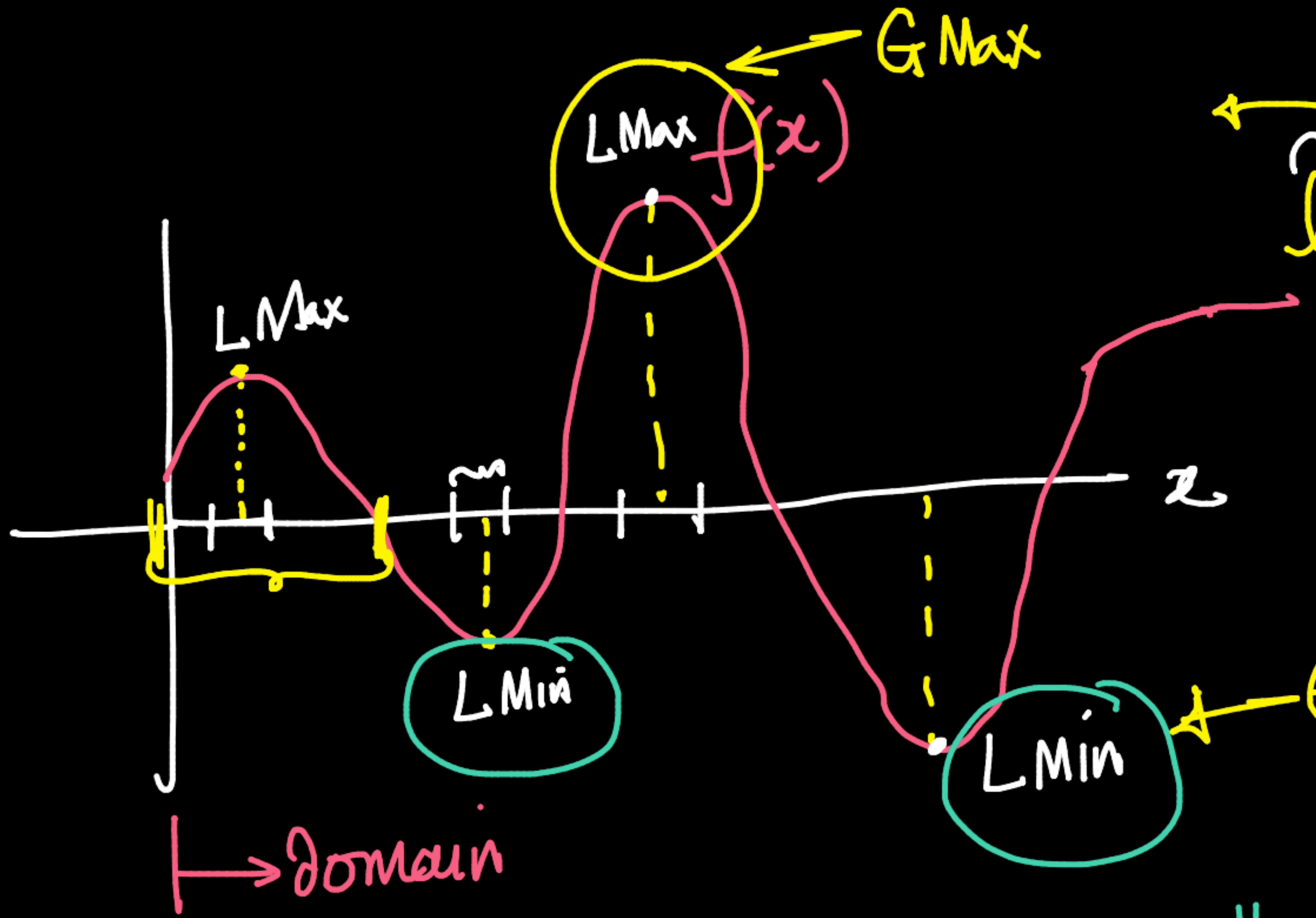
minima @ x_1



g(x) has a maxima @

$x = x_2$

Def:



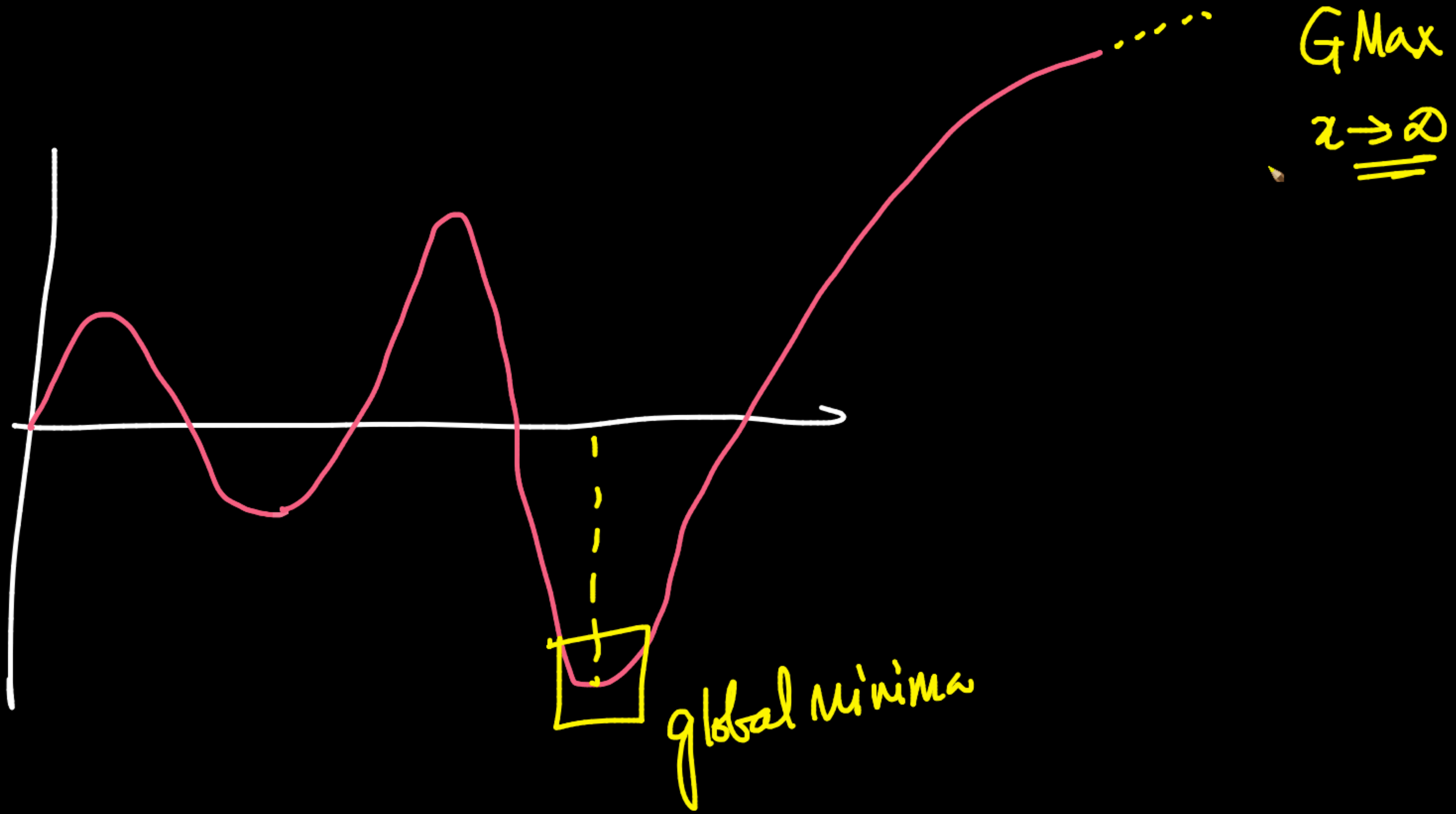
neighborhood

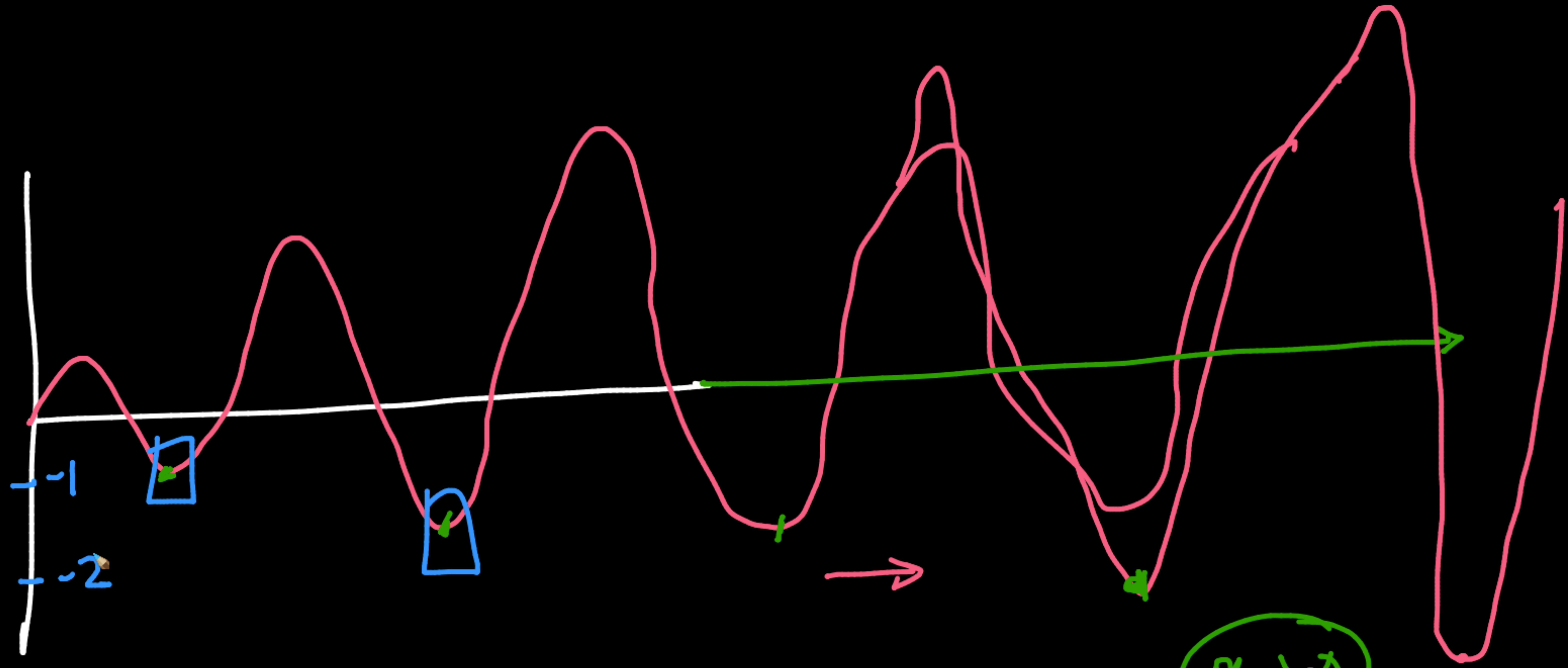
local minima

Global minima

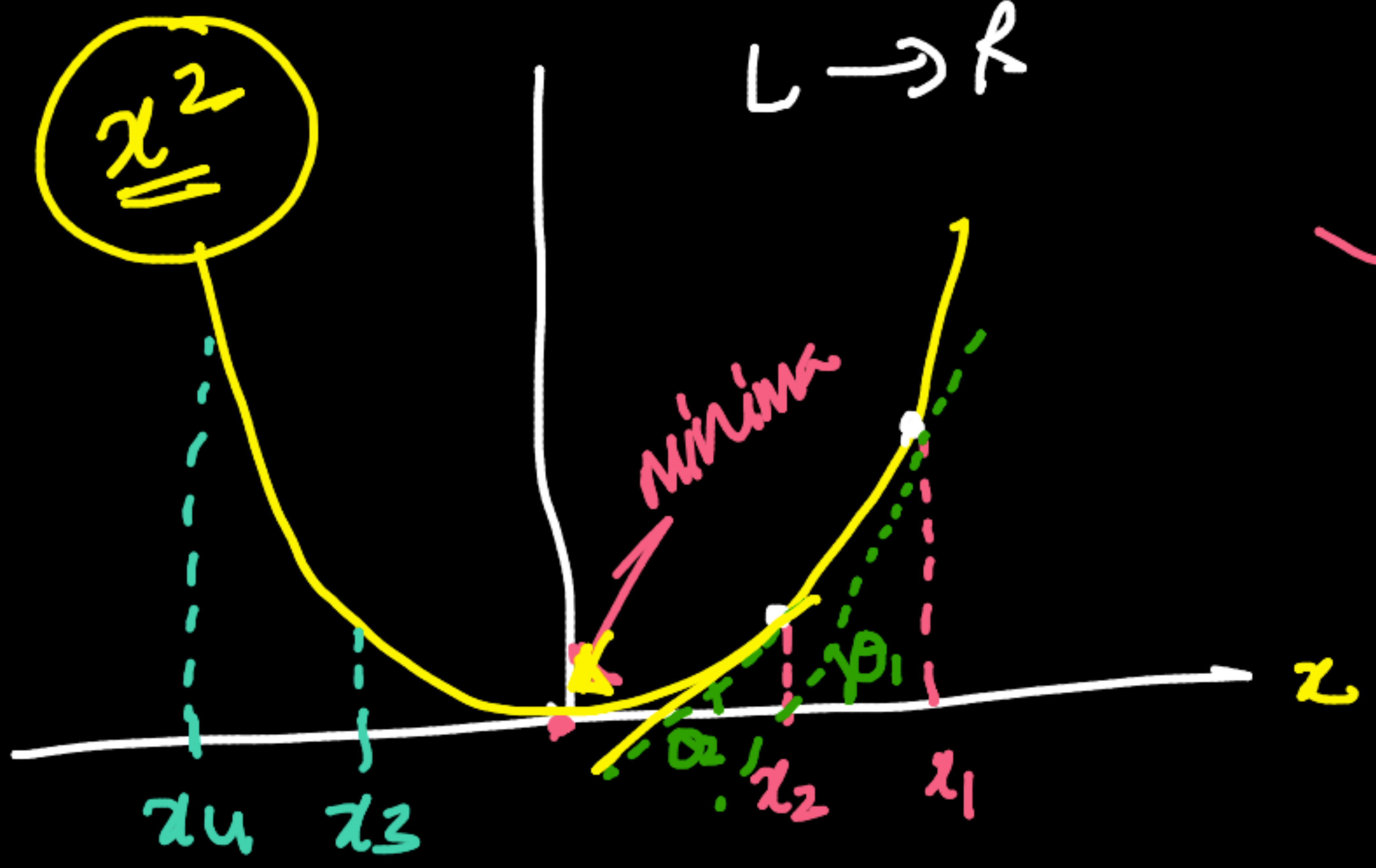
Domain

min amongst all local minima
= Global minima


GMax
$x \to \infty$
global minima



$$\begin{cases} \text{GMin} \rightarrow x \to \infty \\ \text{GMax} \rightarrow x \to \infty + \delta \end{cases}$$



