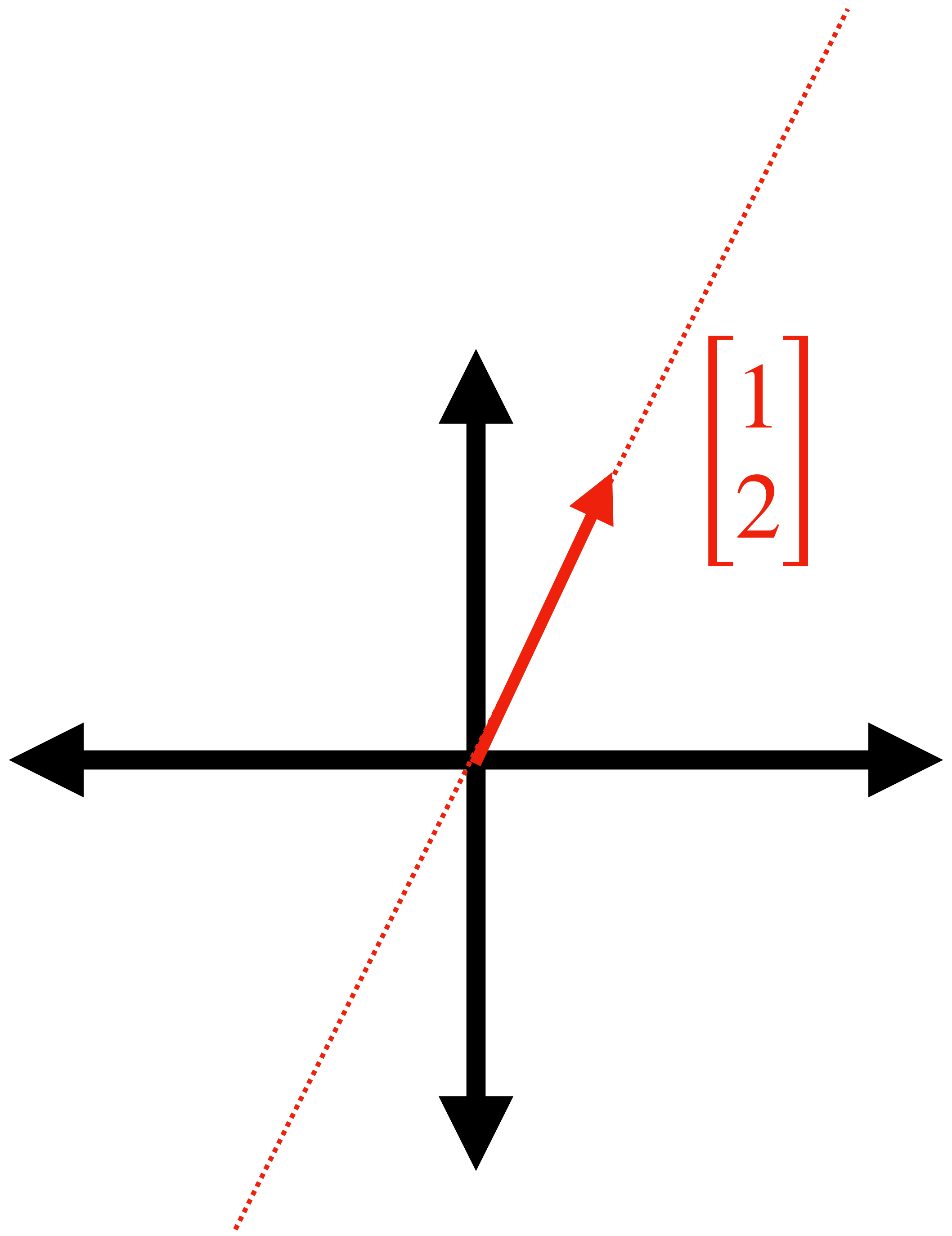
# Calculus I

**Dhruva Venkita Raman**

# My week

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= v_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + v_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$



# My week

They forgot their matrix was low rank!

Heterosynaptic Circuits Are Universal Gradient Machines

Liu Ziyin[1,2], Isaac Chuang[1], Tomaso Poggio[1]

[1]*Massachusetts Institute of Technology*
[2]*NTT Research*

June 12, 2025

**Abstract**

We propose a design principle for the learning circuits of the biological brain. The principle states that almost any dendritic weights updated via heterosynaptic plasticity can implement a generalized and efficient class of gradient-based meta-learning. The theory suggests that a broad class of biologically plausible learning algorithms, together with the standard machine learning optimizers, can be grounded in heterosynaptic circuit motifs. This principle suggests that the phenomenology of (anti-) Hebbian (HBP) and heterosynaptic plasticity (HSP) may emerge from the same underlying dynamics, thus providing a unifying explanation. It also suggests an alternative perspective of neuroplasticity, where HSP is promoted to the primary learning and memory mechanism, and HBP is an emergent byproduct. We present simulations that show that (a) HSP can explain the metaplasticity of neurons, (b) HSP can explain the flexibility of the biology circuits, and (c) gradient learning can arise quickly from simple evolutionary dynamics that do not compute any explicit gradient. While our primary focus is on biology, the principle also implies a new approach to designing AI training algorithms and physically learnable AI hardware. Conceptually, our result demonstrates that contrary to the common belief, gradient computation may be extremely easy and common in nature.

# Previously

**Numbers (fields)**

**Vector spaces, matrices**

**Probability spaces**

**What are they?**

**Operations and algebra?**

Addition, subtraction…

Dot product, matrix multiplication, norm…

Intersection, conditioning, …

**Code and solve concrete questions with them?**

# Today

**Differential quantities**

What are they?

Operations and algebra

How do we compute them?

**Differential equations**

**Optimisation**

**Machine learning**

# Course structure

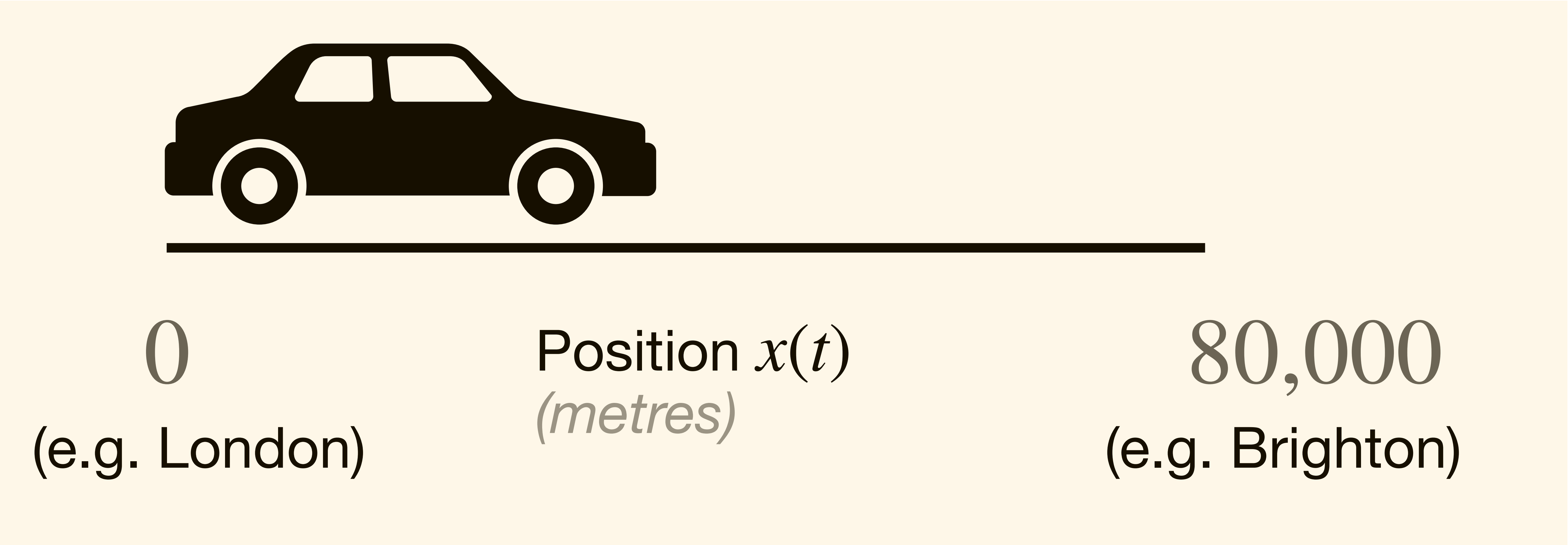
Differentiation is abstract

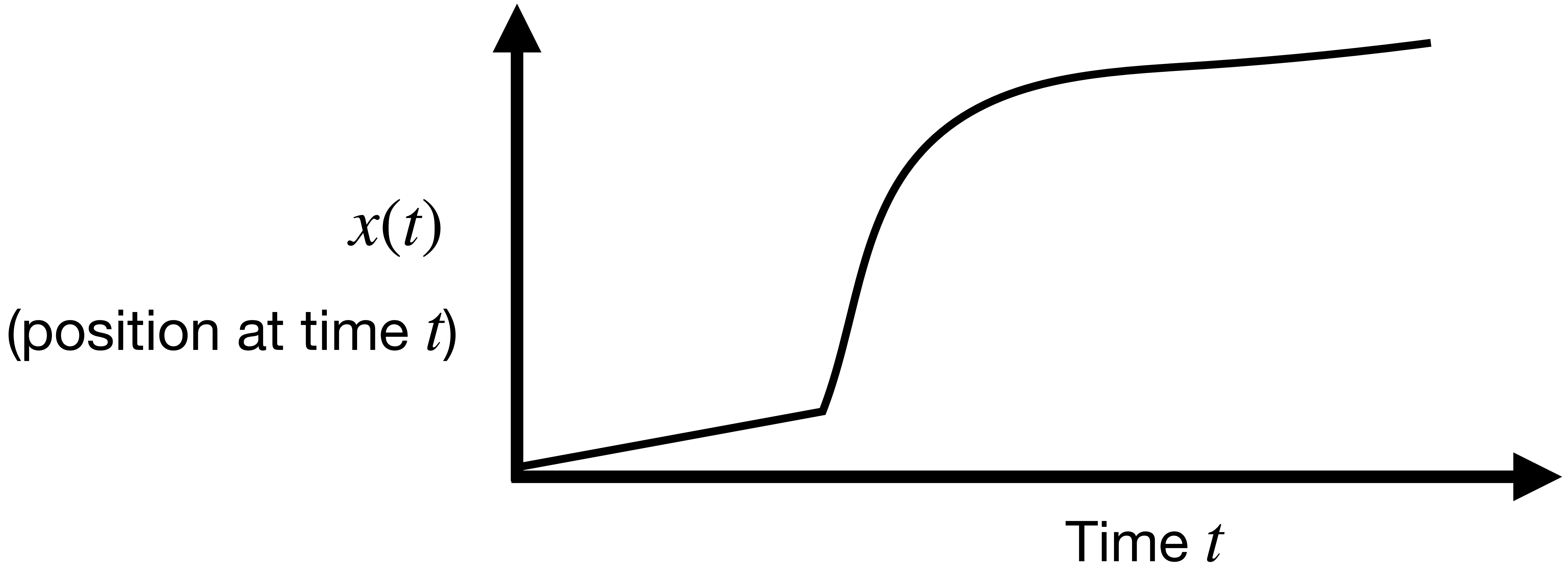
It’s uses are concrete and
numerous

**Modelling**
**dynamical systems**

**Optimisation**
**(AI)**

# Position of car as a function of time



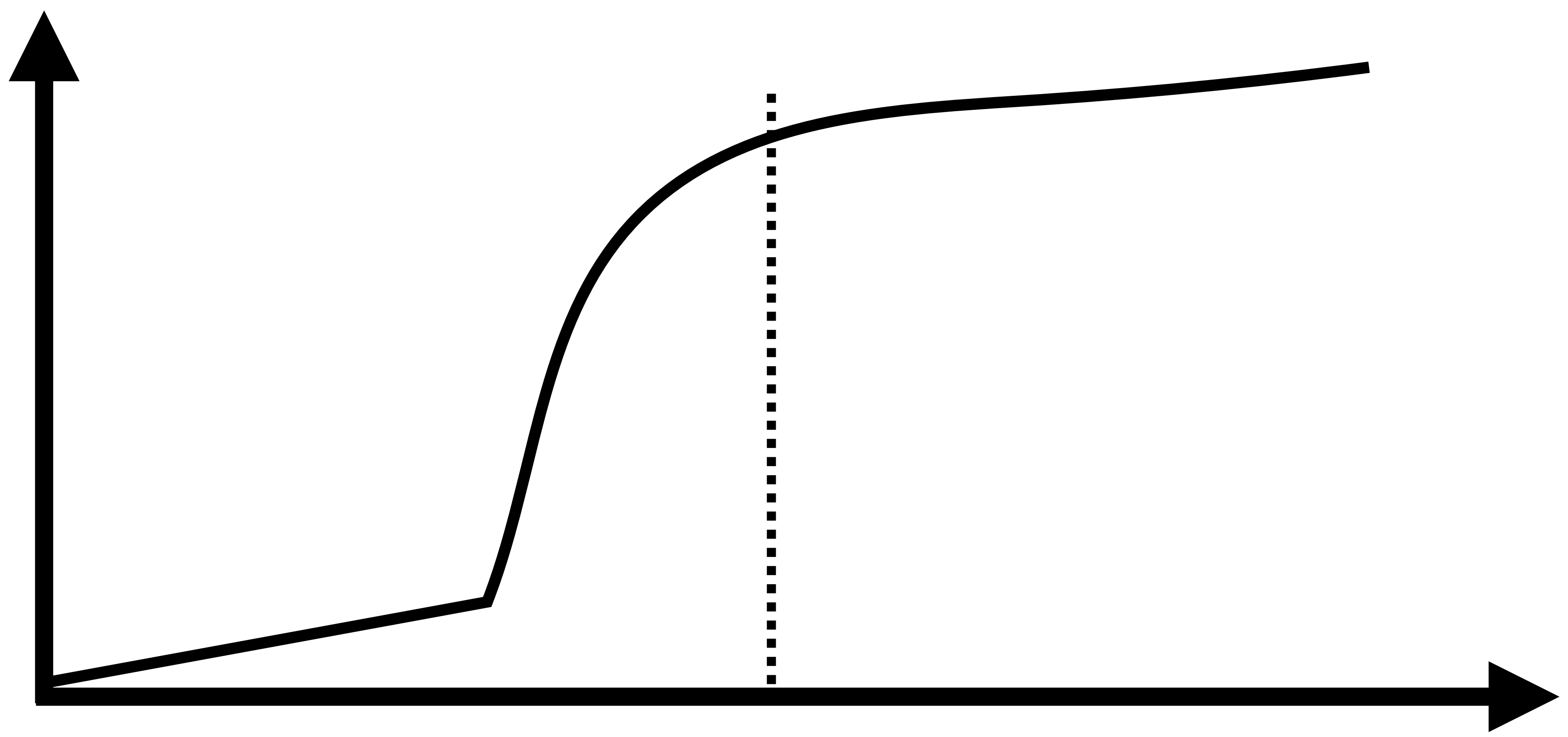



# Velocity of car at a point in time

N.B. Velocity has a direction, so can be negative
Speed is the magnitude of velocity

Freeze-frame

*What does this even mean?*

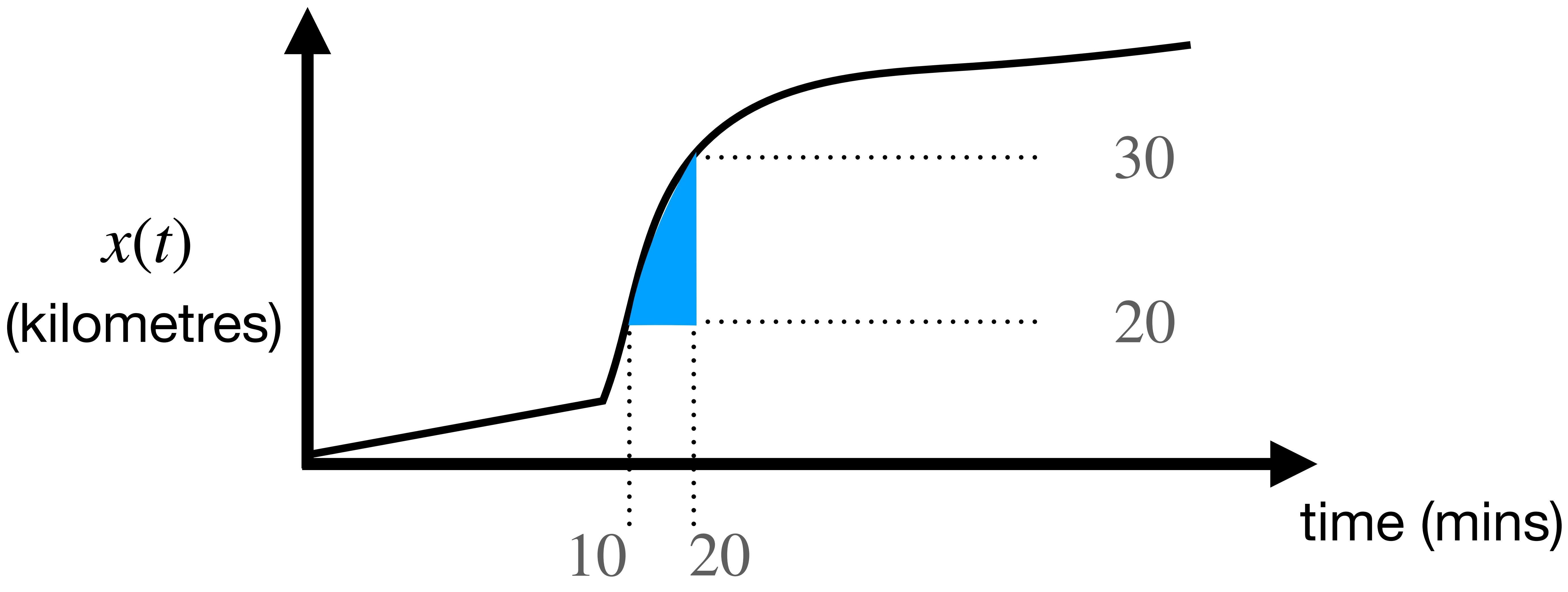



**Differential quantity:**

Velocity requires comparison of position at two points in time!

# Average velocity over a time interval

N.B. $\Delta$ often represents *'change in'*




Units?
Distance
**per** time

Average velocity between times $t$ and $t + \Delta t$ :

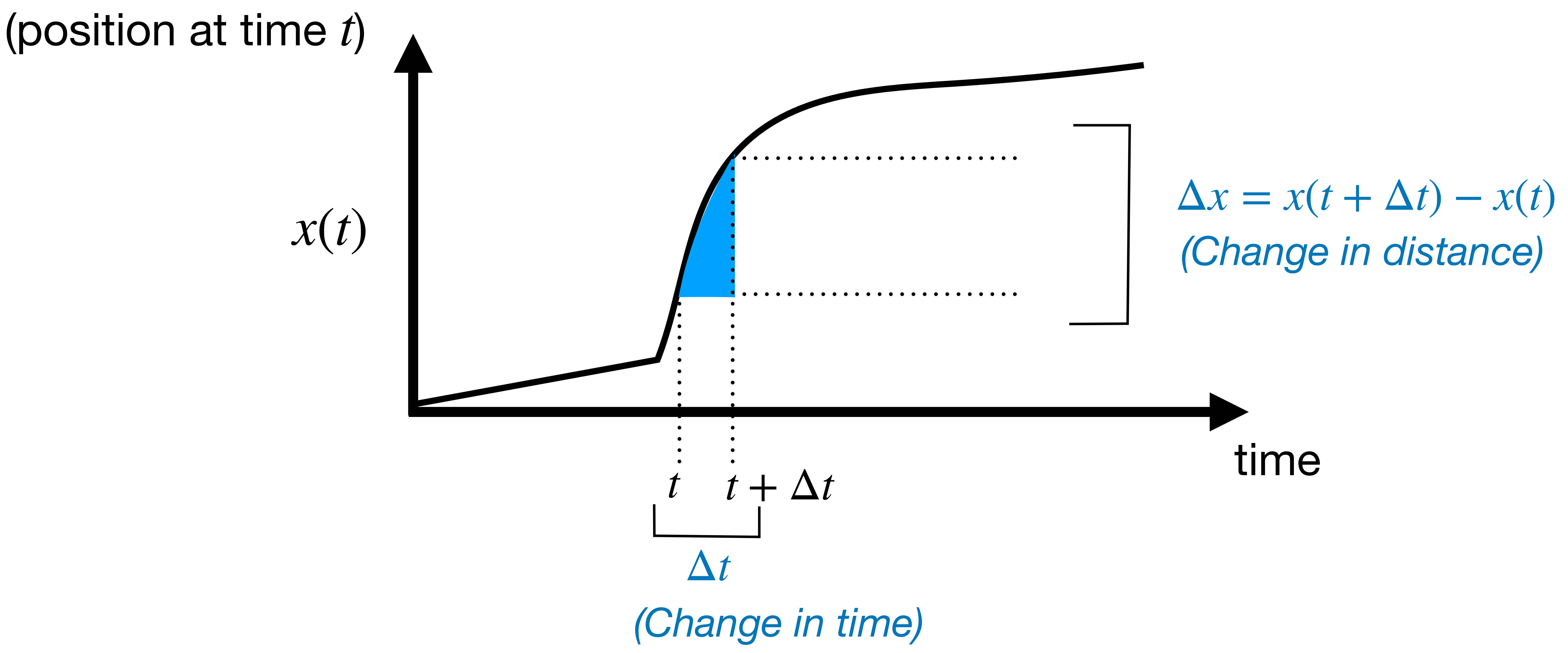
$$\frac{x(20) - x(10)}{20 - 10}$$

"Per"