How to find Maximas & minima

INTUITION

der → 0 ←

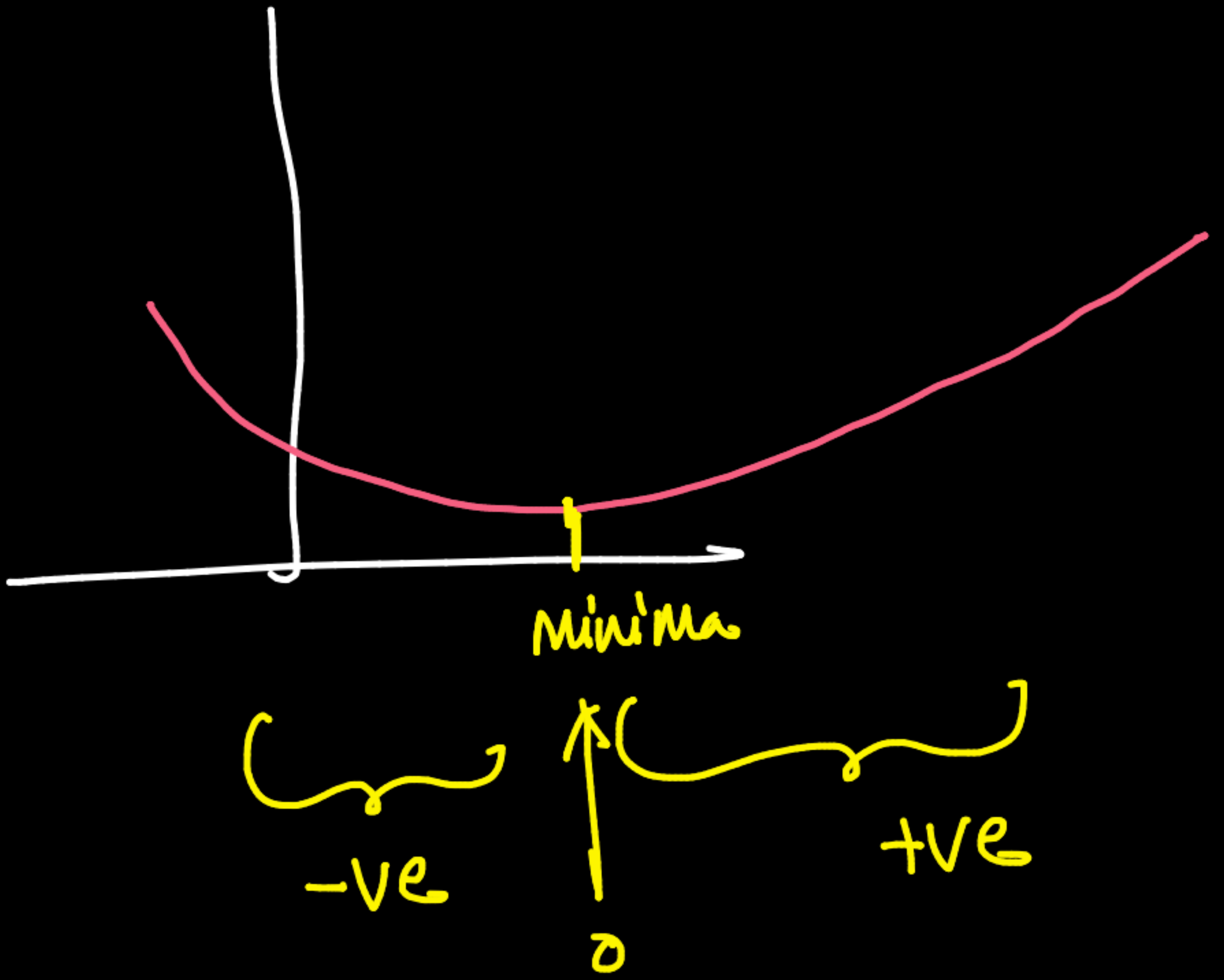
-ve +ve

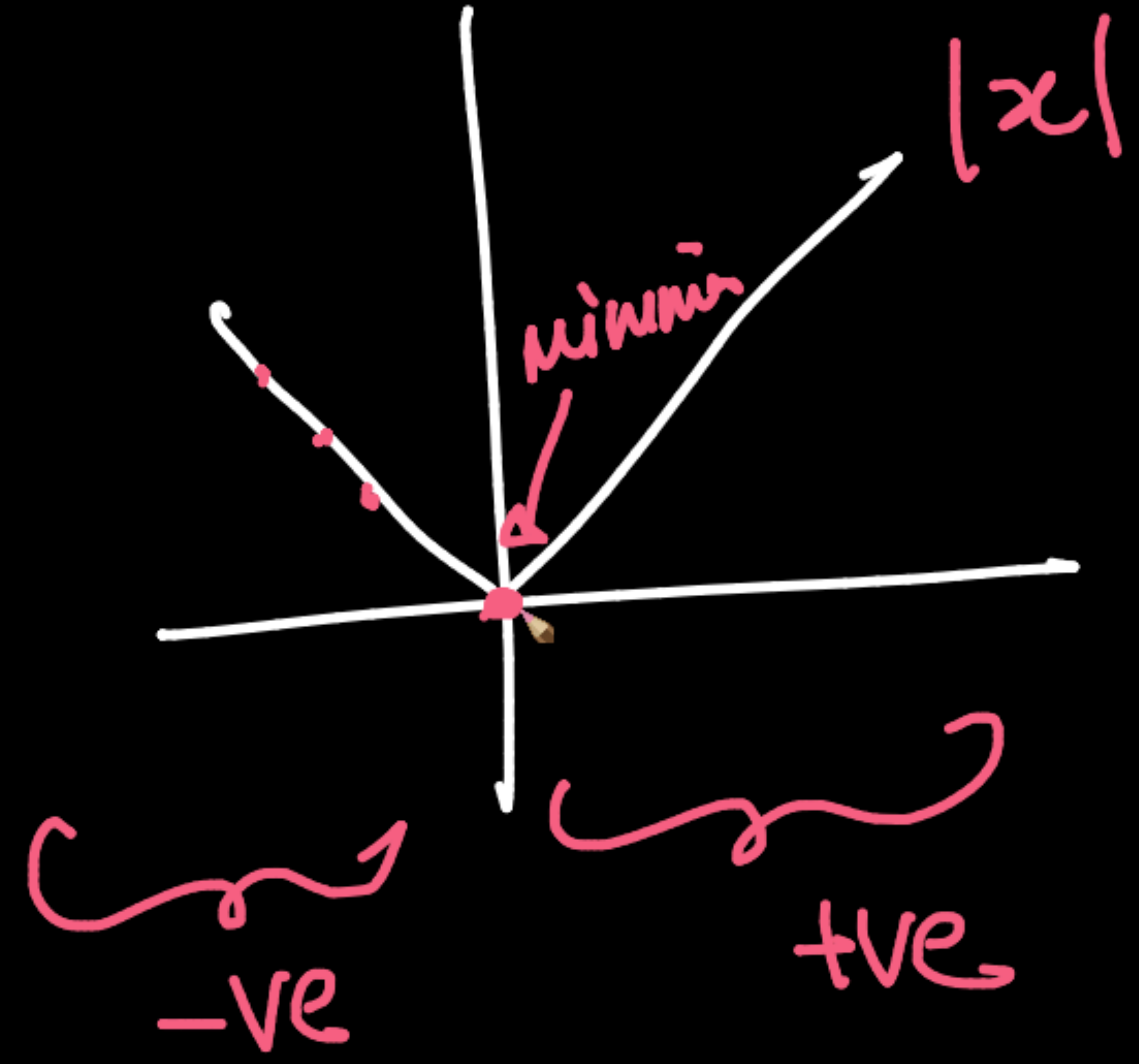
(incr) (incr)