$$= \frac{x(10) - x(20)}{10 - 20}$$

# Average velocity over a time interval

N.B. $\Delta$ often represents *'change in'*

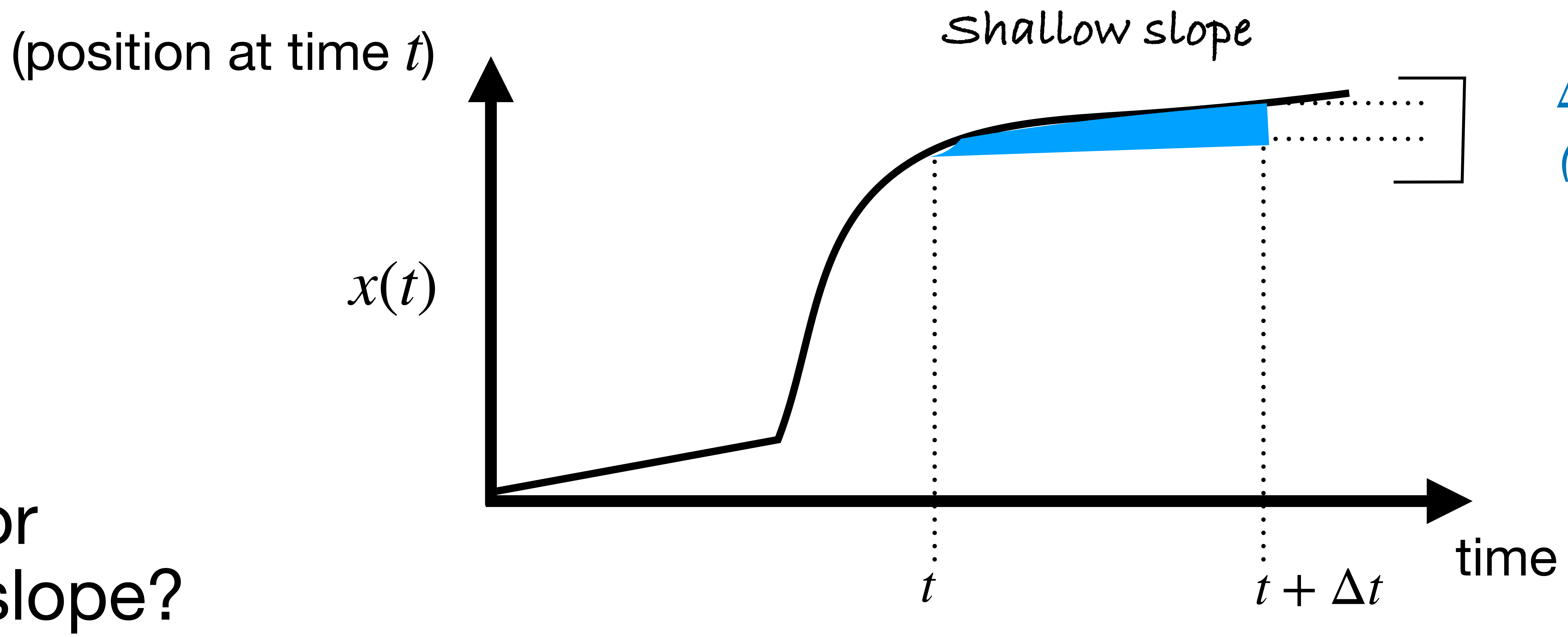



Average velocity between times $t$ and $t + \Delta t$ :

$$\frac{x(t + \Delta t) - x(t)}{\Delta t}$$

# Average velocity is average slope

N.B. $\Delta$ often represents *'change in'*



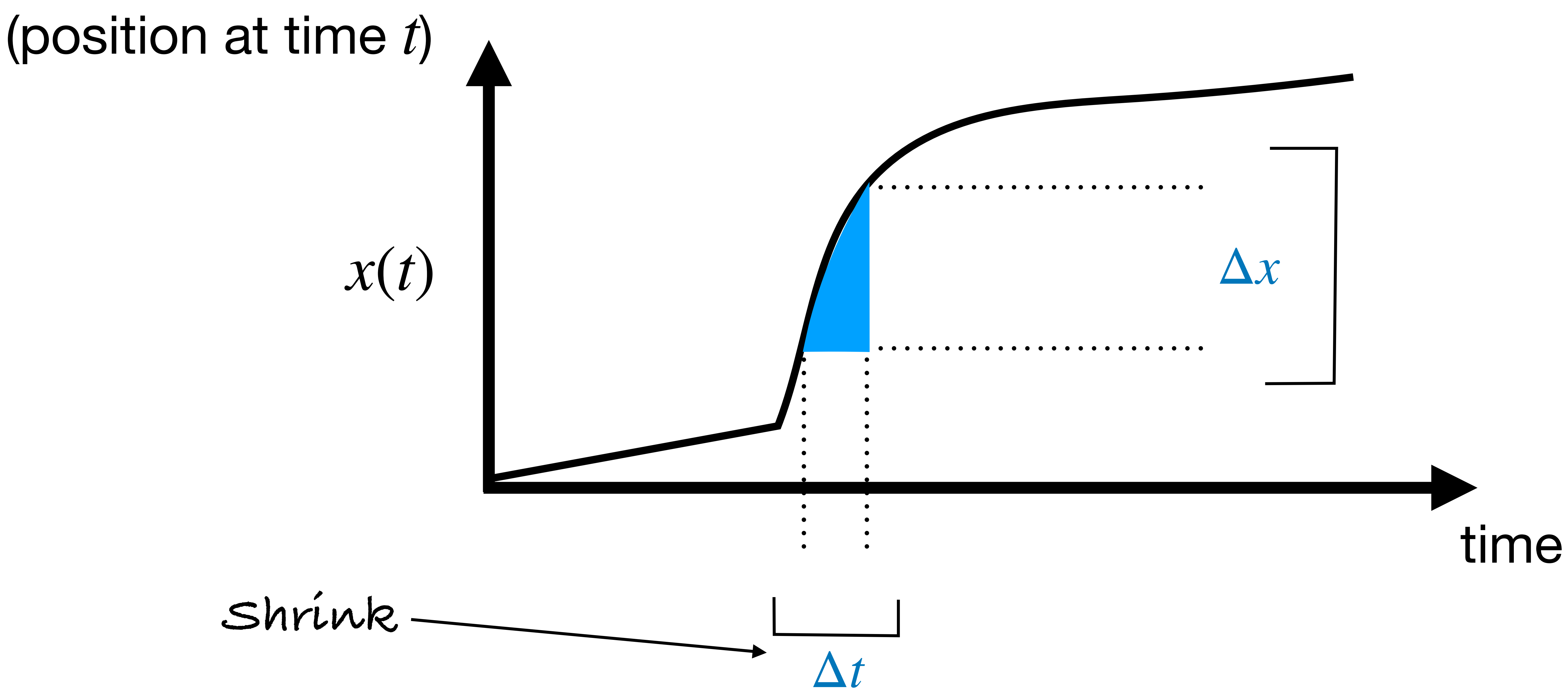

$\Delta x = x(t + \Delta t) - x(t)$

*(Change in distance)*

Velocity for negative slope?

Average velocity between times $t$ and $t + \Delta t$ : $\dfrac{x(t+\Delta t) - x(t)}{\Delta t}$

# Average speed over an infinitesimally small interval




Velocity at time $t$ : $\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$

# Calculating velocity

Suppose $x(t) = t^2$

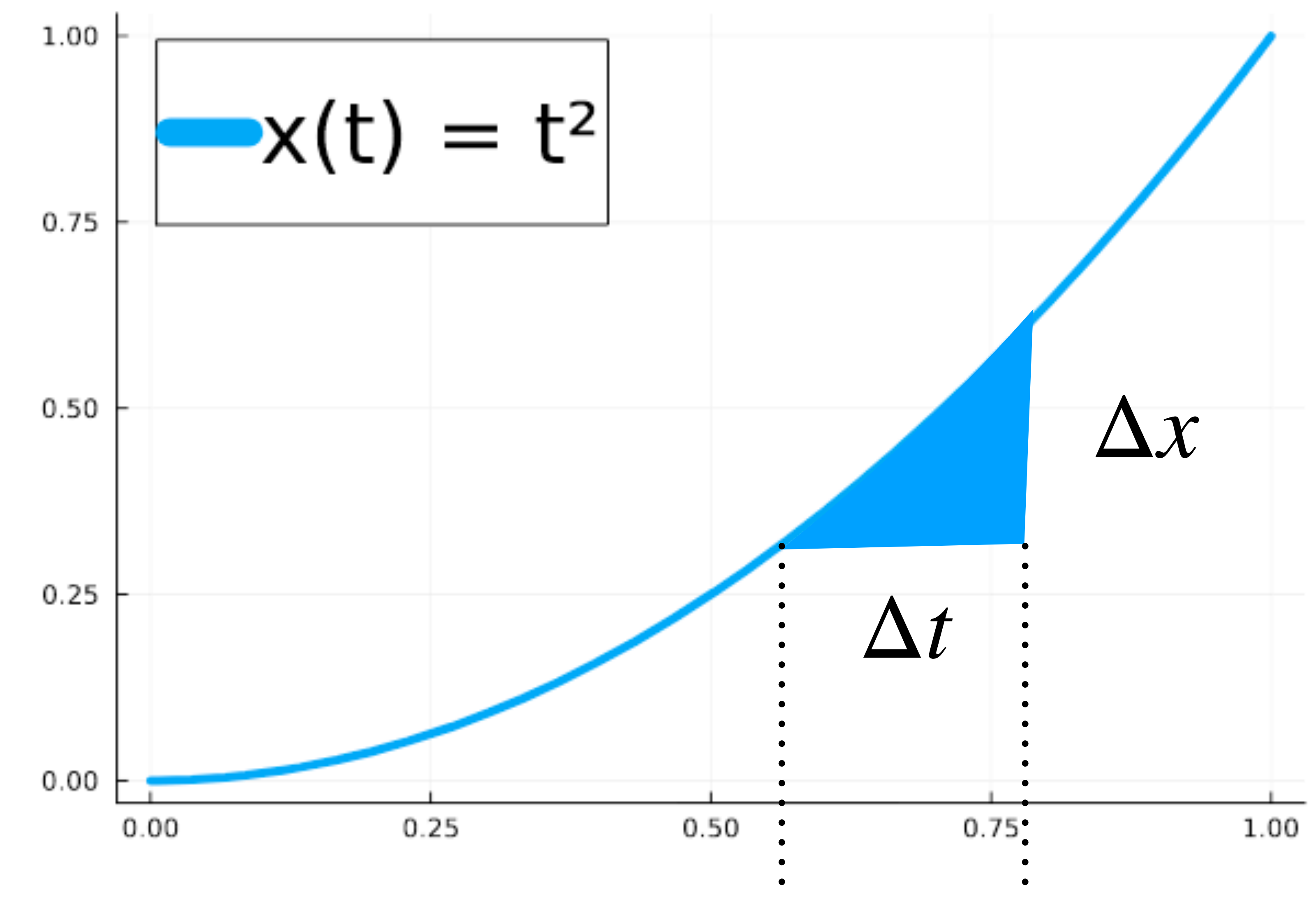
$$\frac{\Delta x}{\Delta t} = \frac{(t+\Delta t)^2 - t^2}{\Delta t}$$