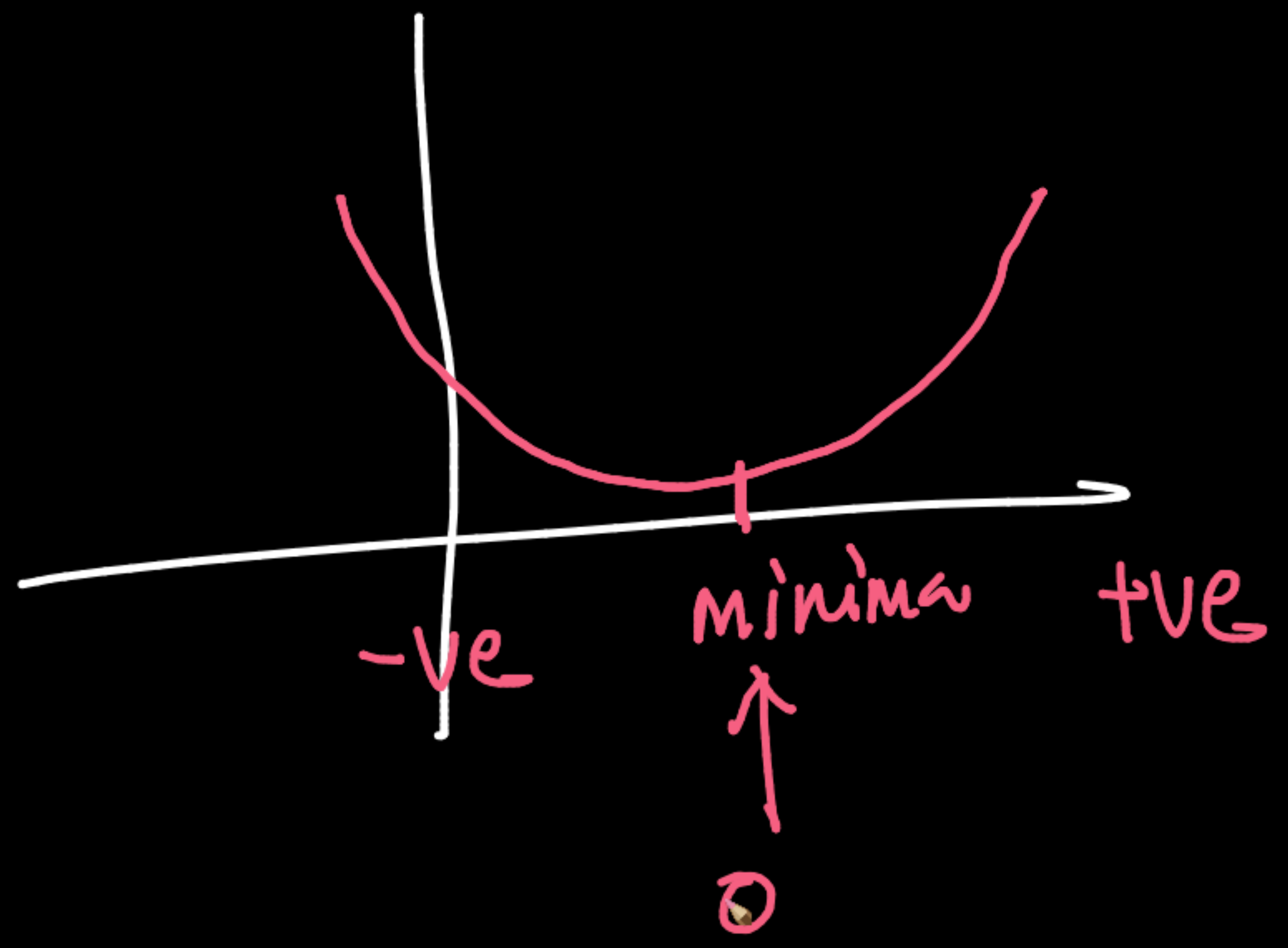
der@ x_1 > der@ x_2

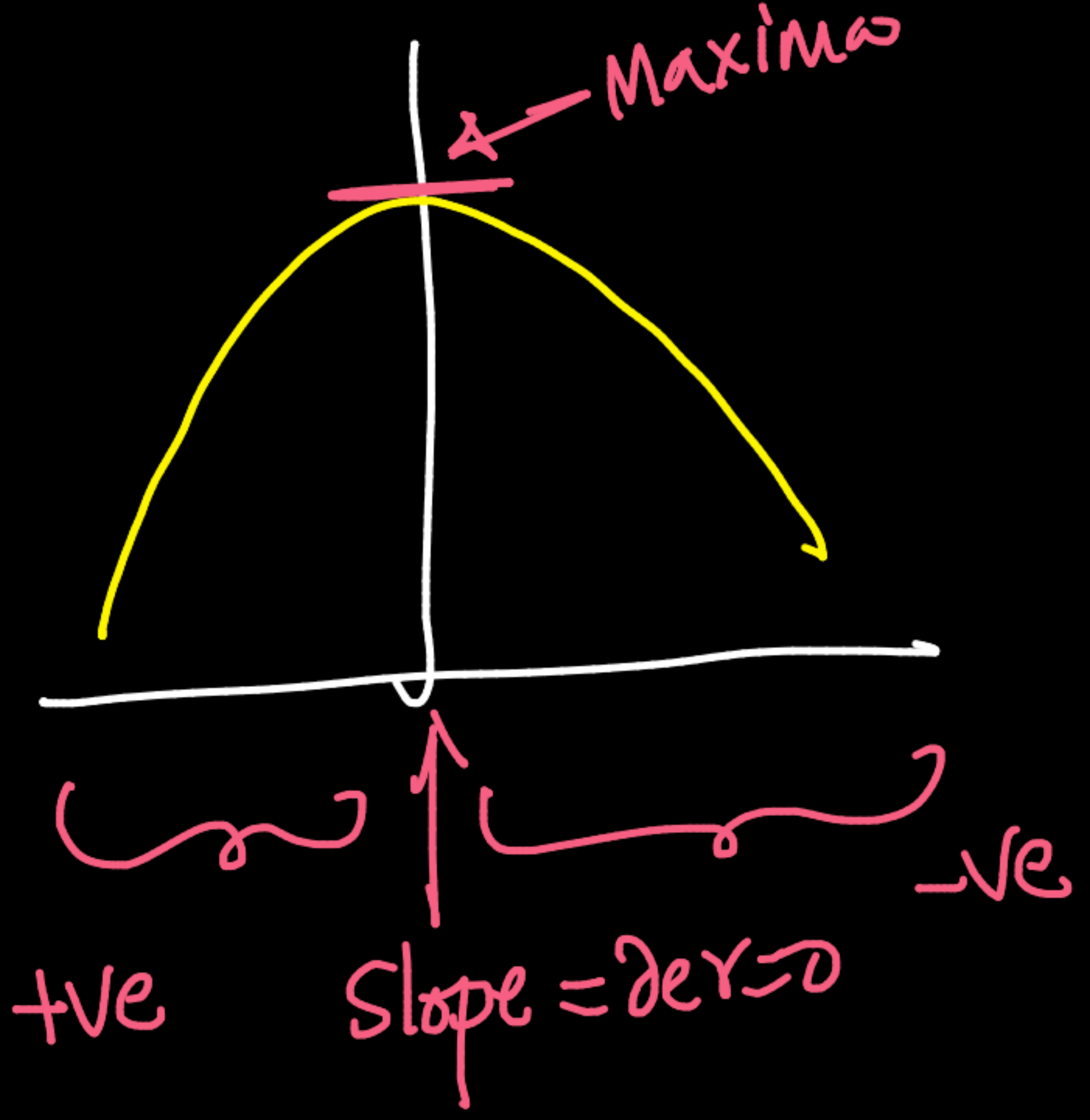
$0 < \theta_2 < \theta_1 < 90^\circ$

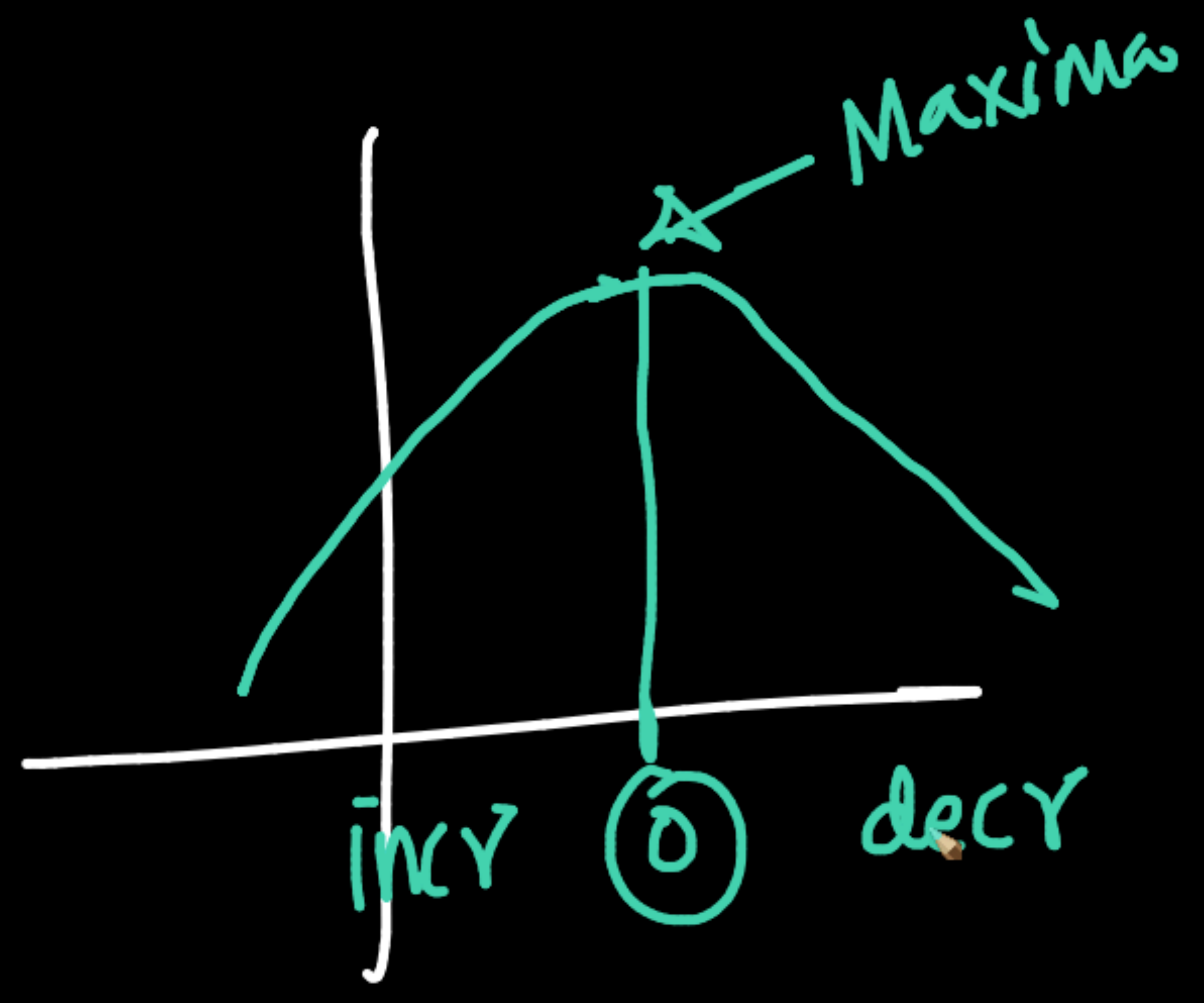
$\tan\theta_2 < \tan\theta_1$


minima
-ve
0
+ve


|x|
minima
-ve
+ve


minima
-ve
+ve
0

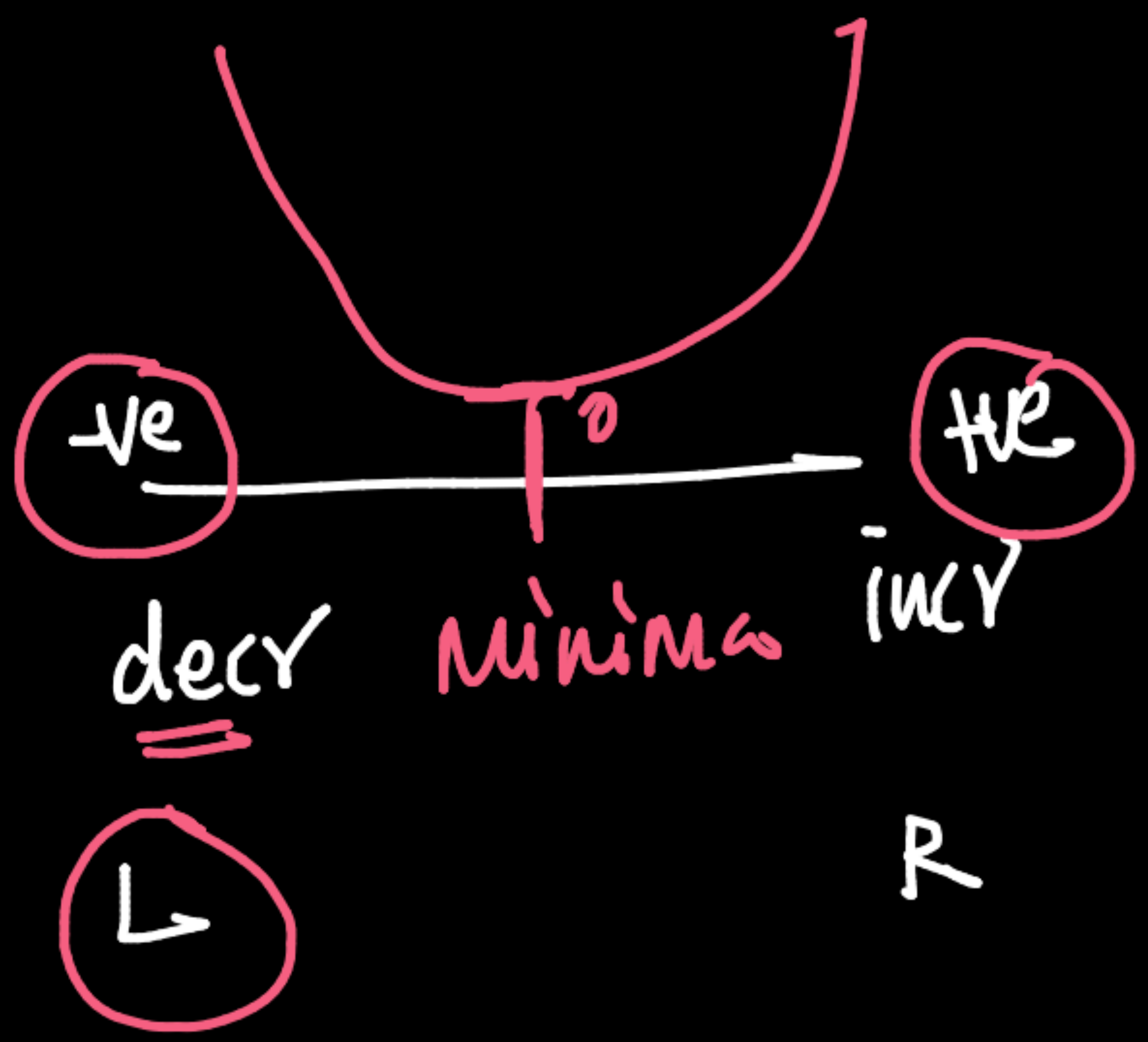

Maxima
+ve
Slope = der = 0
-ve


Maxima
incr
0
decr

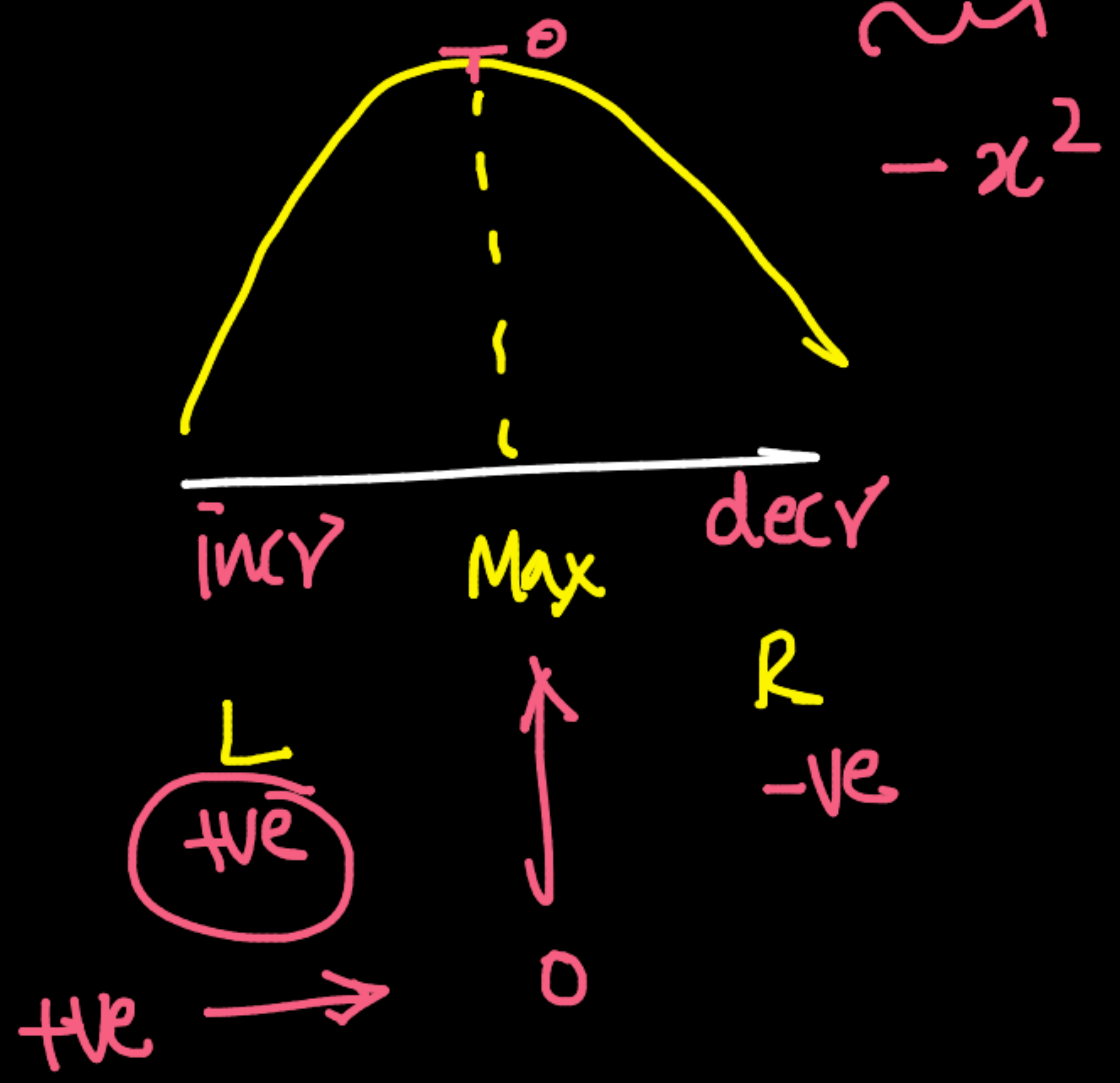
Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:50

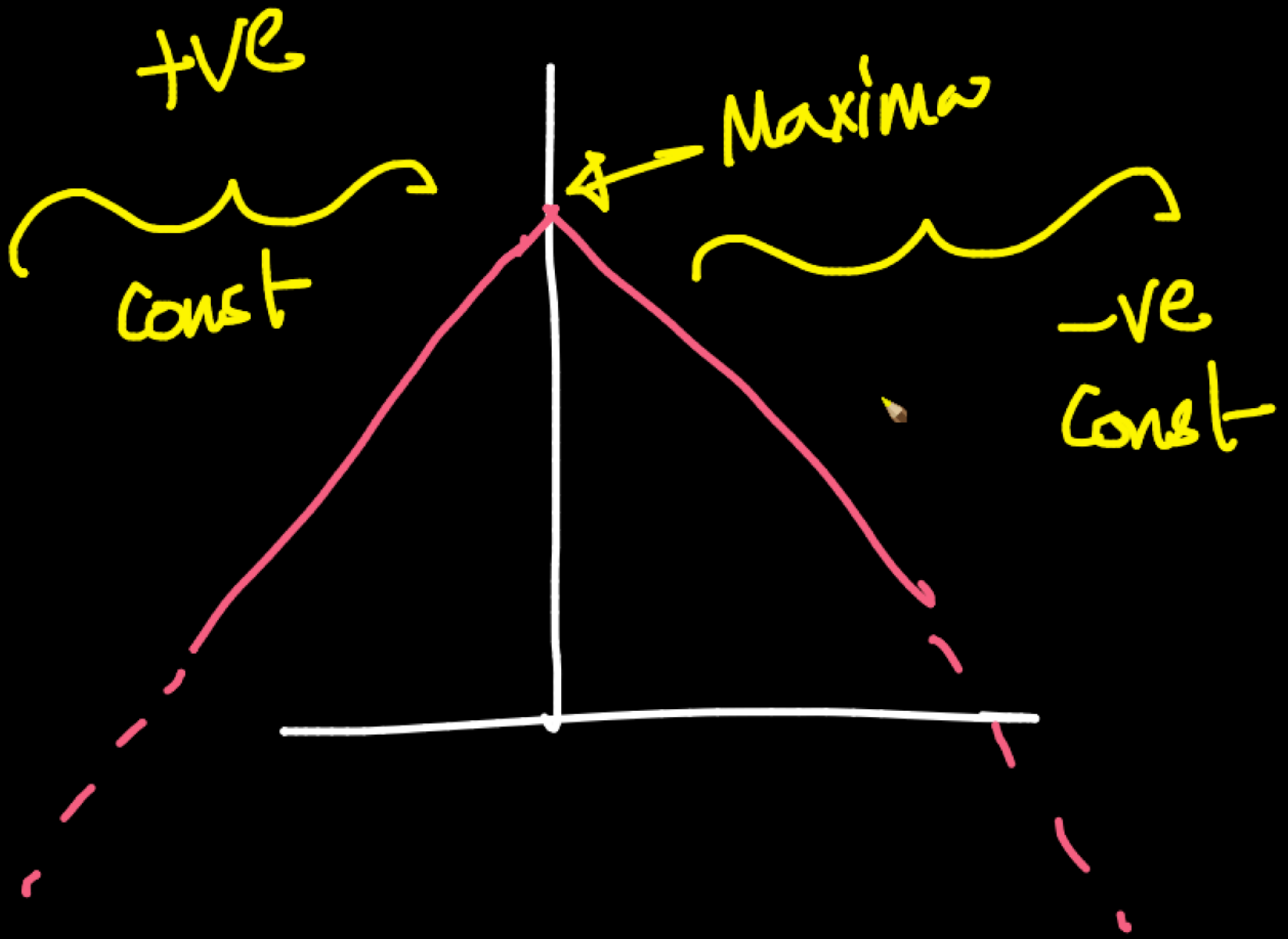
29 / 29

(*) Minima



(*) Max

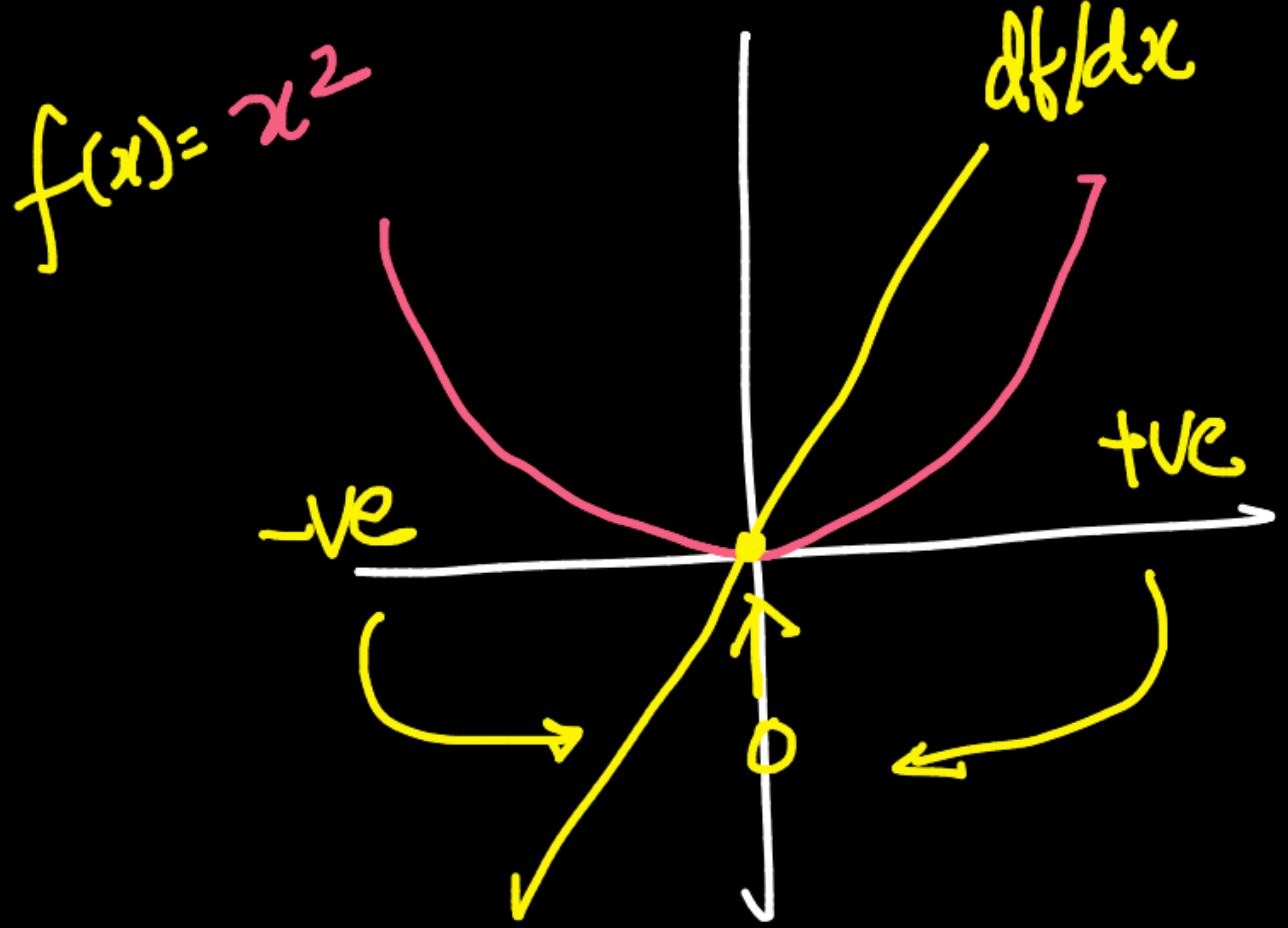



+ve
Const
Maxima
-ve
Const

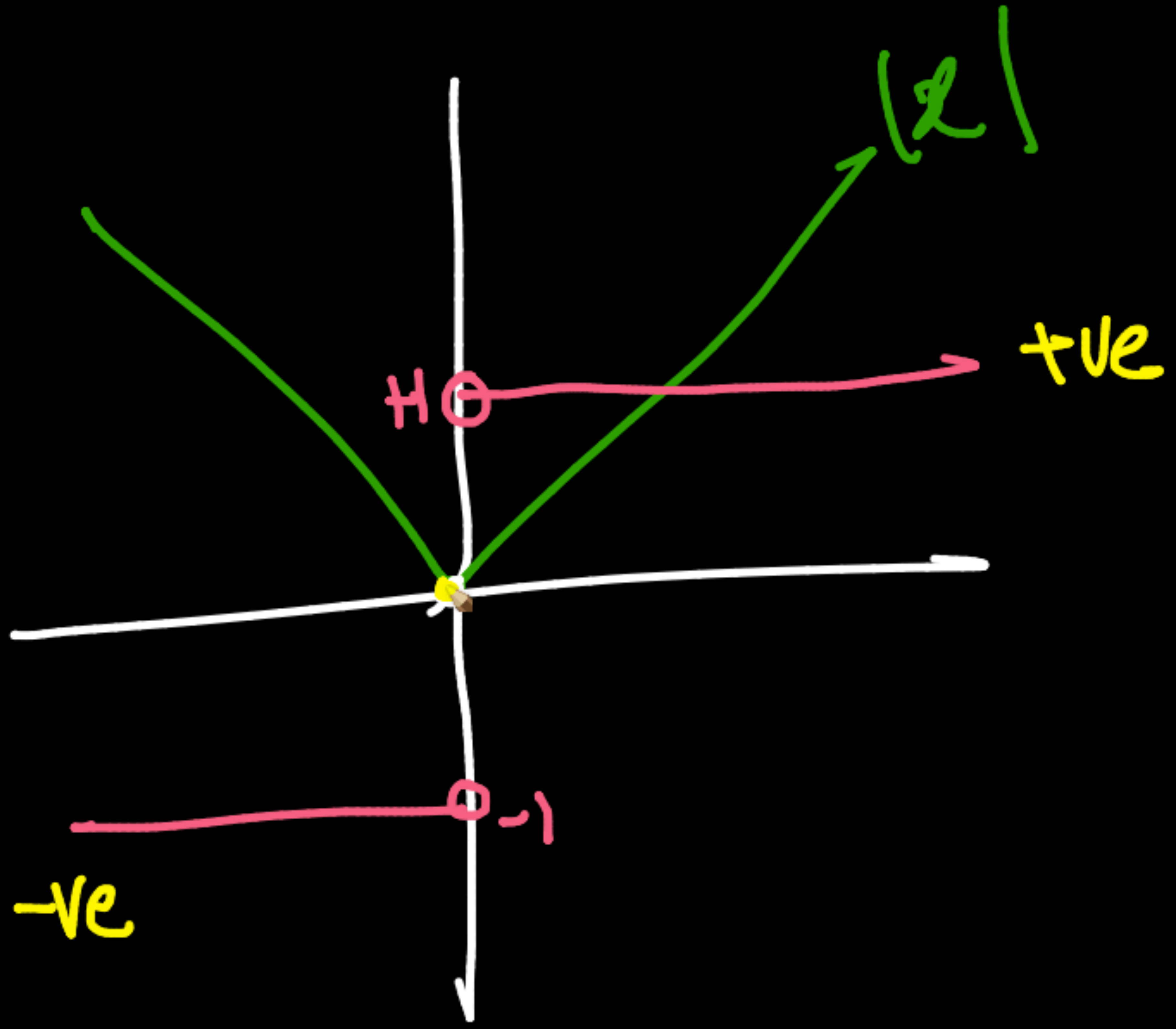
$$f(x) = x^2$$

$$\frac{df}{dx} = 2x$$

$$g(x) = \frac{df}{dx}$$



$$f(x) = |x|$$



$$\left.\frac{df}{dx}\right|_{x=x_1} = 0$$

Then $\overbrace{\text{Maxima or Minima}}^{\text{optima}}$ or $\boxed{\text{saddle}}$ @ x_1