$$= \frac{t^2 + 2t\Delta t + (\Delta t)^2 - t^2}{\Delta t}$$

$$= 2t + \Delta t$$

$$= 2t \qquad \text{(As } \Delta t \to 0\text{)}$$

$x(t)$




time $t$ $t + \Delta t$

Velocity at time $t$ : $\displaystyle\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$

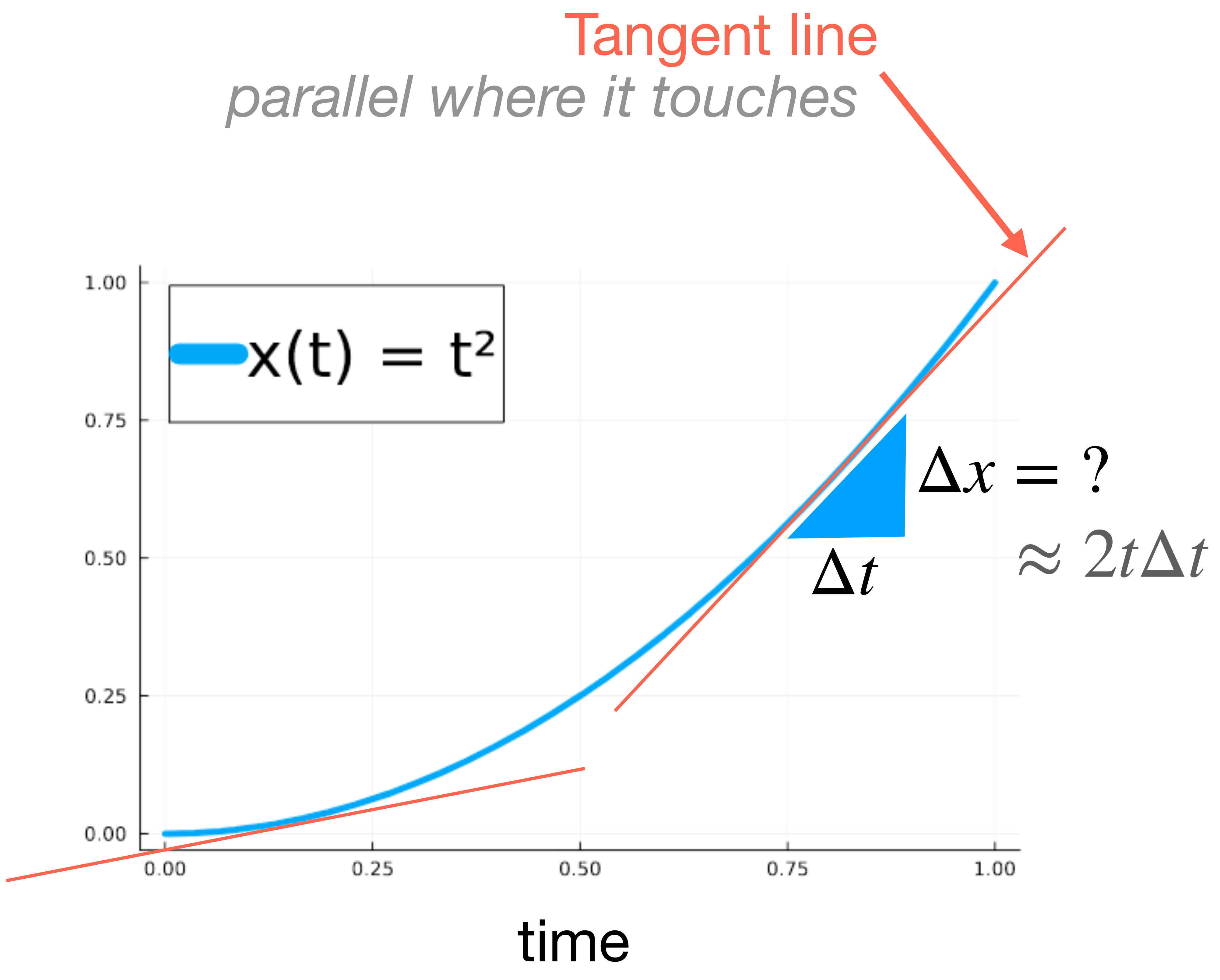
**Exercise**

Did we assume $\Delta t$ is positive?

# Tangents are slopes that kiss a curve

Velocity at time $t$ : $\displaystyle\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$

*(i.e. as triangle width goes to zero)*




Velocity $(2t)$ at time $t$ is the steepness of the tangent

# Velocity is a differential quantity

N.B. $\mathrm{d}$ is \mathrm{d} in LaTeX

*"Velocity is $\dfrac{\Delta x}{\Delta t}$ for an infinitesimally small change in time"*

Time: $t$

Change in time: $\Delta t$ ← Same notation for any variable!

Infinitesimally small change in time: $\mathrm{d}t$ ← Differential/infinitesimal quantity

# Velocity is a differential quantity

Infinitesimal change in position

*"Velocity is* $\frac{\mathrm{d}x(t)}{\mathrm{d}t}$*"*

Infinitesimal change in time

*"Velocity is the rate of change of position with respect to time"*

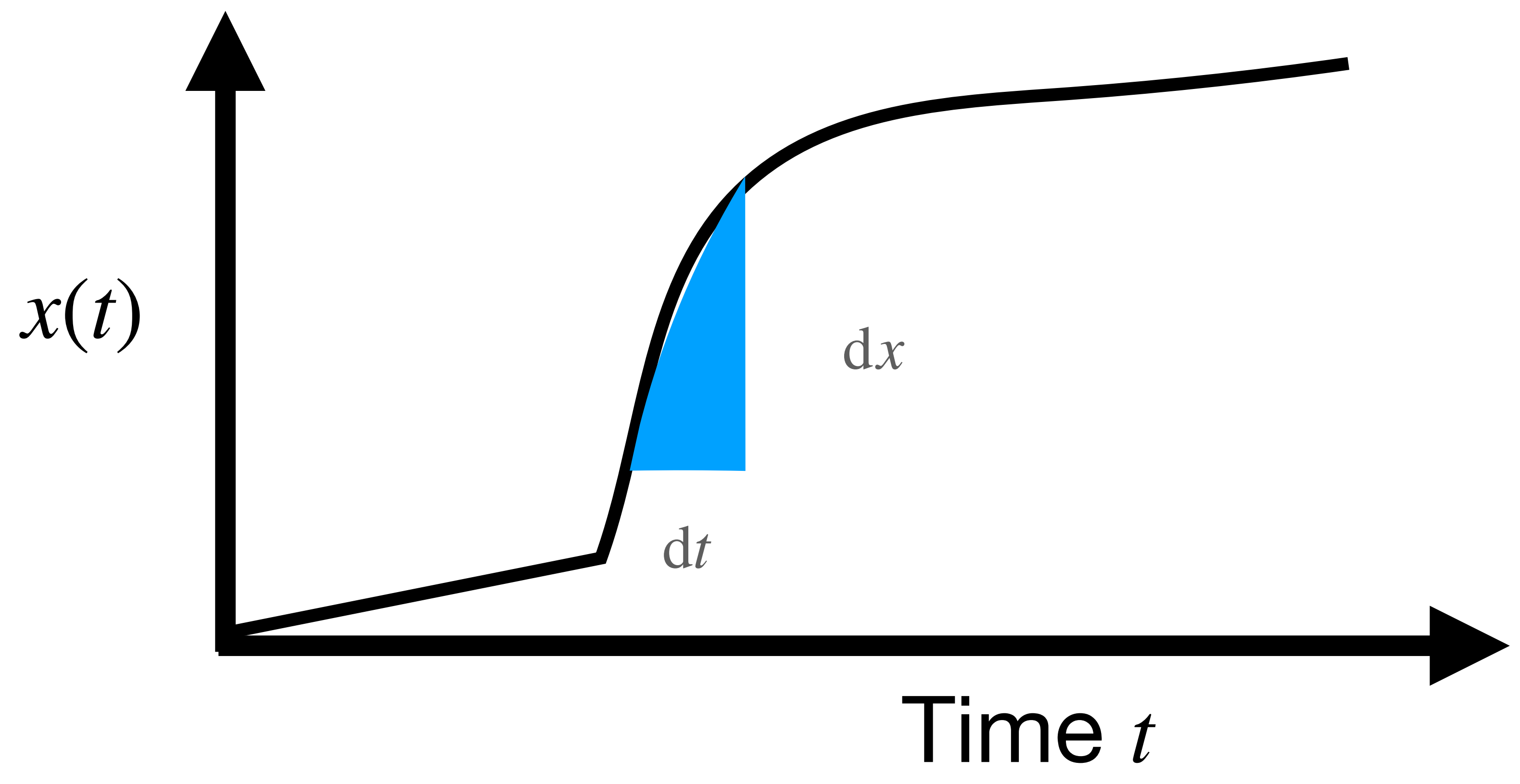
*"Velocity is the derivative of position with respect to time"*

| | |
|---|---|
| Time: | $t$ |
| Change in time: | $\Delta t$ |
| Infinitesimally small change in time: | $\mathrm{d}t$ |

Same notation for any variable!

Differential quantity

# Velocity is a differential quantity



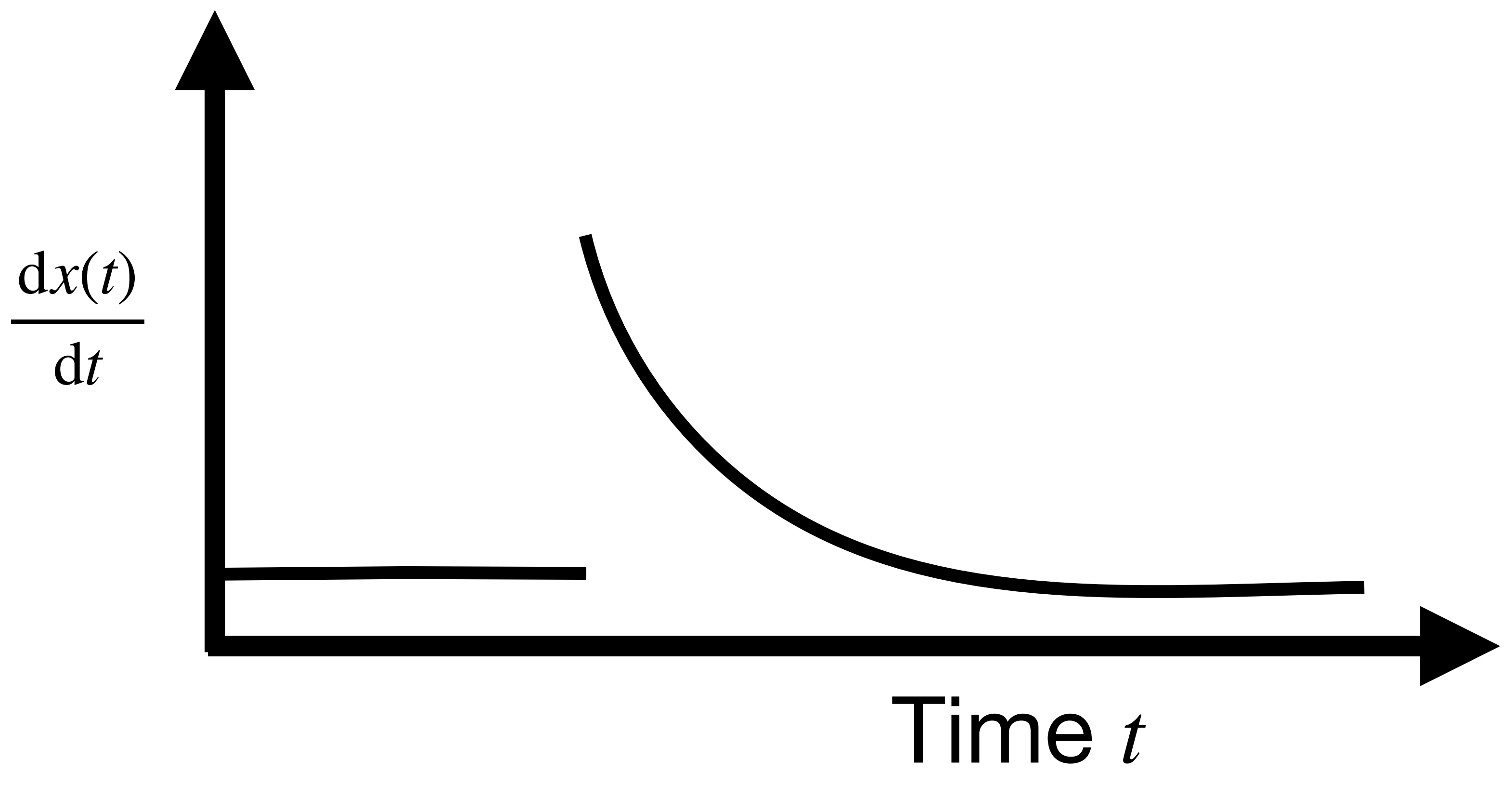

$$\mathrm{d}x = \text{velocity(t)} \times \mathrm{d}t$$

$$\mathrm{d}x(t) = \text{velocity(t)} \times \mathrm{d}t$$

It's a function, depends on $t$!

$$\text{velocity(t)} = \frac{\mathrm{d}x(t)}{\mathrm{d}t}$$

Rough sketch?



# How do I read this?

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}$$

**Tip: read from the bottom**

Infinitesimal change in denominator =>
How much change in numerator?

$$\Downarrow$$

*If $t$ changed infinitesimally, how much would $x(t)$ (infinitesimally) change?*

# Notational chaos

Position: $x(t)$

$$x'(t) = \frac{\mathrm{d}x(t)}{\mathrm{d}t}$$

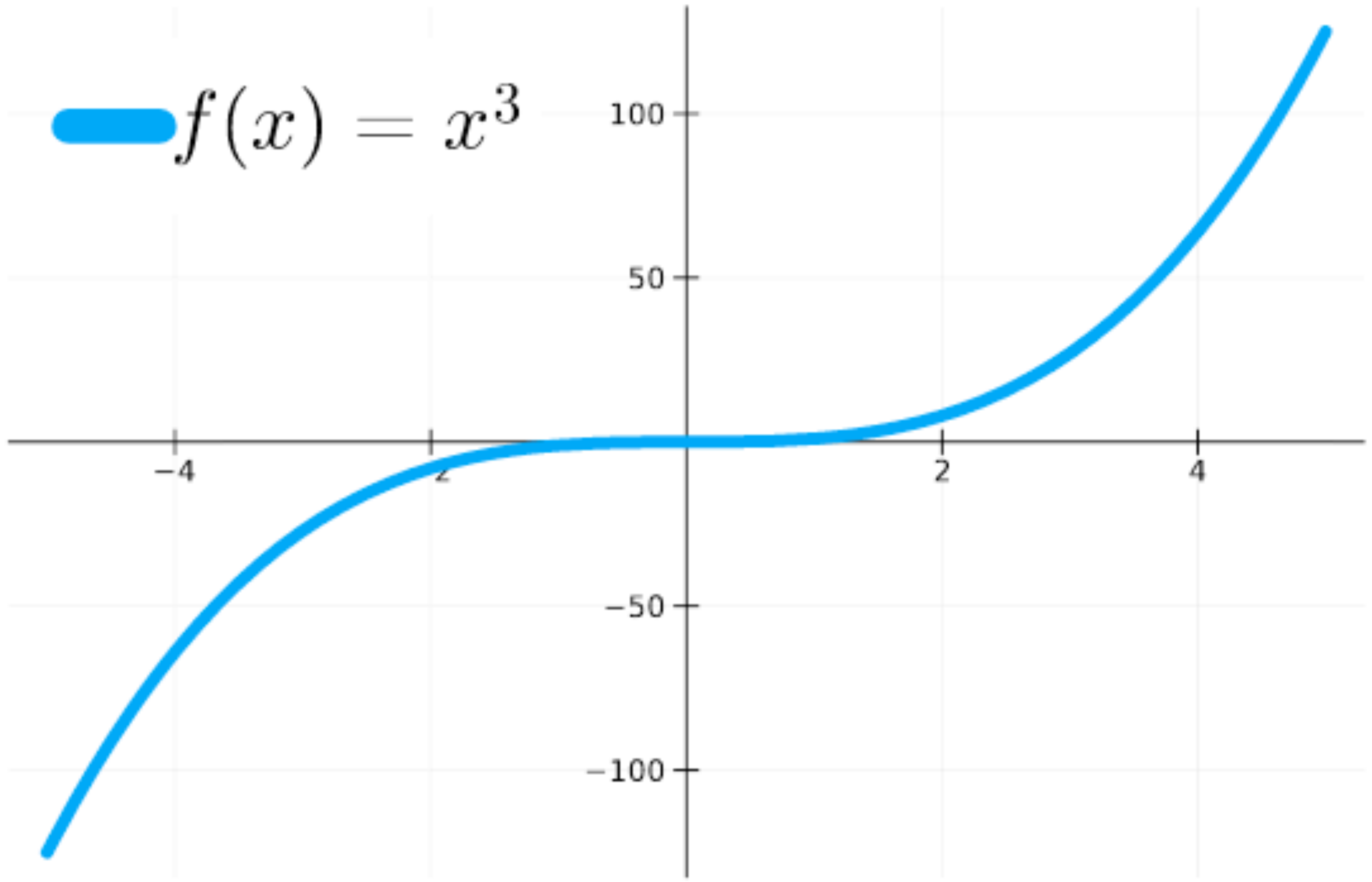
**Other notations for the derivative**
**evaluated at t**

$$\dot{x}(t) \qquad \frac{\mathrm{d}x}{\mathrm{d}t}(t) \qquad \frac{\mathrm{d}x(t)}{\mathrm{d}t}$$