little later

inflection points

Stationary pts

optima

L   R

minima : $-ve \rightarrow 0 \leftarrow +ve$

df/dx

maxima : $+ve \rightarrow 0 \leftarrow -ve$

df/dx

$\dfrac{df}{dx}$ $\uparrow$ incr $\qquad \dfrac{d^2f}{dx^2} > 0$

$\dfrac{df}{dx}$ $\downarrow$ decr

$\Downarrow$

$\dfrac{d^2f}{dx^2} < 0$

$f(x) = x^2$

Eqn

$\frac{df}{dx} = 0$

g(x)

$\frac{d}{dx} x^2 = \boxed{2x = 0}$

$x = 0$

optima

minima:

$\frac{d}{dx}\frac{df}{dx} = \frac{d^2 f}{dx^2}$ is +ve

maxima:

$\frac{d}{dx}\frac{df}{dx} = \frac{d^2 f}{dx^2}$ is -ve

$$f(x) = x^2 - 2x + 6$$

$$\frac{df}{dx} = 2x - 2$$

$$\frac{d^2f}{dx^2} = 2$$

Minima @ x_1

$$\left.\frac{df}{dx}\right|_{x_1} = 0$$

& $\left.\frac{d^2f}{dx^2}\right|_{x_1}$ is +ve

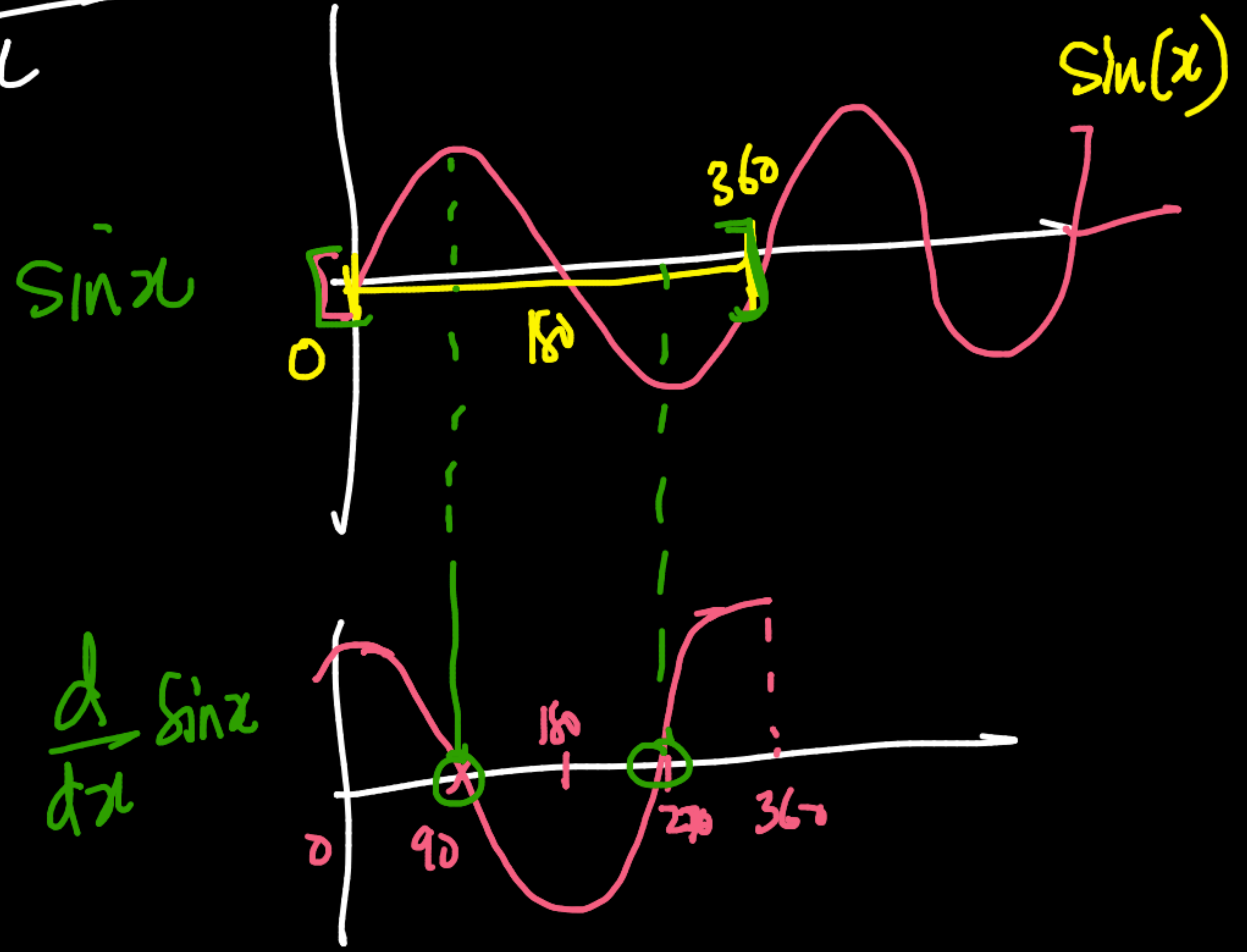
Maxima @ x_1

$\left.df/dx\right|_{x_1} = 0$ &

$\left.\frac{d^2f}{dx^2}\right|_{x_1}$ is -ve

$$\frac{d\sin(x)}{dx} = \cos(x)$$

∞-many minimas & Maximas.



(e.g)

$$\log(\underbrace{1+\exp(x)}_{u}) = f(x)$$

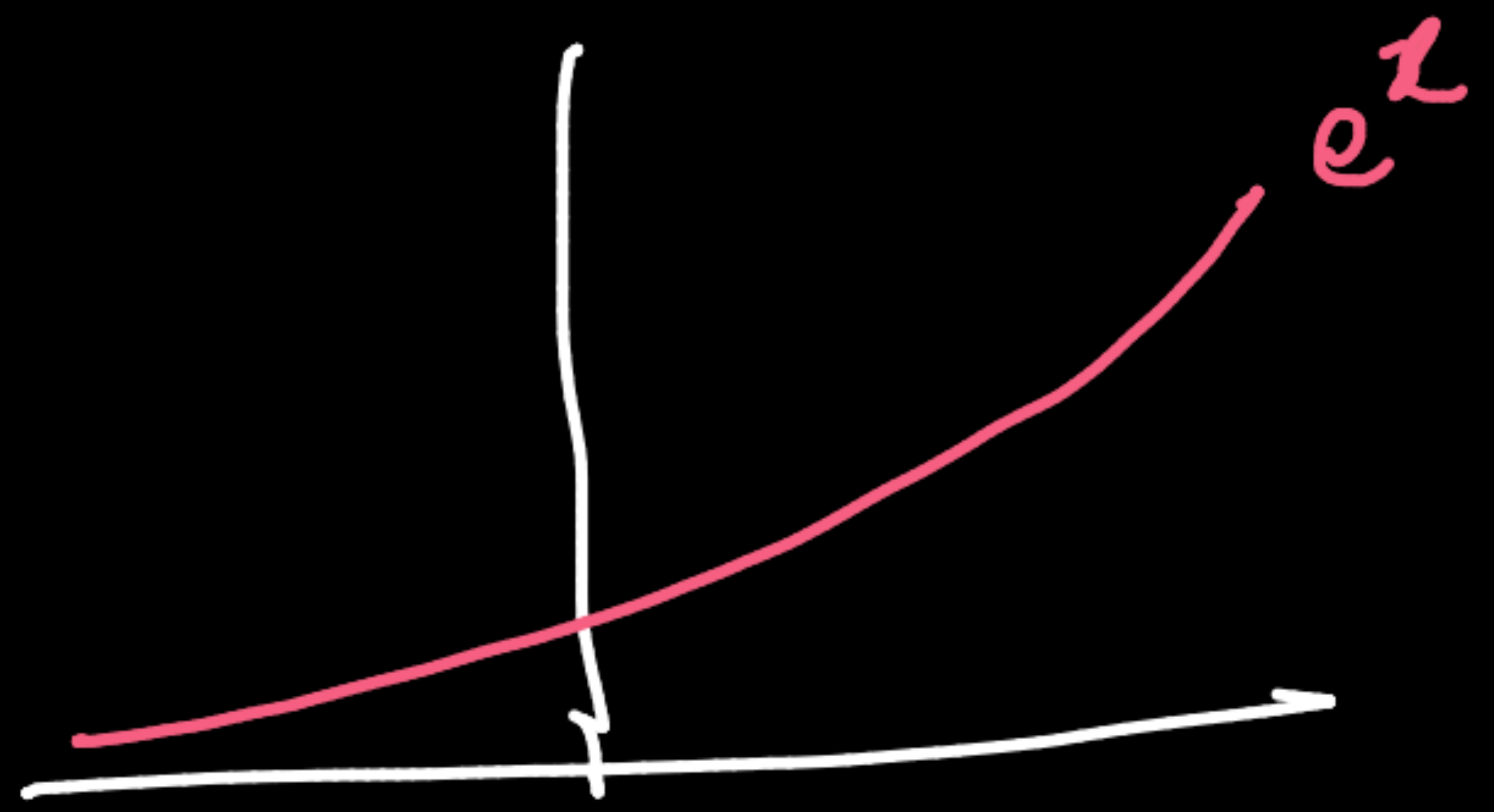
(Minima & Maxima)

$$\frac{df}{dx} = \frac{1}{1+\exp(x)} \cdot \exp(x) = 0$$

$$\left\{ \frac{\exp(x)}{1+\exp(x)} = 0 \right.$$

if $\exp(x) = 0$ ← optima

$e^x = 0$



$e^x \to 0$ as $x \to -\infty$

Q

(optima)

$$f(x) = \left(\overbrace{10 - (2x+3)}\right)^2$$

$$\frac{df}{dx} = 2\left(10 - (2x+3)\right) \cdot (-2)$$

$$= \boxed{-\cancel{4}\left(10 - (2x+3)\right) = 0}$$

(one)

$$2x+3 = 10 \Rightarrow x = 7/2$$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:17

$$\frac{d}{dx}\left(\frac{df}{dx}\right) = \frac{d}{dx}\left(-4\left(10-(2x+3)\right)\right)$$

$$= \underbrace{0+8+0} = 8 > 0$$

minima @ $x = 7/2$

Let

$$\frac{d^2f}{dx^2} = 8x\Big|_{x=7/2} = 8\times\frac{7}{2}$$

41 / 41

$$\boxed{\max f(x)}$$

one-variable

Classification

$$\max f(\underbrace{w, w_0}_{d+1 \text{ variables}}) = \sum_{i=1}^{n} y_i \frac{(w^T x_i + w_0)}{\|w\|}$$

$w_1, w_2 \ldots w_d$

MULTI VARIABLE CALCULUS

2D

3D $z = f(x, y) = x^2 + y^2$

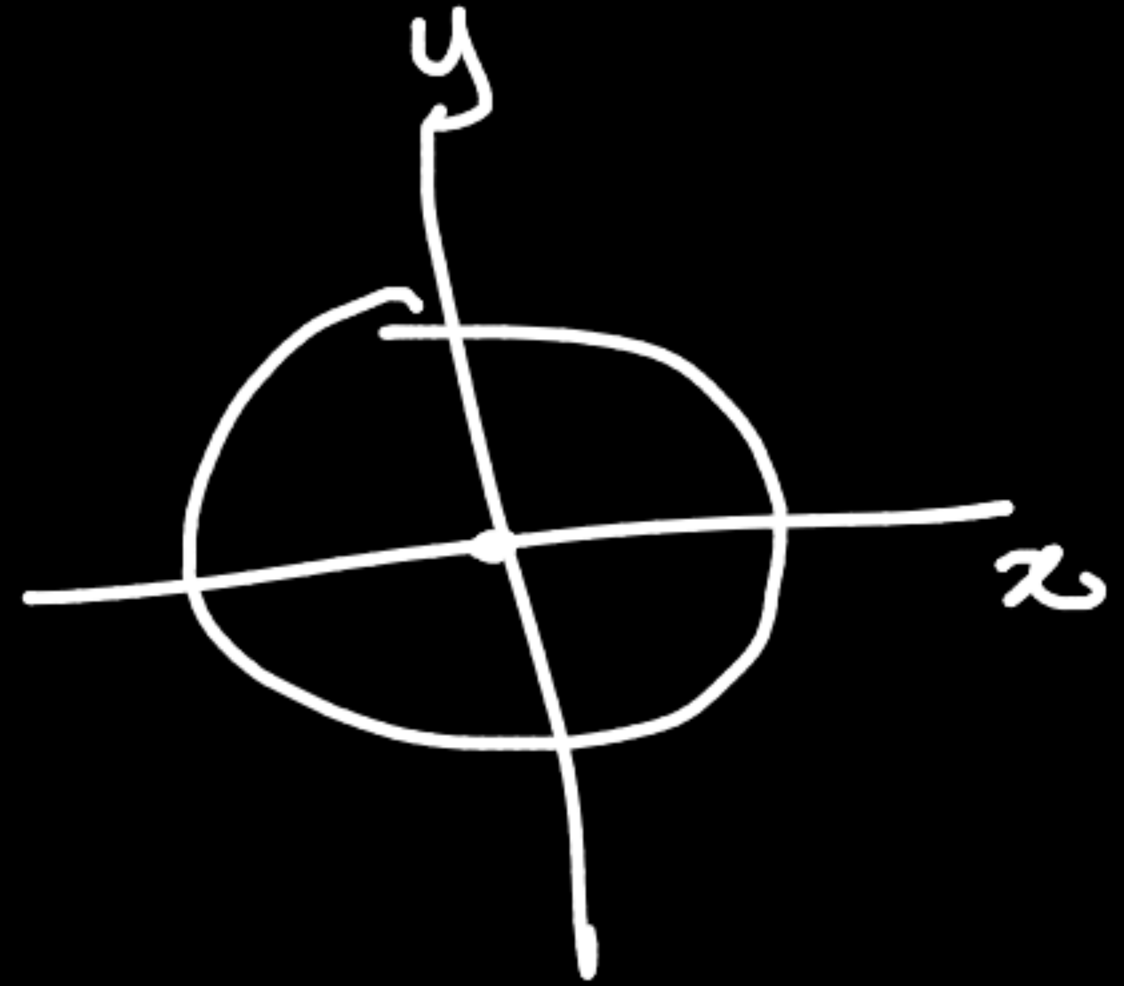
$$\boxed{z = x^2 + y^2}$$

3D

Circle!

2D

$$x^2 + y^2 = r^2$$



dependent on x & y

$$z = f(x, y) = x^2 + y^2$$

x & y are indep of one another

$$\frac{df}{dx} \qquad \frac{df}{dy}$$

ddn

2D-vector

Notations: $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$

Del

partial-der

Del

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

$$f(\underbrace{w_0, w_1, w_2 \dots w_d}_{d+1\text{-var}})$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial w_0} \\ \frac{\partial f}{\partial w_1} \\ \vdots \\ \frac{\partial f}{\partial w_d} \end{bmatrix}_{d+1}$$

minima for $f(x,y) = x^2 + y^2$

$$\frac{\partial f}{\partial x} = 2x + 0 = 0$$

y as a constant

$$\frac{\partial f}{\partial y} = 0 + 2y$$

x as constant

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:31

One-var!