**Other notations for the derivative**
**as a function**

$$\frac{\mathrm{d}x}{\mathrm{d}t} \qquad \dot{x} \qquad x'$$

# Lots of* functions have derivatives




*Mathematical notation for derivative?* $f'(x)$

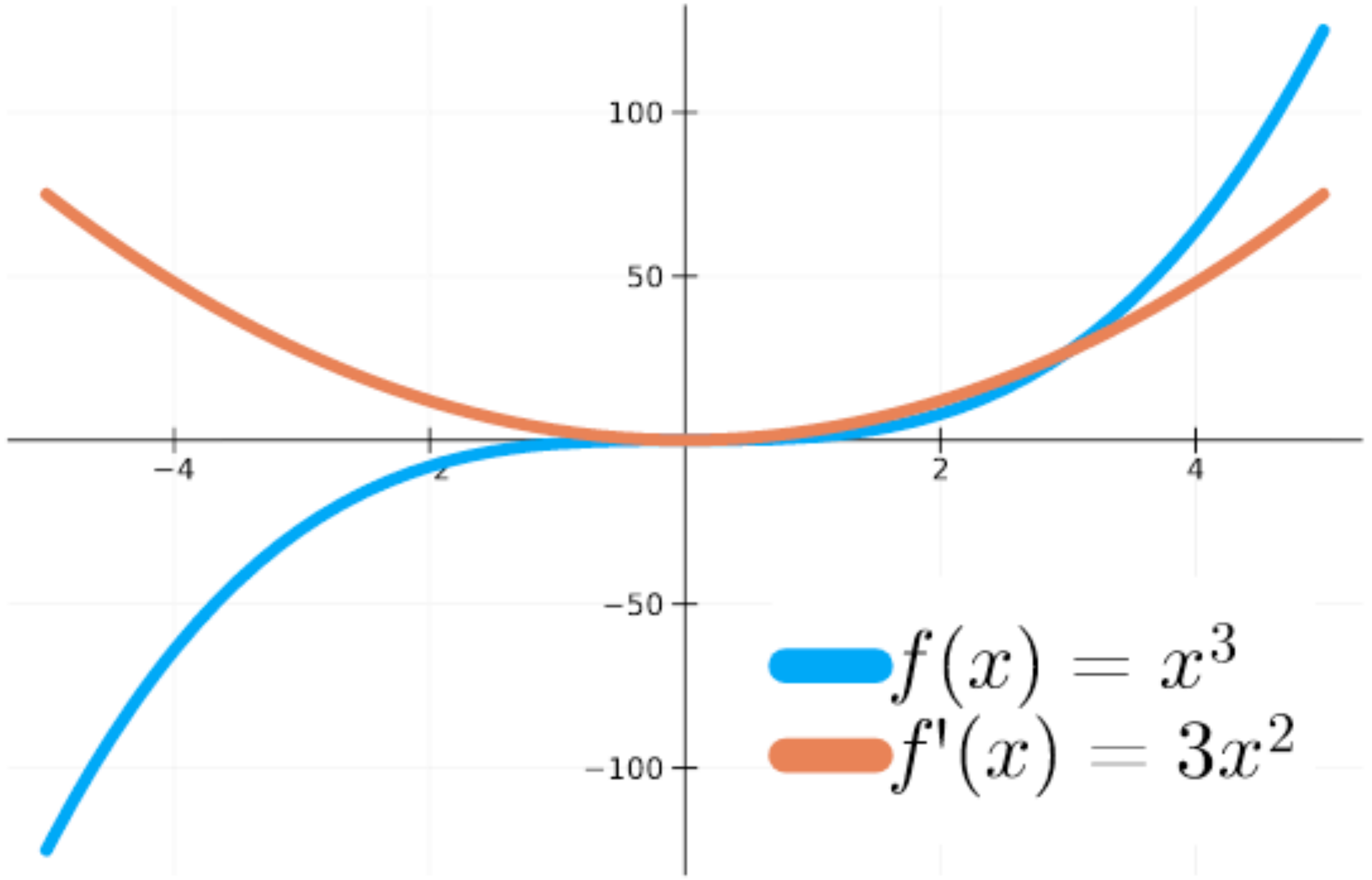
*In/dependent variable?*

$x$ Unlabelled, e.g. $y$

*Sketch the derivative from intuition*

***Terms and conditions apply**

# Lots of functions have derivatives




At each value $x$:

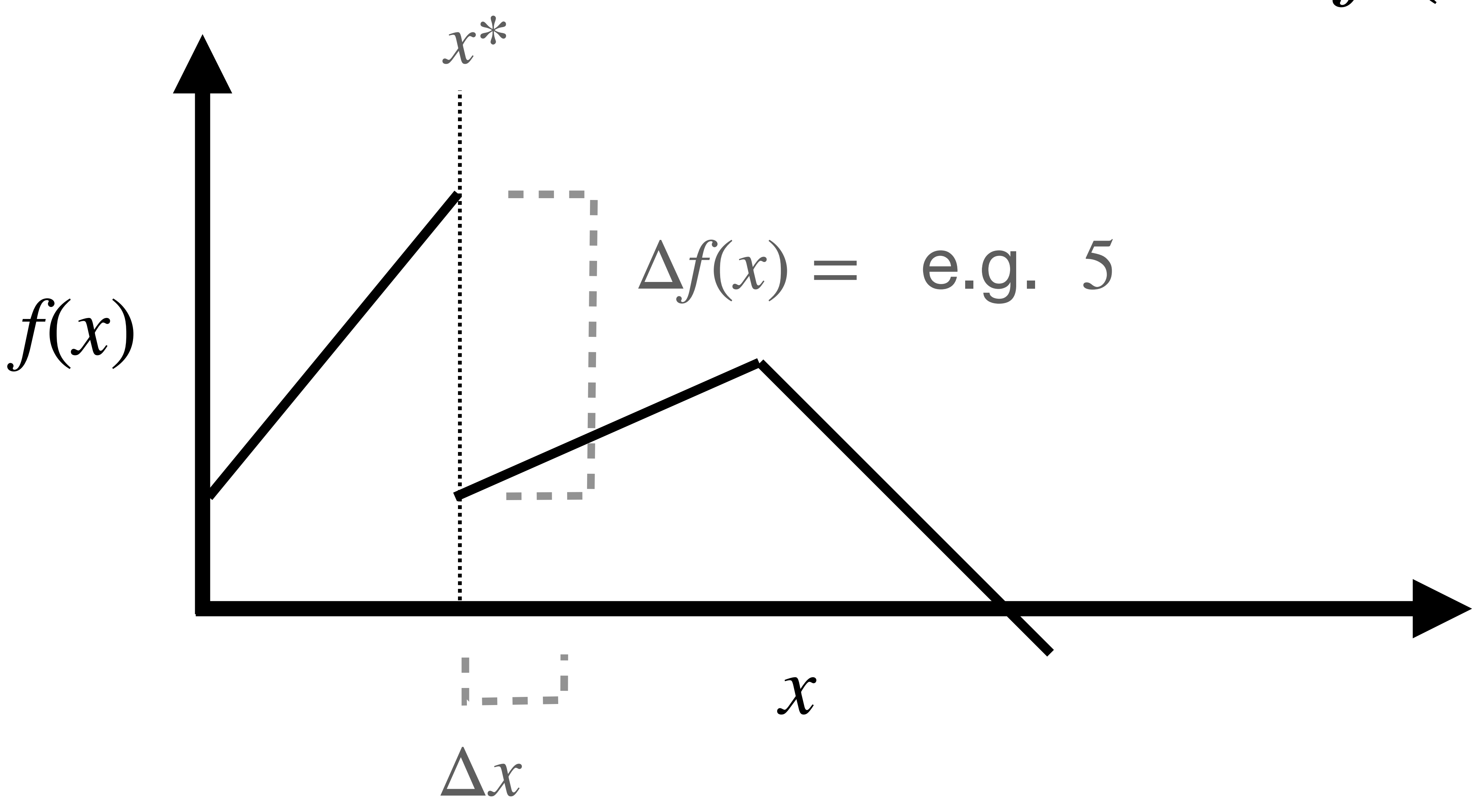
$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

Easier (less proper) notation:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x}$$

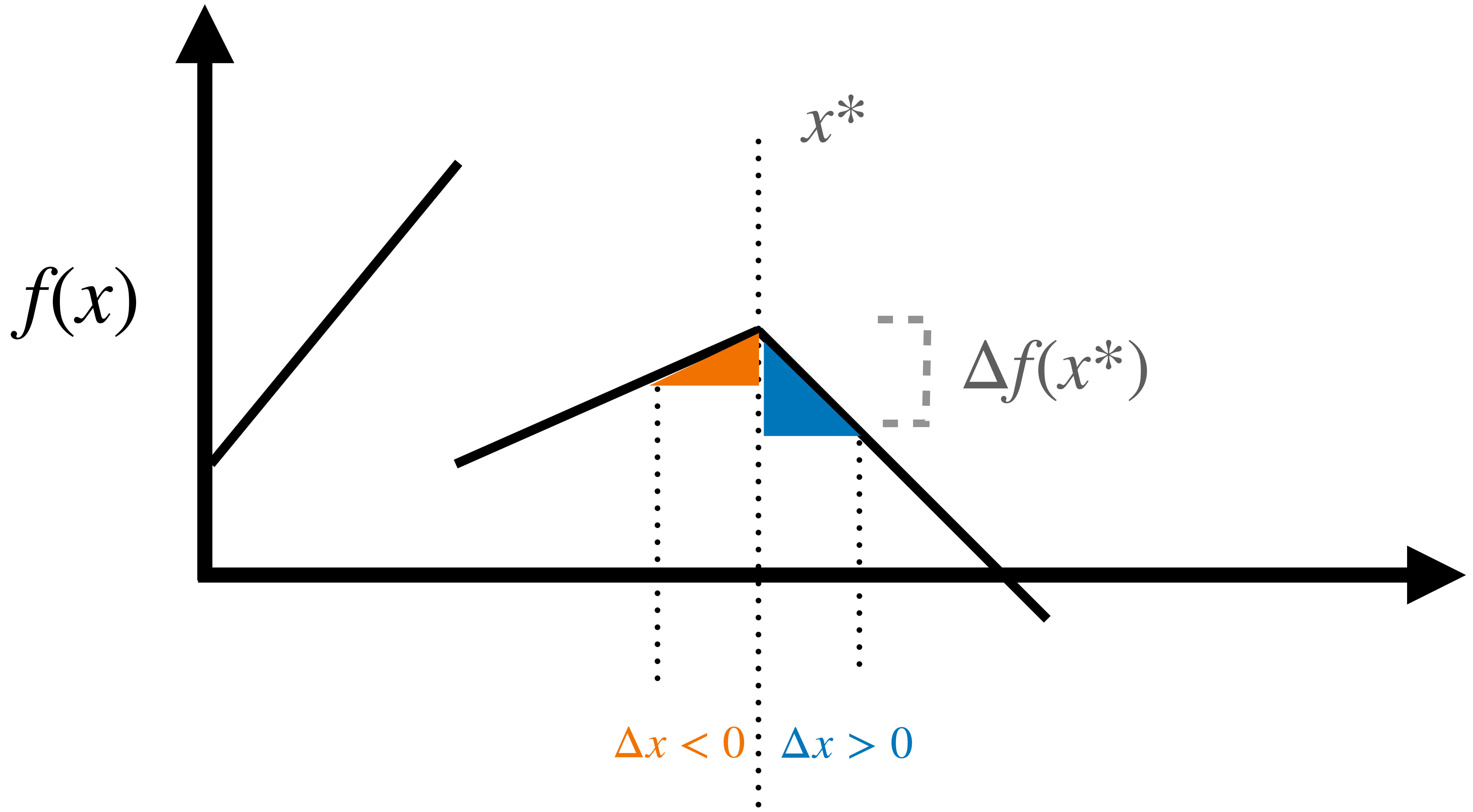
# Derivatives are undefined at corners and jumps