@ optima: $\frac{df}{dx} = 0$

MVC!

@ optima: $\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \end{bmatrix}$

47 / 47

$$f(x) = x^2 + y^2$$

ignoring y

$$\begin{bmatrix} \frac{\partial f}{\partial x} \\[2em] \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x \\[2em] 2y \end{bmatrix} = \begin{bmatrix} 0 \\[2em] 0 \end{bmatrix}$$

ignoring x

$\Rightarrow x, y = (0, 0)$

optimum

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:35

Geom:-

minima → each axis

✓ visualize → hide on axis

49 / 49

Notation:

$$[x_1, \; x_2]^T = x$$

$$Z = x_1^2 + x_2^2 \quad \text{(let)}$$

$$\nabla_x Z = \begin{bmatrix} \frac{\partial z}{\partial x_1} \\ \frac{\partial z}{\partial x_2} \end{bmatrix}$$

$$f(w_0, w_1, \ldots, w_d)$$

$$\nabla_{\underline{w_0 \ldots w_d}} f = \nabla_{\underline{w, w_0}} f = \begin{bmatrix} \dfrac{\partial f}{\partial w_0} \\ \dfrac{\partial f}{\partial w_1} \\ \vdots \\ \dfrac{\partial f}{\partial w_d} \end{bmatrix}$$

OPTIMA

$$\nabla_x z = 0$$

vector

$\boxed{Q1}$

$$z = x^2 + y^2 - 2x + 2y + 6$$

$$\nabla z = \begin{bmatrix} 2x - 2 \\ 2y + 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} x = 1 \\ y = -1 \end{matrix}$$

Maxima ; Minima or saddle

(MVC)

tests

optima

Most fns in ML maxima/minima

$$f(x,y)$$

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial x} \qquad \frac{\partial}{\partial y}\frac{\partial f}{\partial x}$$

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial y} \qquad \frac{\partial}{\partial x}\frac{\partial f}{\partial y}$$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:49

minima: $\frac{df}{dx} = 0$ & $\frac{d^2f}{dx^2}$ +ve

maxima: $\frac{df}{dx} = 0$ & d^2f/dx^2 -ve

Exception: (Saddle points)

$f(x) = x^3$

$\frac{df}{dx} = 3x^2 = 0 \Rightarrow x = 0$

$\frac{d^2f}{dx^2} = 6x$ @ $x=0$ $\boxed{= 0}$

neither +ve nor -ve

$\frac{df}{dx} = 0$

NOT Maxima

not Minima

$\nabla f = \vec{0}$

55 / 58

plot(x^3)

All Shopping Images News Maps More Tools

About 69,80,00,000 results (0.52 seconds)

Graph for x^3

More info

People also ask

How do I plot x 3?

What does X 3 mean on a graph?

What is an x 3 graph called?

How do you plot in math?

Feedback

https://www.wolframalpha.com › examples › math

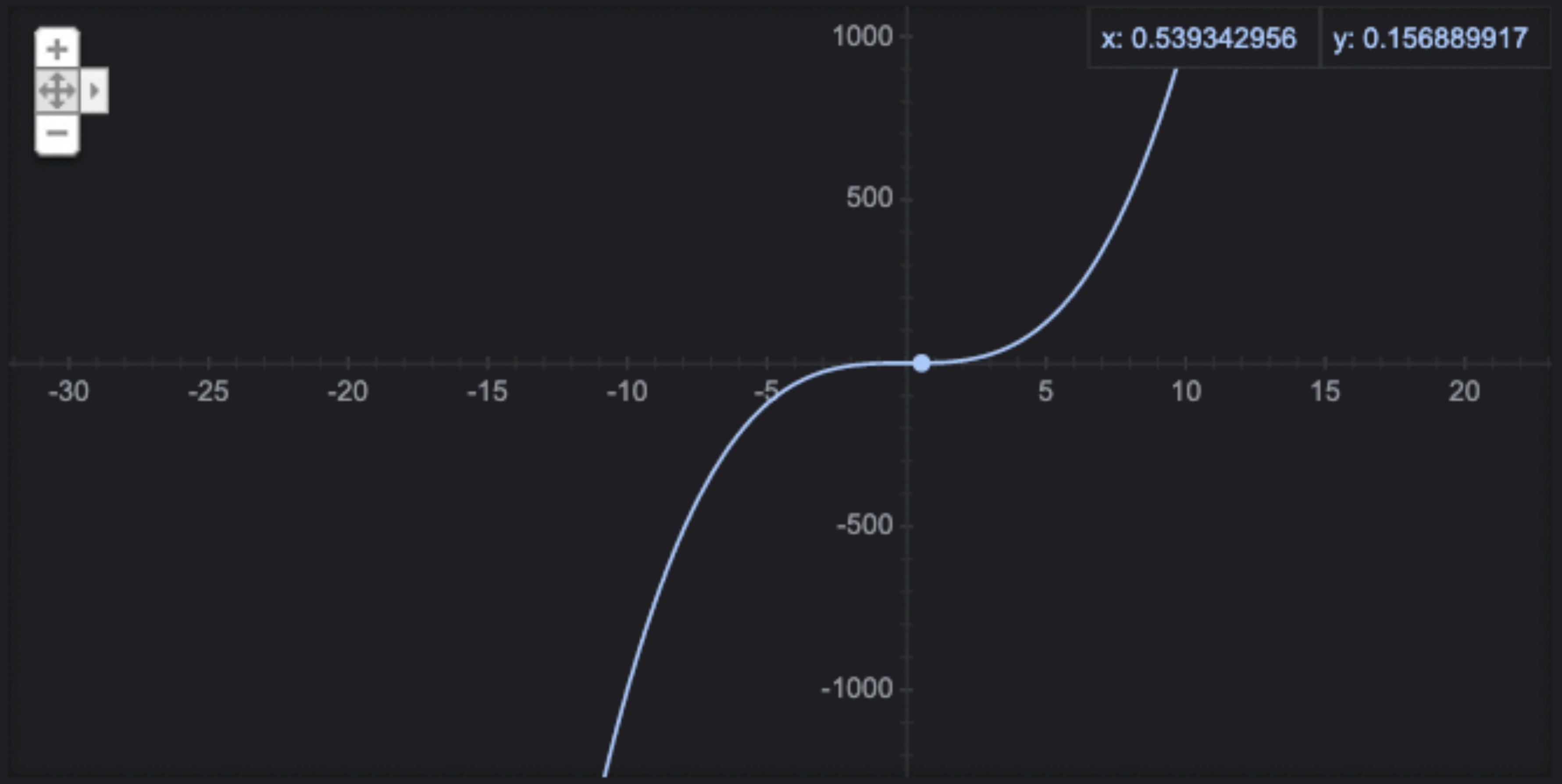
56 / 56

plot(x^3)

All Shopping Images News Maps More Tools

About 69,80,00,000 results (0.52 seconds)

Graph for x^3



More info

People also ask

How do I plot x 3?

What does X 3 mean on a graph?

What is an x 3 graph called?

How do you plot in math?

Feedback

https://www.wolframalpha.com › examples › mat

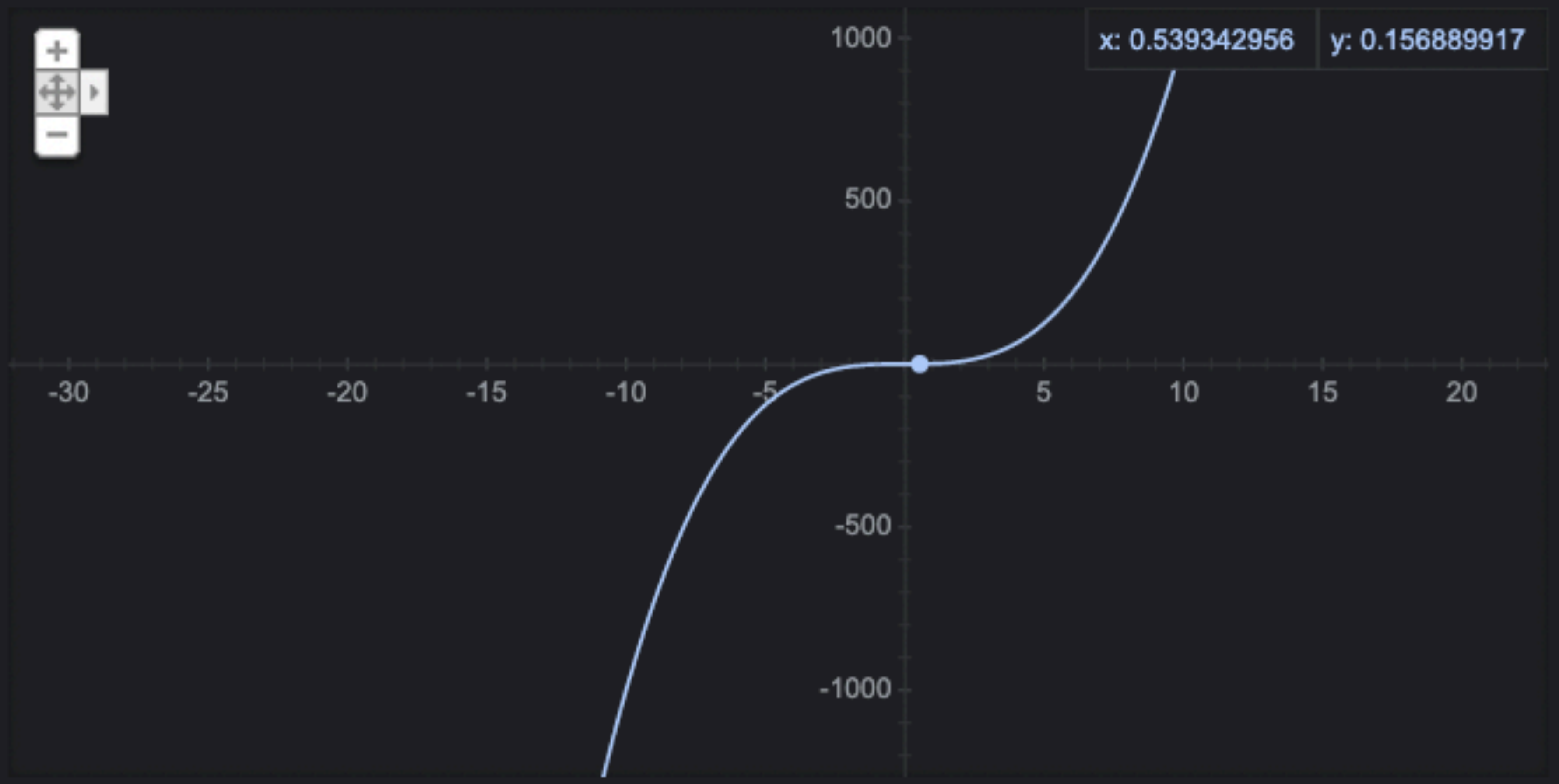
57 / 59

plot(x^3)

All · Shopping · Images · News · Maps · More · Tools

About 69,80,00,000 results (0.52 seconds)

Graph for x^3



More info

People also ask

- How do I plot x 3?
- What does X 3 mean on a graph?
- What is an x 3 graph called?
- How do you plot in math?

Feedback

https://www.wolframalpha.com › examples › matl

58 / 58

3D

$$f(x) =$$

$$f(x) = x^4$$

$$\frac{df}{dx} = 4x^3 = 0 \quad @ \quad x = 0$$

$$\frac{d^2f}{dx^2} = 12x^2 \quad @ \quad x = 0 \quad = \boxed{0}$$

$\rightarrow$ +ve $\rightarrow$ $\boxed{\text{minima}}$

$\rightarrow$ -ve

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:52

GradientDescent.ipynb - Colab | Untitled1.ipynb - Colaboratory | plot(x^4) - Google Search

google.com/search?q=plot%28x%5E4%29&rlz=1C5CHFA_enIN958IN958&ei=Zu5BYoXME6ixmAXus5igAg&ved=0ahUKEwiFl8O_qOn2AhWoGKYKHe4ZBiQQ4dUDCA8&uact=5&oq=plot%28x%5E4%29&gs_lcp=Cgdnd3Mtd2l6EAMy...

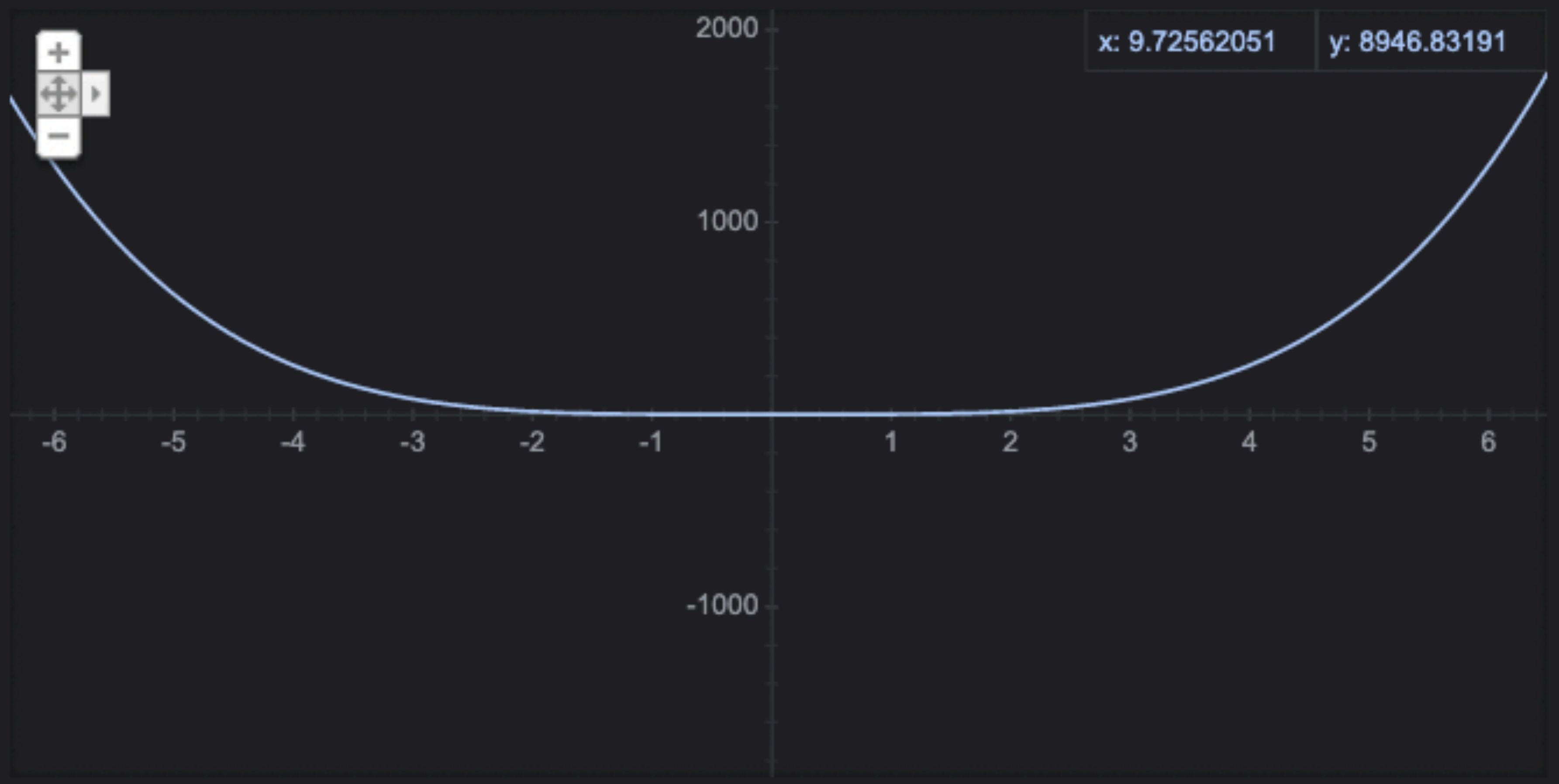
Google

plot(x^4)

All Shopping Images Maps News More Tools

About 6,02,00,00,000 results (0.54 seconds)

Graph for x^4



More info

https://www.mathway.com › popular-problems › Algebra

Graph x=4 | Mathway

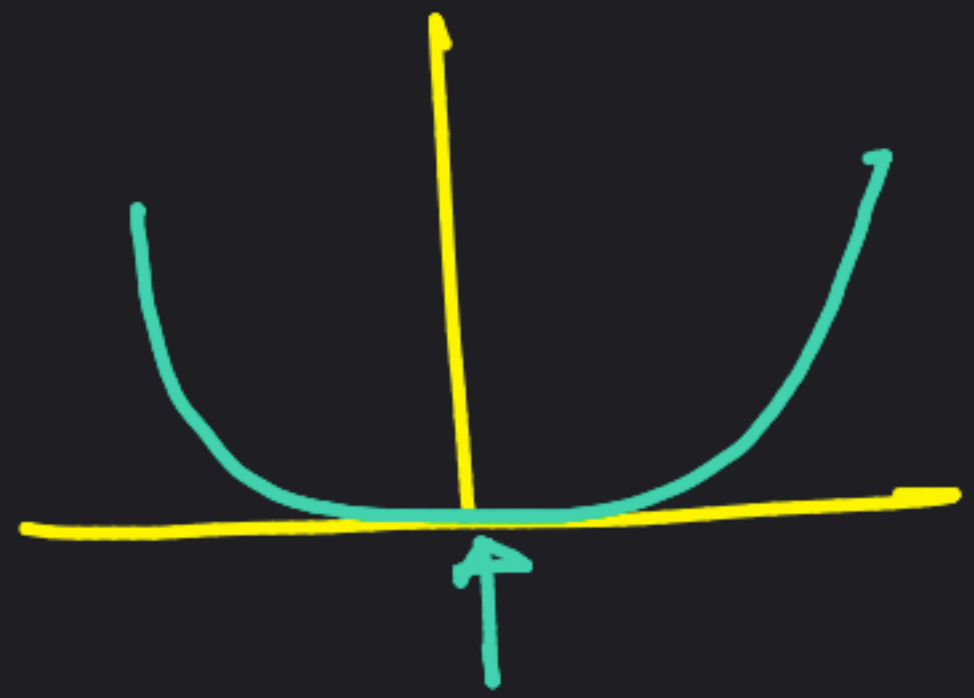
Since **x=4 x = 4** is a vertical line, there is no y-intercept and the slope is undefined. Slope: Undefined. y-intercept: No y-intercept. image of **graph**.

People also ask

How do you plot x 4?

Is x 4 horizontal or vertical?

How do you plot in math?



$$f(x) = x^4$$

$$\frac{\partial f}{\partial x} =$$

61 / 61

Special-cases:

Minima & Maxima → First non-zero higher order derivatives

+ve

-ve

$\boxed{\text{Sabuj}}$

$f(x) = x^4$

$\frac{df}{dx} = 4x^3 = 0$ $\quad \boxed{x=0}$

$\frac{d^2f}{dx^2} = 12x^2 \rightarrow 0$

$\frac{\partial^3 f}{\partial x^3} = 24x \rightarrow 0$

$\frac{\partial^4 f}{\partial x^4} = \boxed{24} > 0$

$\downarrow$

minima

$f(x) = x^3$

→ odd

↳ even

2nd order test ✓

higher order test

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:58

65 / 66

3D

Minima

Maxima

$$f(x,y) = x^2 - y^2$$

$x = 0$
$y = 0$

$$\nabla f = \begin{bmatrix} \frac{\partial z}{\partial x} \\ \frac{\partial z}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x \\ -2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Handle:

→ saddle point

→ ✓ higher order derivatives

Comp. algo

hacks

Gradient Descent

2o

Gradient Descent

→ Comp. algo to find the minimas/Maximas

↳ Math → built on it

$\boxed{Q}$

$$f(x) = \overbrace{x^8 + x^7 + x^4}$$

(optima)

$$\frac{df}{dx} = 8x^7 + 7x^6 + 4x^3 = 0$$

$$\boxed{8x^4 + 7x^3 + 4 = 0}$$ ← solving this eqn

$\log(x)$

$\exp(x)$

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Gradient Descent

23:11

Higher-order derivative test [edit]

The *higher-order derivative test* or *general derivative test* is able to determine whether a function's critical points are maxima, minima, or points of inflection for a wider variety of functions than the second-order derivative test. As shown below, the second-derivative test is mathematically identical to the special case of $n = 1$ in the higher-order derivative test.

Let f be a real-valued, sufficiently differentiable function on an interval $I \subset \mathbb{R}$, let $c \in I$, and let $n \geq 1$ be a natural number. Also let all the derivatives of f at c be zero up to and including the n-th derivative, but with the $(n+1)$th derivative being non-zero:

$$f'(c) = \cdots = f^{(n)}(c) = 0 \quad \text{and} \quad f^{(n+1)}(c) \neq 0.$$

There are four possibilities, the first two cases where c is an extremum, the second two where c is a (local) saddle point:

- If n is odd and $f^{(n+1)}(c) < 0$, then c is a local maximum.
- If n is odd and $f^{(n+1)}(c) > 0$, then c is a local minimum.
- If n is even and $f^{(n+1)}(c) < 0$, then c is a strictly decreasing point of inflection.
- If n is even and $f^{(n+1)}(c) > 0$, then c is a strictly increasing point of inflection.

Since n must be either odd or even, this analytical test classifies any stationary point of f, so long as a nonzero derivative shows up eventually.

Example [edit]

Say, we want ... $x^6 + 5$ at the point $x = 0$. To do this, we

Higher-order derivative test [edit]

The *higher-order derivative test* or *general derivative test* is able to determine whether a function's critical points are maxima, minima, or points of inflection for a wider variety of functions than the second-order derivative test. As shown below, the second-derivative test is mathematically identical to the special case of $n = 1$ in the higher-order derivative test.

Let f be a real-valued, sufficiently differentiable function on an interval $I \subset \mathbb{R}$, let $c \in I$, and let $n \geq 1$ be a natural number. Also let all the derivatives of f at c be zero up to and including the n-th derivative, but with the $(n+1)$th derivative being non-zero:

$$f'(c) = \cdots = f^{(n)}(c) = 0 \quad \text{and} \quad f^{(n+1)}(c) \neq 0.$$

There are four possibilities, the first two cases where c is an extremum, the second two where c is a (local) saddle point:

- If n is odd and $f^{(n+1)}(c) < 0$, then c is a local maximum.
- If n is odd and $f^{(n+1)}(c) > 0$, then c is a local minimum.
- If n is even and $f^{(n+1)}(c) < 0$, then c is a strictly decreasing point of inflection.
- If n is even and $f^{(n+1)}(c) > 0$, then c is a strictly increasing point of inflection.

Since n must be either odd or even, this analytical test classifies any stationary point of f, so long as a nonzero derivative shows up eventually.

Example [edit]

Say, we want ... $x^6 + 5$ at the point $x = 0$. To do this, we

$f(x) = x^3$

$f' = 3x^2 = 0$

$x = 0$

$\overset{2}{f''} = 6x \to 0$

$\overset{3}{f'''} = 6$

Higher-order derivative test [edit]

The *higher-order derivative test* or *general derivative test* is able to determine whether a function's critical points are maxima, minima, or points of inflection for a wider variety of functions than the second-order derivative test. As shown below, the second-derivative test is mathematically identical to the special case of $n = 1$ in the higher-order derivative test.

Let f be a real-valued, sufficiently differentiable function on an interval $I \subset \mathbb{R}$, let $c \in I$, and let $n \geq 1$ be a natural number. Also let all the derivatives of f at c be zero up to and including the n-th derivative, but with the $(n+1)$th derivative being non-zero:

$$f'(c) = \cdots = \overset{2}{f^{(n)}}(c) = 0 \quad \text{and} \quad \overset{3}{f^{(n+1)}}(c) \neq 0.$$

There are four possibilities, the first two cases where c is an extremum, the second two where c is a (local) saddle point:

- If n is odd and $f^{(n+1)}(c) < 0$, then c is a local maximum.
- If n is odd and $f^{(n+1)}(c) > 0$, then c is a local minimum.
- If n is even and $f^{(n+1)}(c) < 0$, then c is a strictly decreasing point of inflection.
- If n is even and $f^{(n+1)}(c) > 0$, then c is a strictly increasing point of inflection.

Since n must be either odd or even, this analytical test classifies any stationary point of f, so long as a nonzero derivative shows up eventually.

Example [edit]

Say, we want ... $x^6 + 5$ at the point $x = 0$. To do this, we

$x=0$

$f(x)=x^4$

$f^1=3x^3$

$f^2=9x^2$

$f^3=18x$

$f^4=18$

Higher-order derivative test [edit]

The *higher-order derivative test* or *general derivative test* is able to determine whether a function's critical points are maxima, minima, or points of inflection for a wider variety of functions than the second-order derivative test. As shown below, the second-derivative test is mathematically identical to the special case of $n = 1$ in the higher-order derivative test.

Let *f* be a real-valued, sufficiently differentiable function on an interval $I \subset \mathbb{R}$, let $c \in I$, and let $n \geq 1$ be a natural number. Also let all the derivatives of *f* at *c* be zero up to and including the *n*-th derivative, but with the $(n+1)$th derivative being non-zero:

$$f'(c) = \cdots = f^{(n)}(c) = 0 \quad \text{and} \quad f^{(n+1)}(c) \neq 0.$$

3 4

$\boxed{f'' = 0}$

There are four possibilities, the first two cases where *c* is an extremum, the second two where *c* is a (local) saddle point:

- If *n* is odd and $f^{(n+1)}(c) < 0$, then *c* is a local maximum.
- If *n* is odd and $f^{(n+1)}(c) > 0$, then *c* is a local minimum.
- If *n* is even and $f^{(n+1)}(c) < 0$, then *c* is a strictly decreasing point of inflection.
- If *n* is even and $f^{(n+1)}(c) > 0$, then *c* is a strictly increasing point of inflection.

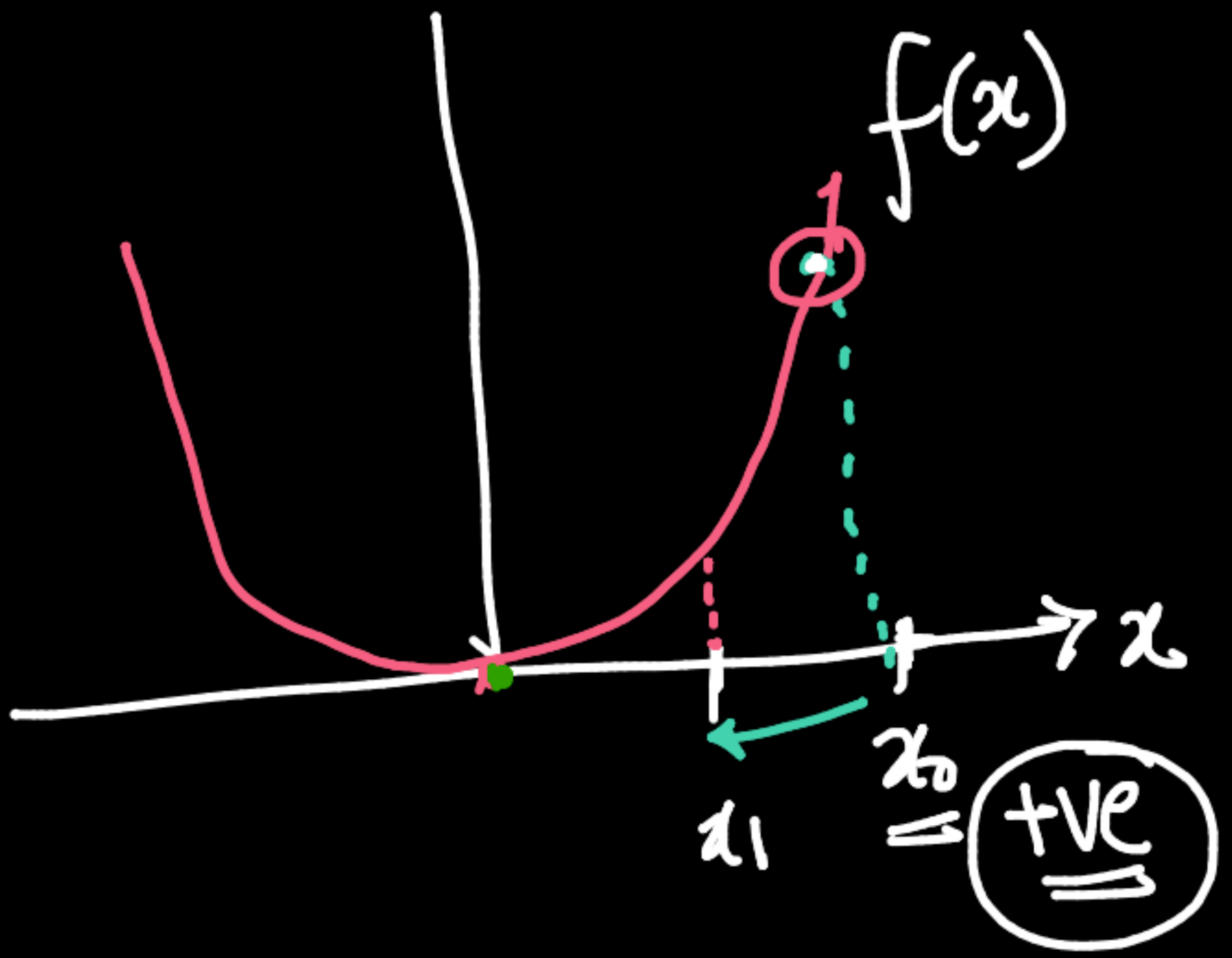
Since *n* must be either odd or even, this analytical test classifies any stationary point of *f*, so long as a nonzero derivative shows up eventually.

Example [edit]

Say, we want ... $x^6 + 5$ at the point $x = 0$. To do this, we

74 / 74

[minima]



$x_1 < x_0$

iterative ← Gradient Descent for min

(1) pick x_0 randomly

$\left.\frac{df}{dx}\right|_{x_0} = +ve$ [incr fn]

(1)