$f'(x)$ is undefined at $x^*$ only




$$\lim_{\Delta x \to 0} \frac{\Delta f(x^*)}{\Delta x} ? \quad = \lim_{\Delta x \to 0} \frac{5}{\Delta x}$$

# Derivatives are undefined at sharp corners and jumps




$$\lim_{\Delta x \to 0} \frac{\Delta f(x^*)}{\Delta x}:$$

$\Delta x$ *can be negative or positive. Limit must be the same or doesn't exist!*

$$\lim_{\Delta x \to 0, \Delta x < 0} \frac{\Delta f(x)\quad <0}{\Delta x\quad <0}$$

"Limit from below"

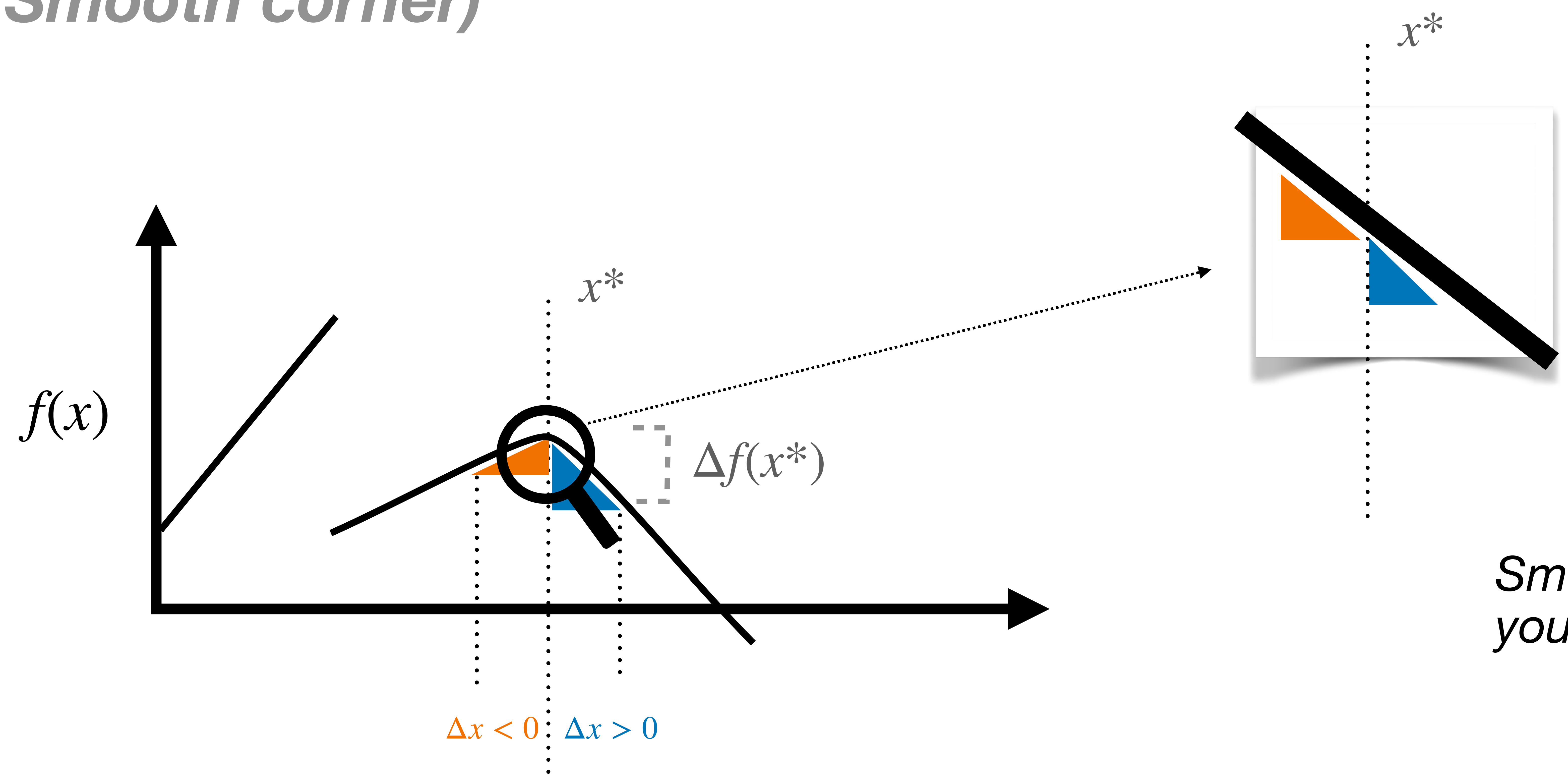
$$\neq$$

$$\lim_{\Delta x \to 0, \Delta x > 0} \frac{\Delta f(x)\quad <0}{\Delta x\quad >0}$$

"Limit from above"

# Why isn’t this undefined?

## *(Smooth corner)*

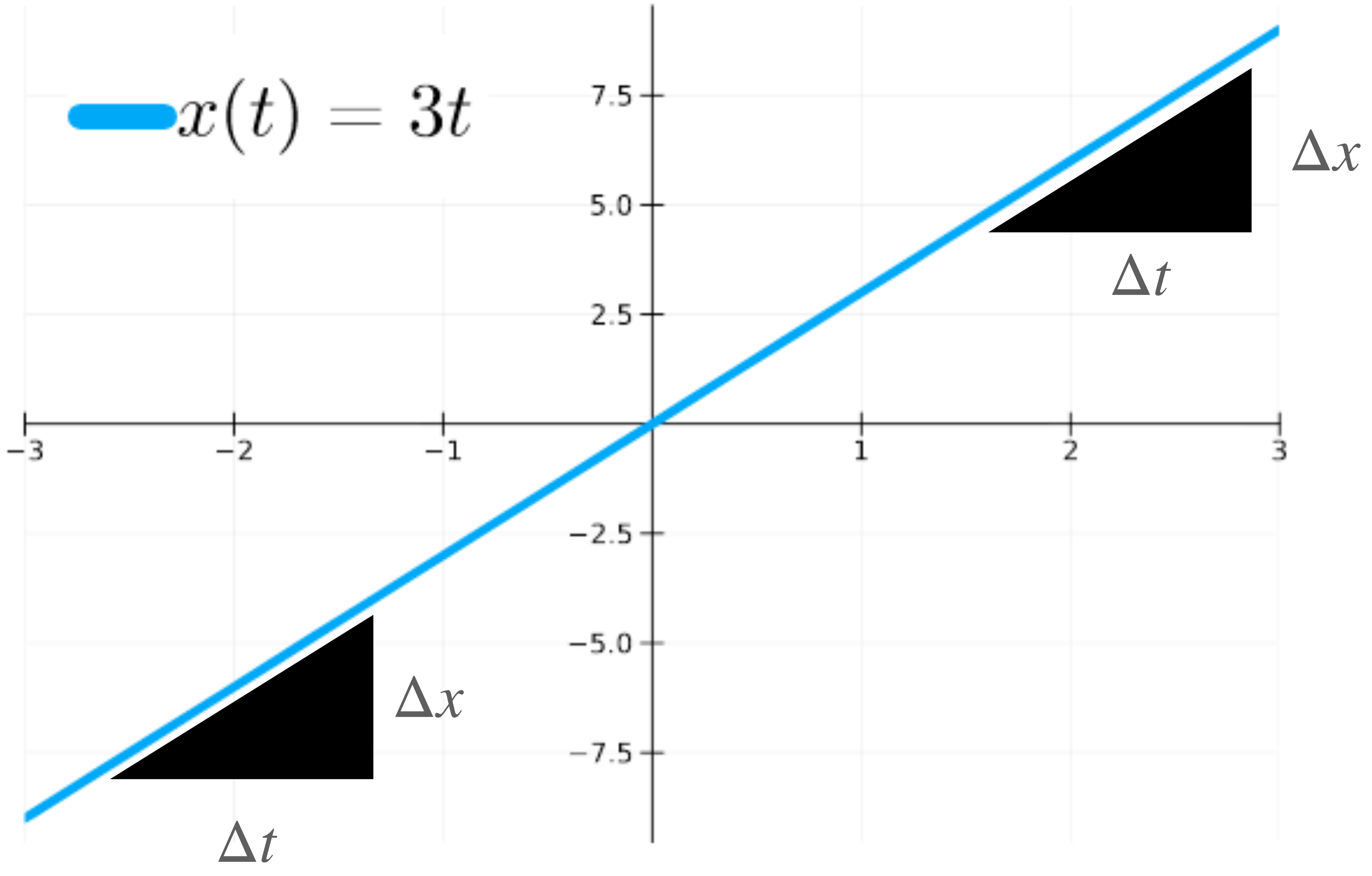



*Smooth corners are straight if you zoom in enough!!*

$$\lim_{\Delta x \to 0} \frac{\Delta f(x^*)}{\Delta x} :$$ *Same whether* $\Delta x < 0$ *or* $> 0$*!*