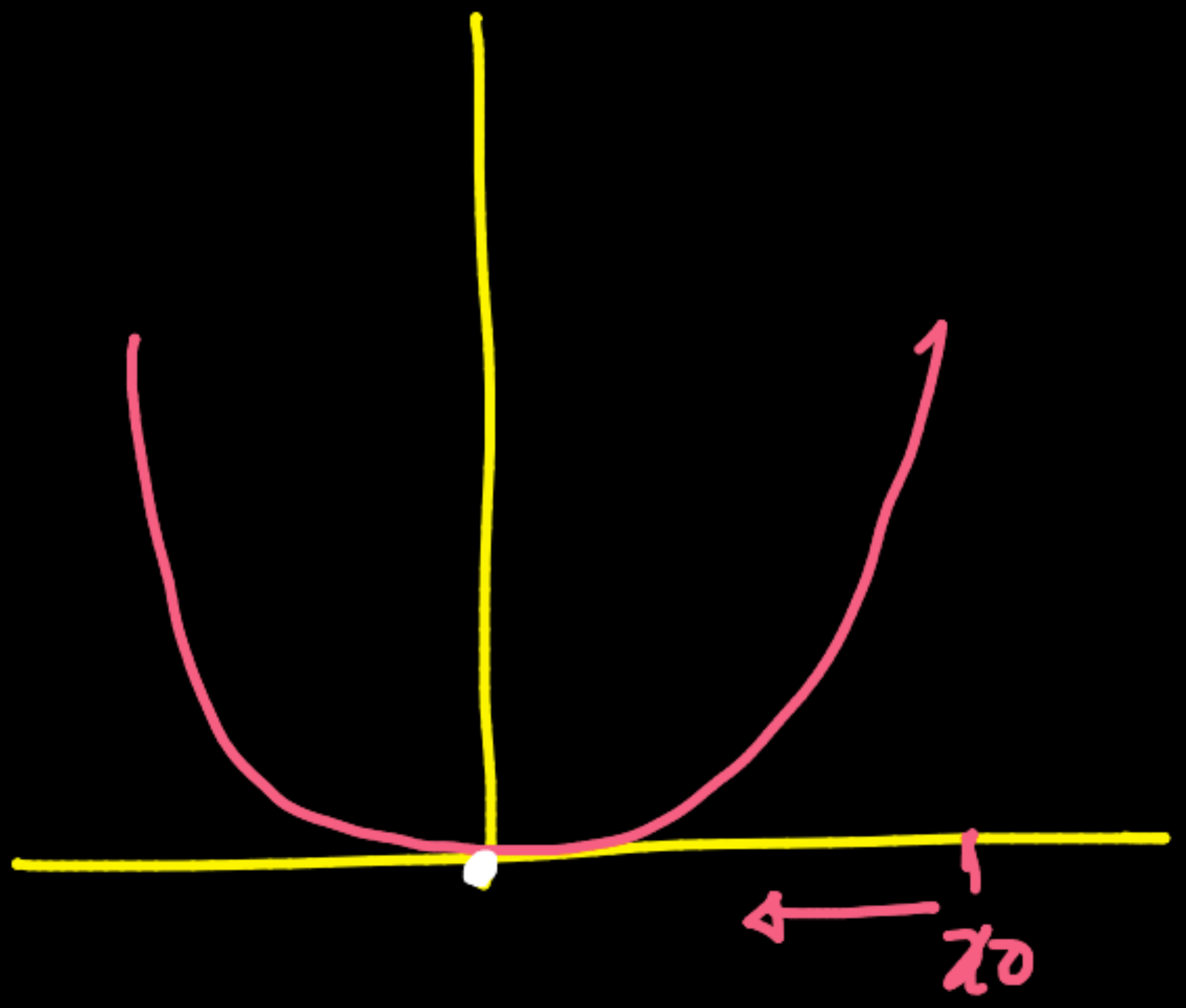
const : 0.1

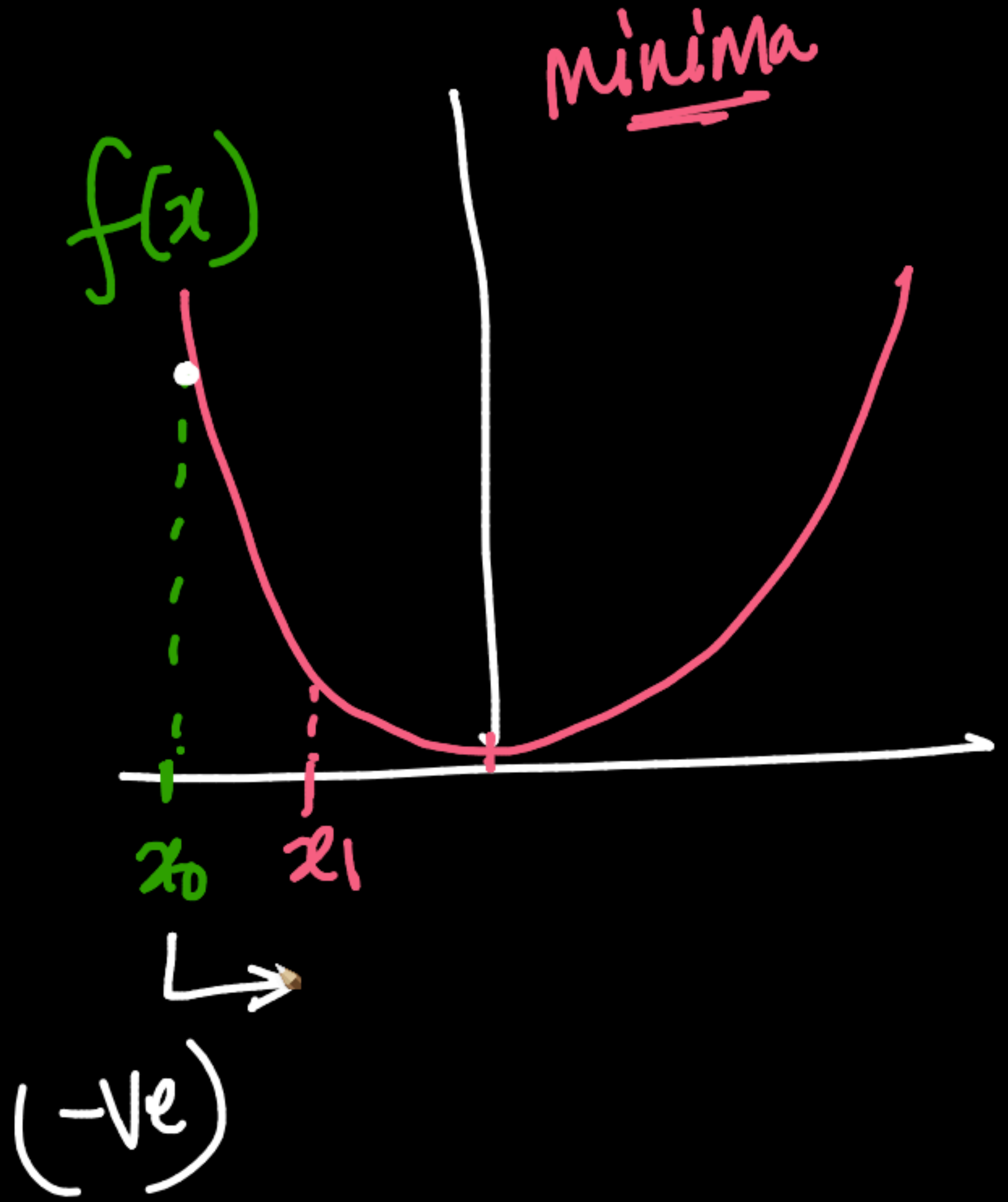
$x_1 = x_0 + \eta \left(-\frac{\partial f}{\partial x}\right)_{x=x_0}$

learning rate

Update eqn:

$$x_1 = x_0 + \overset{0.1}{\eta}\left(-\frac{\partial f}{\partial x}\right)_{x=x_0}$$





$$\left.\frac{df}{dx}\right|_{x_0} : \text{decr} : -ve$$

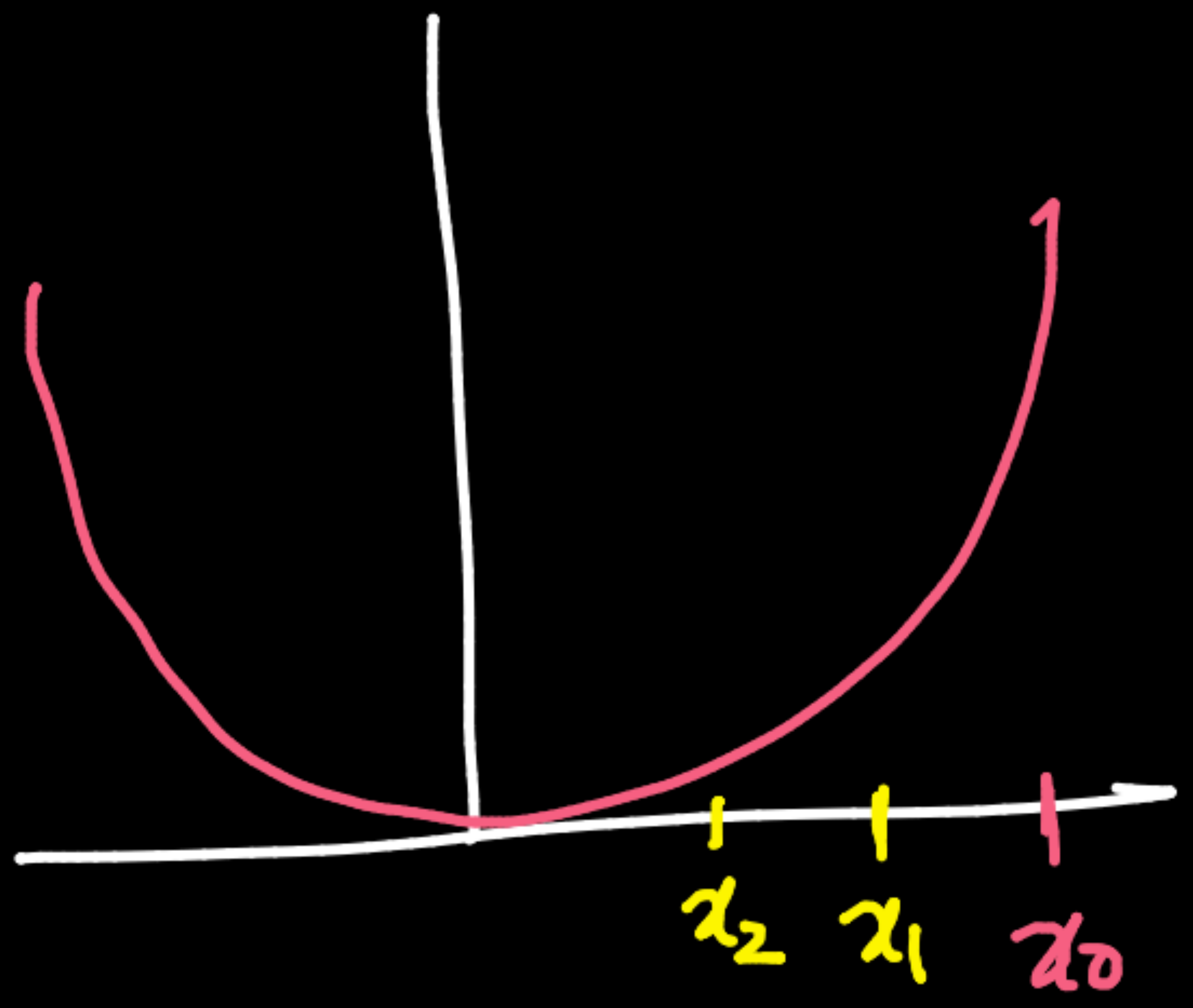
-ve

0.1

$$x_1 = x_0 + \eta \left(-\frac{\partial f}{\partial x} \right)_{x=x_0}$$

$\eta > 0$ +ve

small



2nd-derivatives ✓

iterate

x_0 : randomly

$$x_1 = x_0 + \eta \left(-\frac{\partial f}{\partial x}\right)_{x_0}$$

$$x_2 = x_1 + \eta \left(-\frac{\partial f}{\partial x}\right)\Big|_{x_1}$$

$\vdots$

K

$$\left.\frac{\partial f}{\partial x}\right|_{x_K} \approx 0$$

Update Eqn:

$$x_{i+1} = x_i + \eta \overbrace{-\frac{\partial f}{\partial x}}\Big|_{x_i}$$

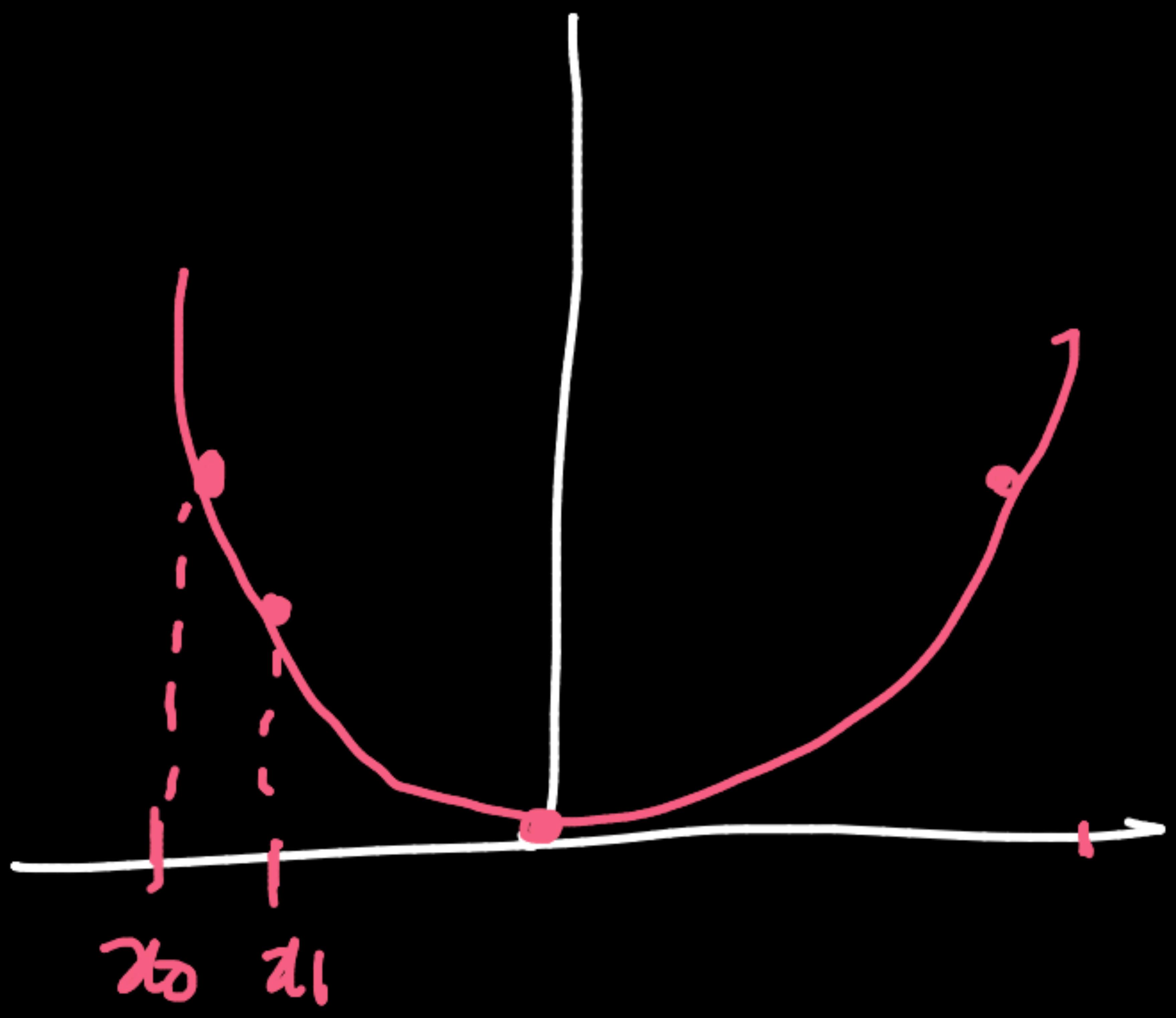
$$x_{i+1} = x_i - \eta \frac{\partial f}{\partial x}\Big|_{x_i}$$

$$\max_{w, w_0} f(w, w_0) = \min_{w, w_0} -f(w, w_0)$$

$\{$ min. misclassification rate

Gradient → slope

Descent → down



ball rolling

$f(x)$: ____ (minima)

$f'(x)$: ____

$\frac{df}{dx} \approx 0$

-0.0001 —— 0.0001

$\eta = 0.01$

x_0 = randomly initialize

$df\_dx = f'(x_0)$

$x\_old = x_0$

while $|df\_dx - 0| > 0.0001$

τ

$x\_new = x\_old - \eta f'(x\_old)$

```
[1] import random
    import math
    import numpy
```

```
[ ] #define f(x)
    def f(x):
      return math.pow(x,2)+2*x+10

    # df_dx; we could have used the limits concepts form prev class too.
    def derF(x):
      return 2*x+2
```

$x^2+2x+10$

$\frac{\partial f}{\partial x} = 2x+2 = 0$

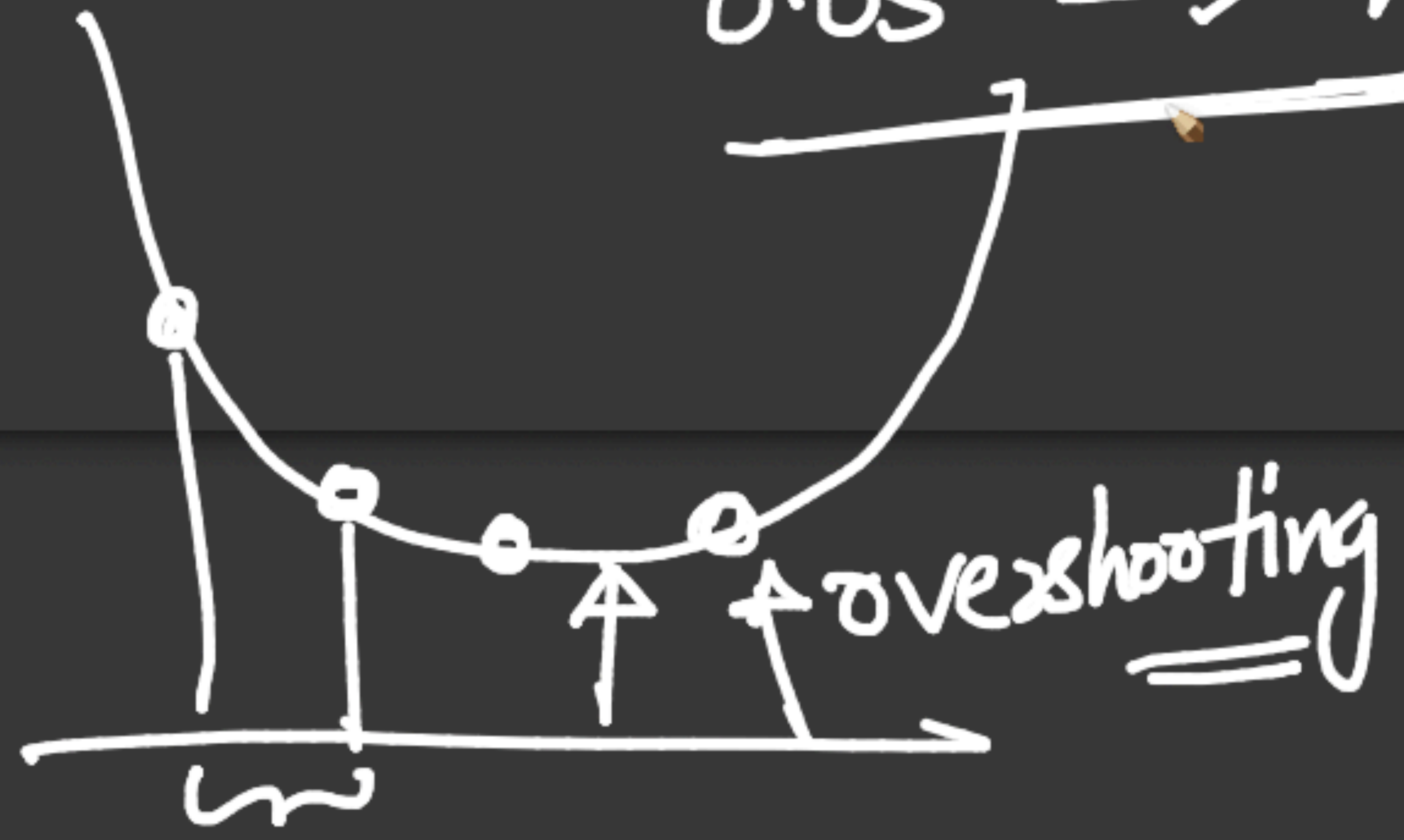
$\boxed{x=-1}$

```
[2] def gradientDescentUnivariate():

      x0 = random.random()
      i=0 # iterations
      eta = 0.01 # learning rate

      xi = x0
      df_dx = derF(xi) # df_dx at x0

      while abs(df_dx - 0.0) > 0.001: # threshold = 0.001
```

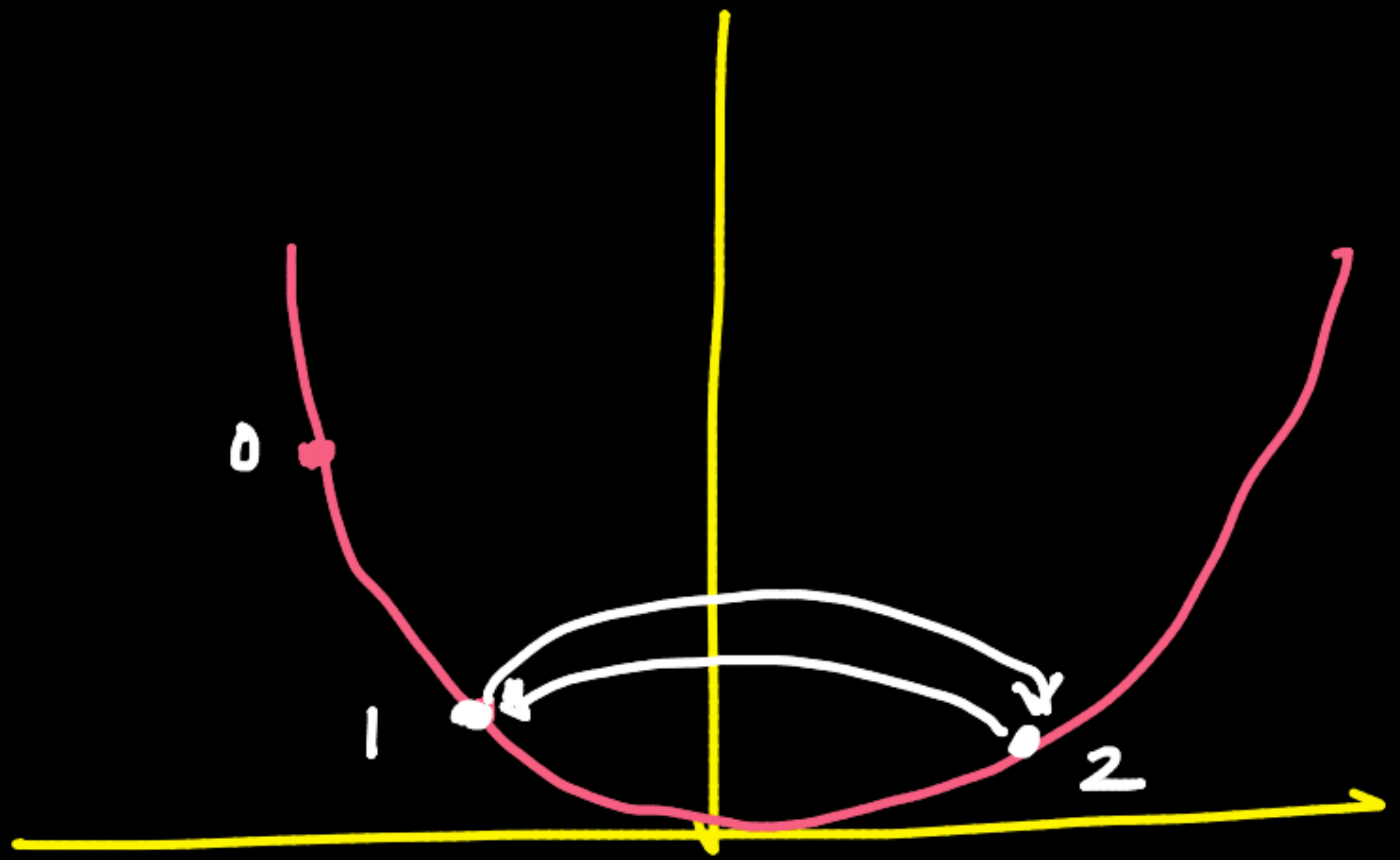
```
71
-0.999285227519035
0.0015883832910332796
**********************
72
-0.999356704767131 6
0.0014295449619299294
**********************
73
-0.9994210342904184
0.0012865904657368699
**********************
74
-0.9994789308613765
0.0011579314191632 05
**********************
75
-0.9995310377752389
0.0010421382772469734
**********************
-0.999577933997715
```

$$x_{i+1} = x_i - \eta \left.\frac{\partial f}{\partial x}\right|_{x_i}$$

η	#iterations
0.01	→ 377
0.1	→ ?
0.05	→ 75



Oscillating



boundary cases

overshooting:

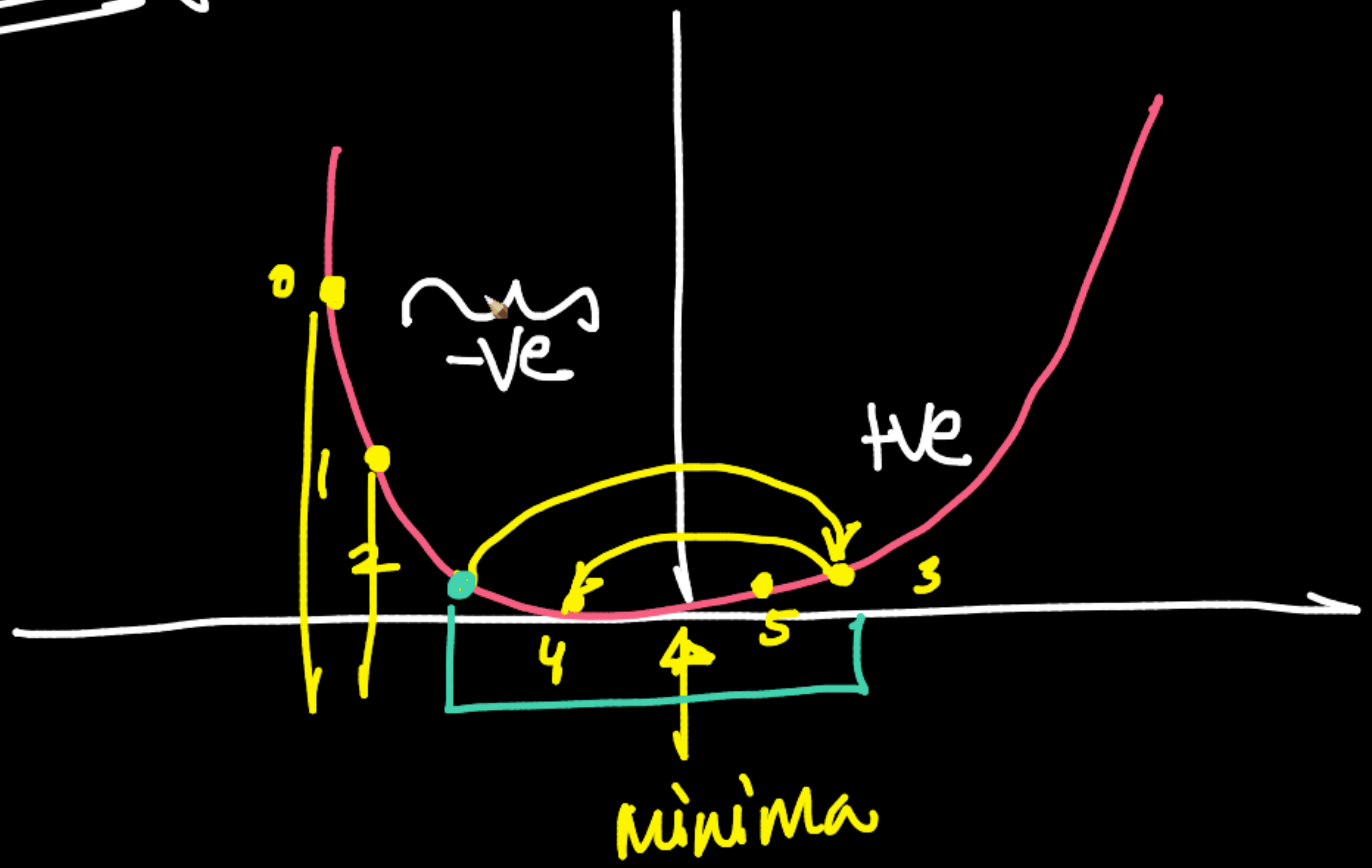
run faster?

comp. hack
↓

0.1 ← η : control on

η a function of iter-number

(iter ↑ η ↓)

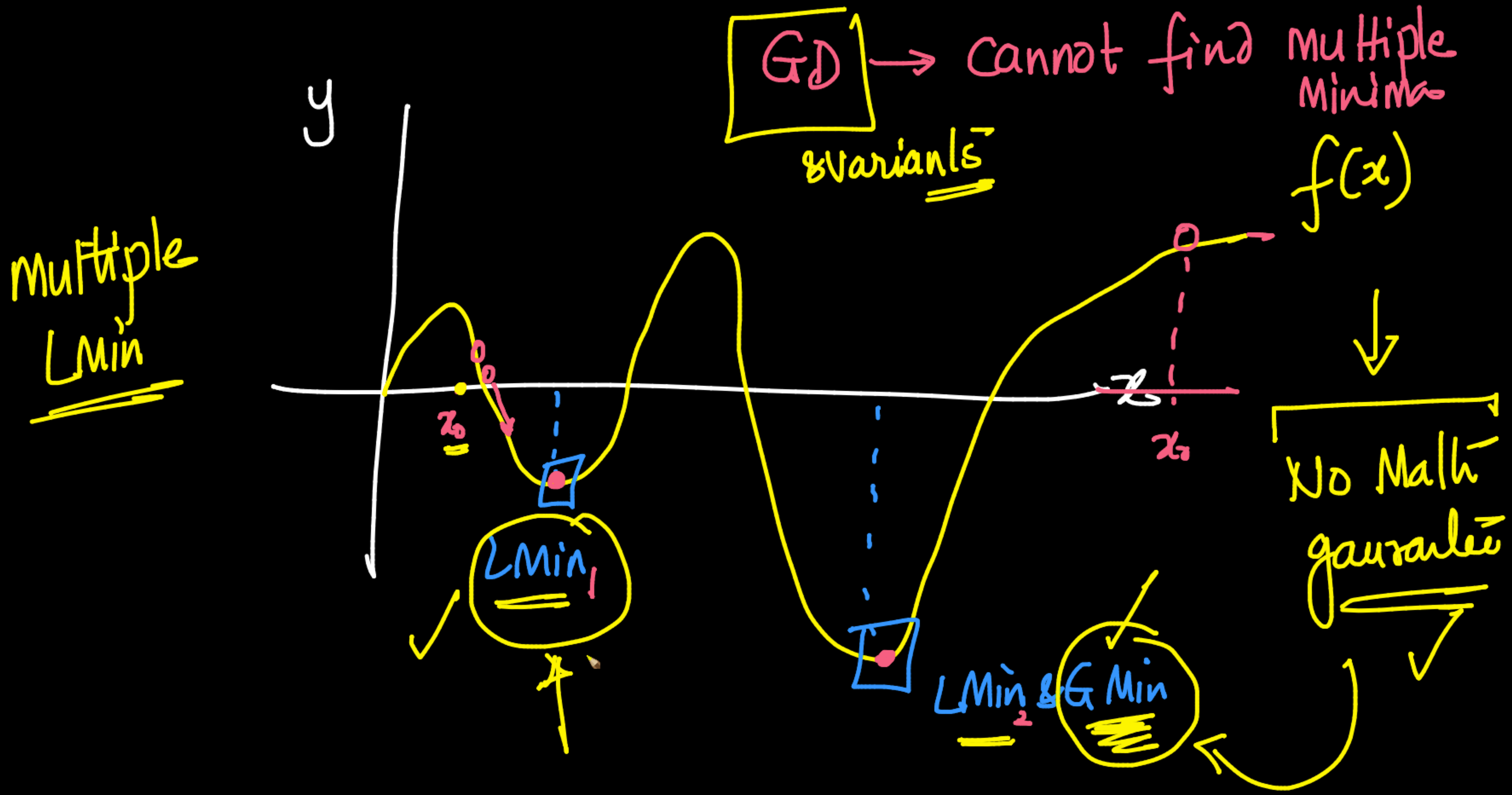


Next-class:

→ GD in MVC

→ Classfn ;

Constrained optmzn


GD → cannot find multiple minima
variants
f(x)
No Math guarantee
Multiple LMin
y
x_0
x_3
LMin₁
LMin₂ & G Min

GD or variant

[prog; internally

Classfn problem → Sep HP (w, w_0) → Optmzn prob

↓

Maxima & minima

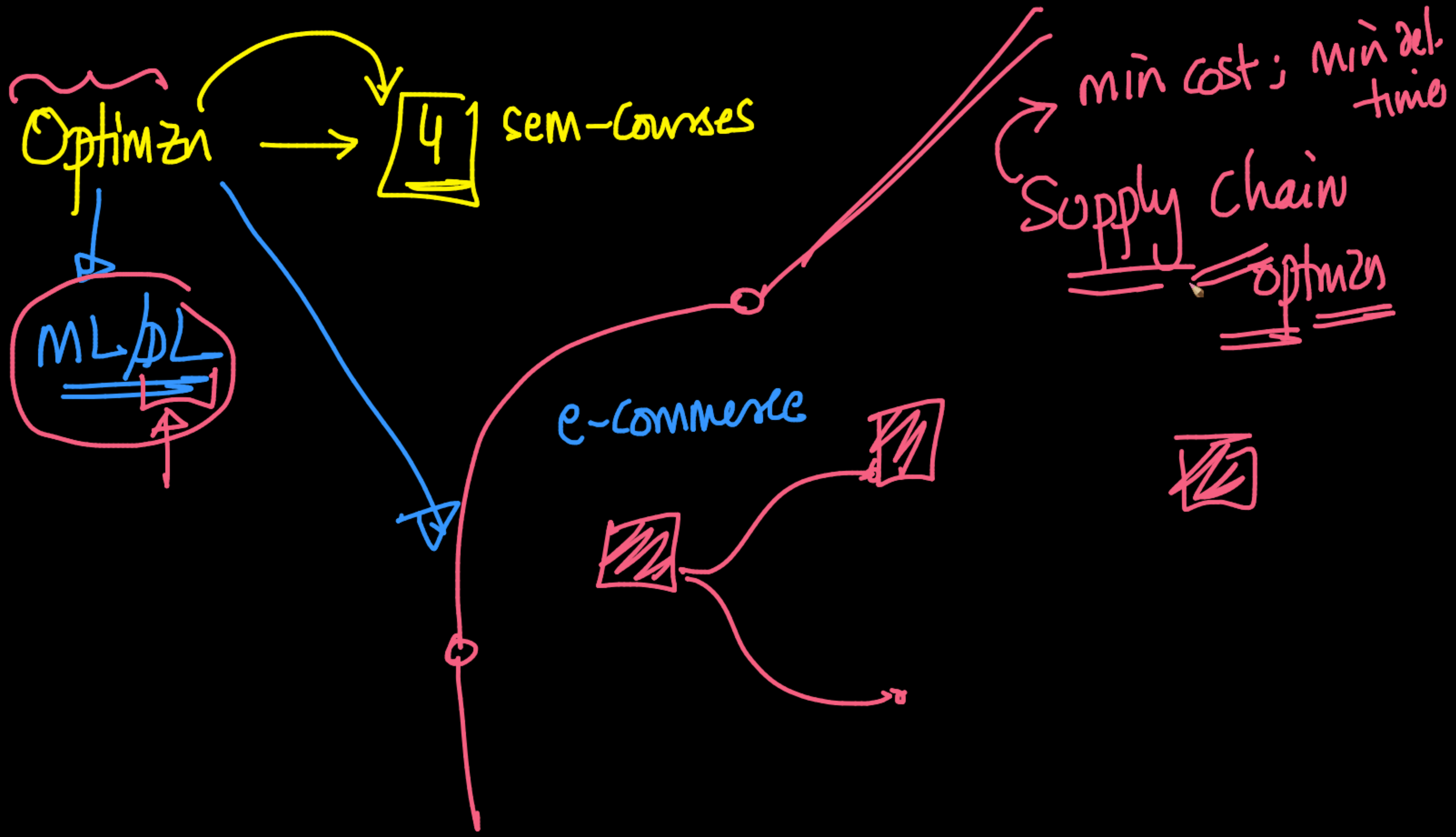
↓

GD

MVC

optimzn

$$\left\{ \max_{w, w_0} \sum_{i=1}^{n} y_i \left(\frac{w^T x_i + w_0}{\|w\|} \right) \right.$$


Optimzn
4
sem-Courses
ML/DL
e-Commerce
min cost ; min del. time
Supply chain
optimzn


Chrome File Edit View History Bookmarks Profiles Tab Window Help
0:11
GradientDescent.ipynb - Colab
Untitled1.ipynb - Colaboratory
Python Get Random Float Num
New Tab
Search Google or type a URL
Update
Gmail
Images
Google
Search Google or type a URL
My Drive
Learning
YouTube
InterviewBit S...
GitHub
Scaler Acade...
Colaboratory
Google Maps
(99+) Feed
Add shortcut
Photo by NASA Image Library
94 / 94