# Derivative?

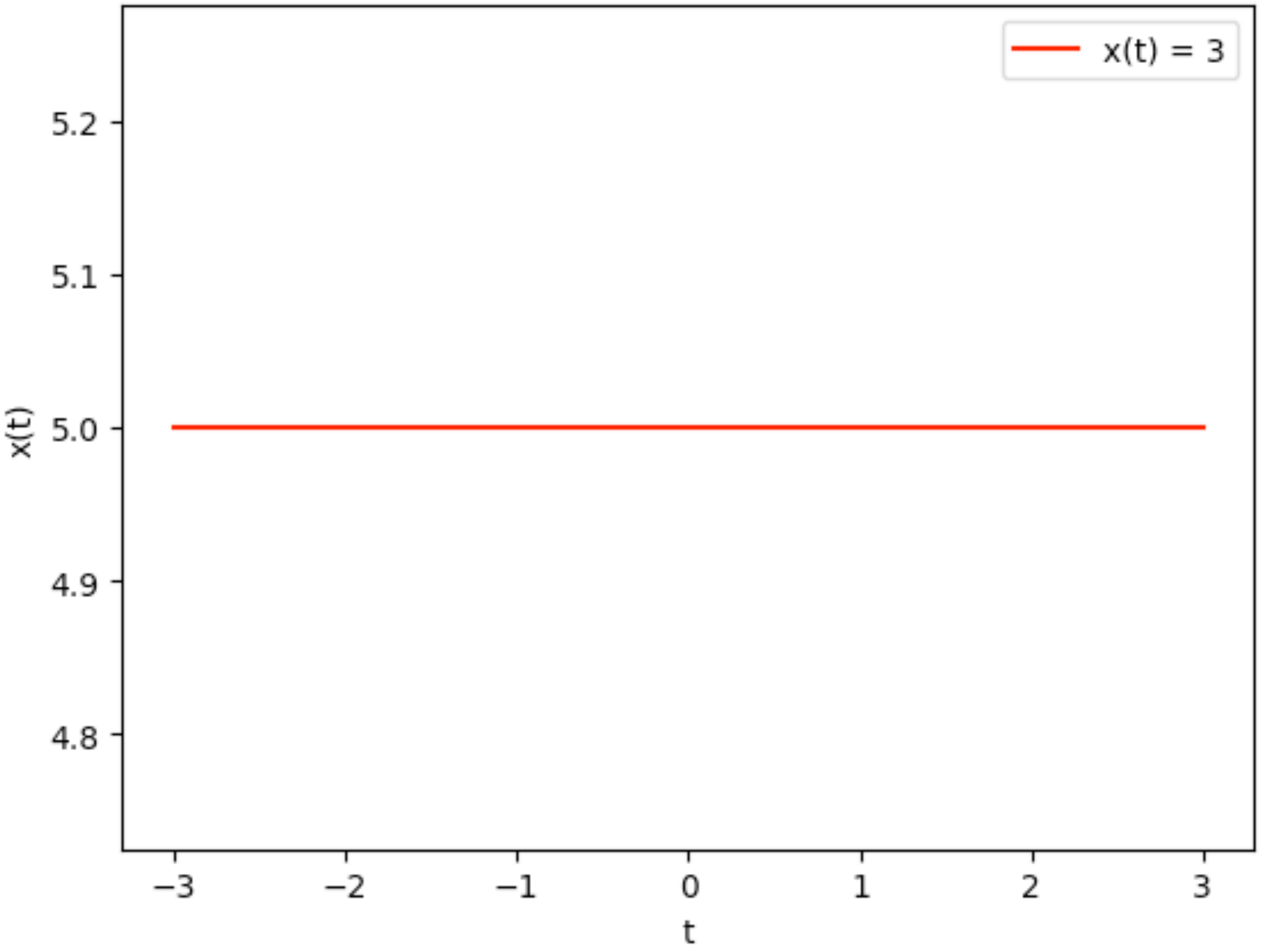
*Linear functions have*
*constant derivatives*



# Derivative?

*Constant functions have zero derivatives*

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



# Machine learning motivation

Neural network $f$

Data batch $x$
(matrix)

Weights $w$
(change to train)

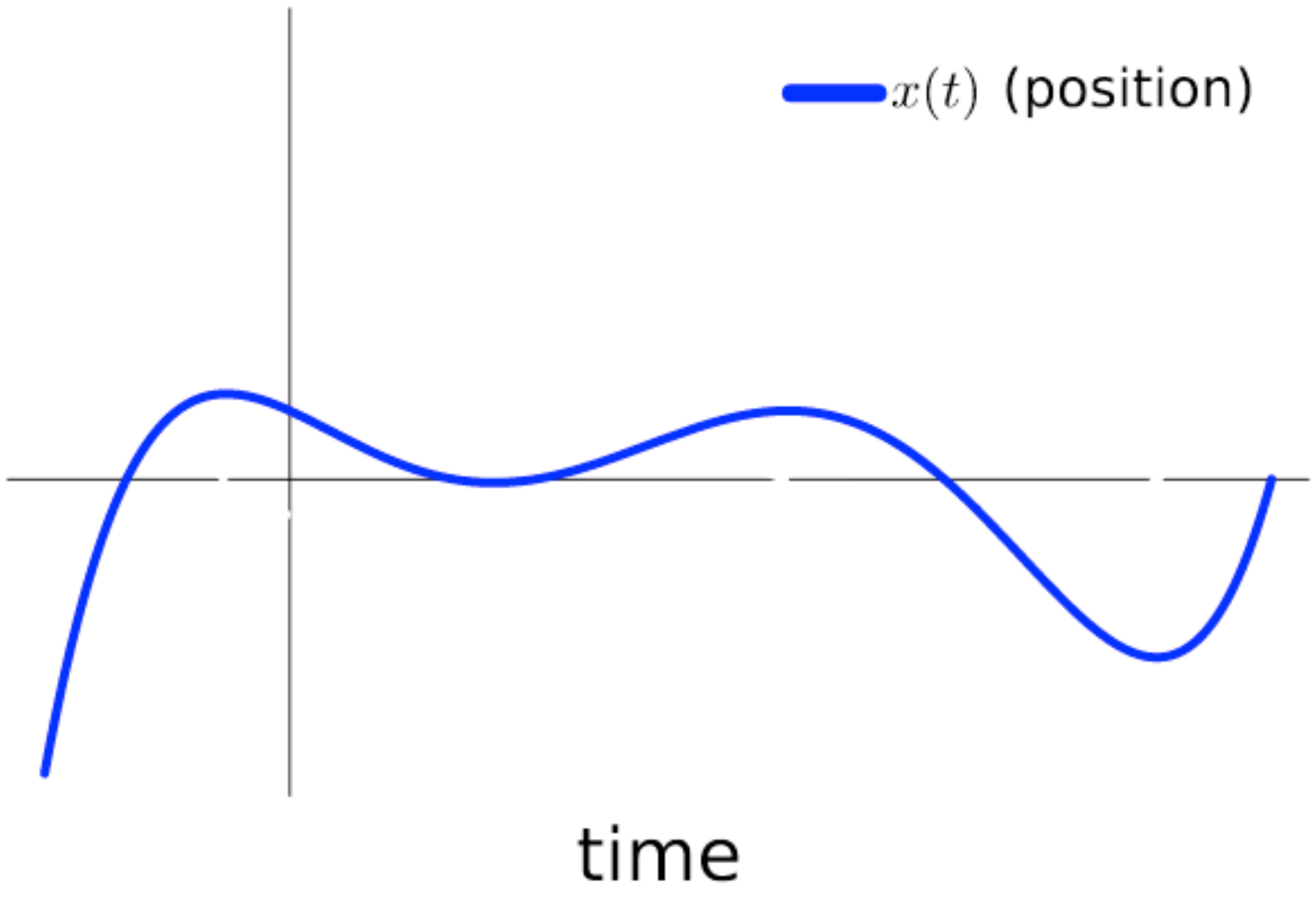
$$f(x, w)$$

*Calculate (automatic differentiation/backpropagation)*

$$\frac{\mathrm{d}f}{\mathrm{d}w}(x, w)$$

*Use derivative to change weights*

# Acceleration



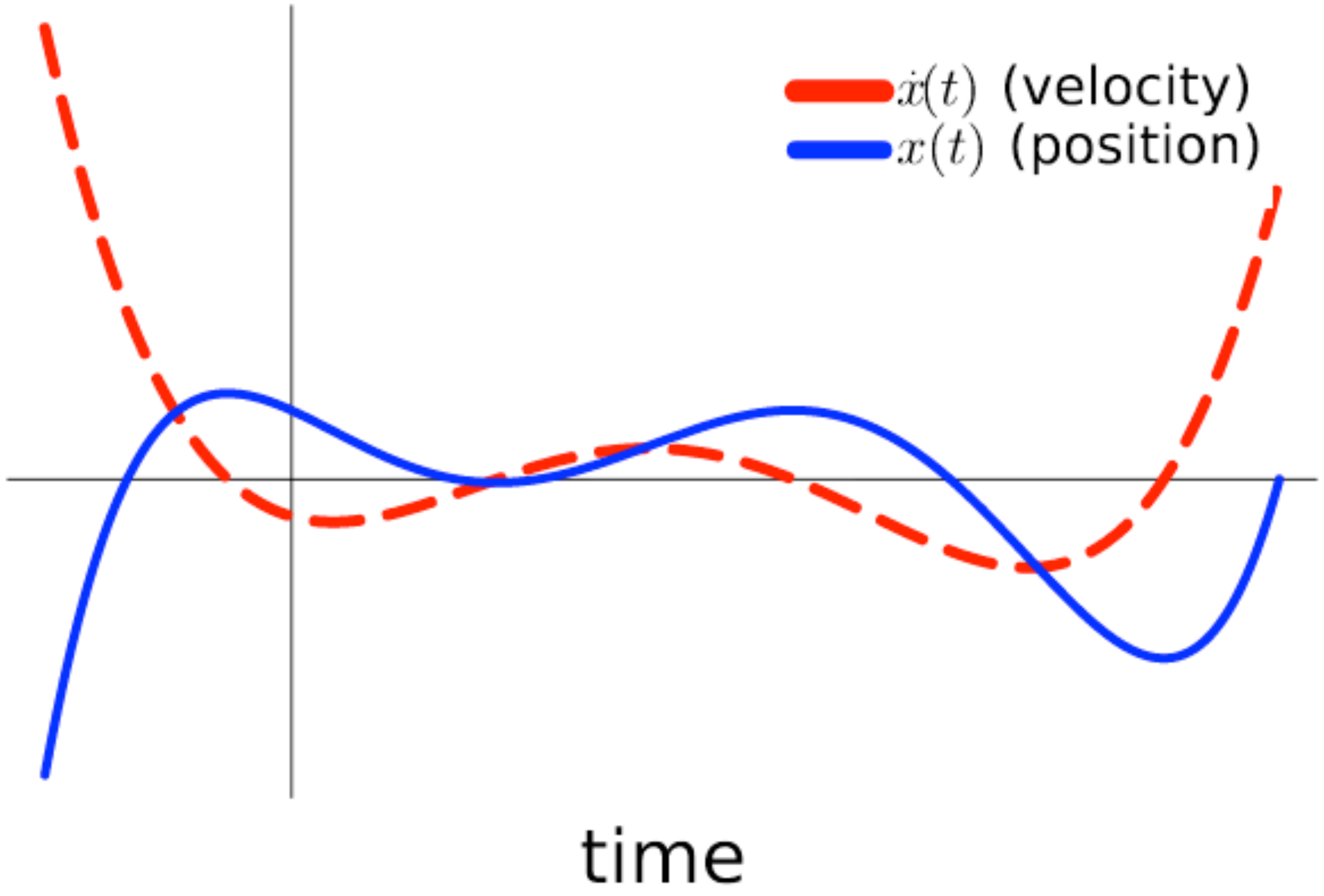

Sketch derivative (velocity)?

# Acceleration

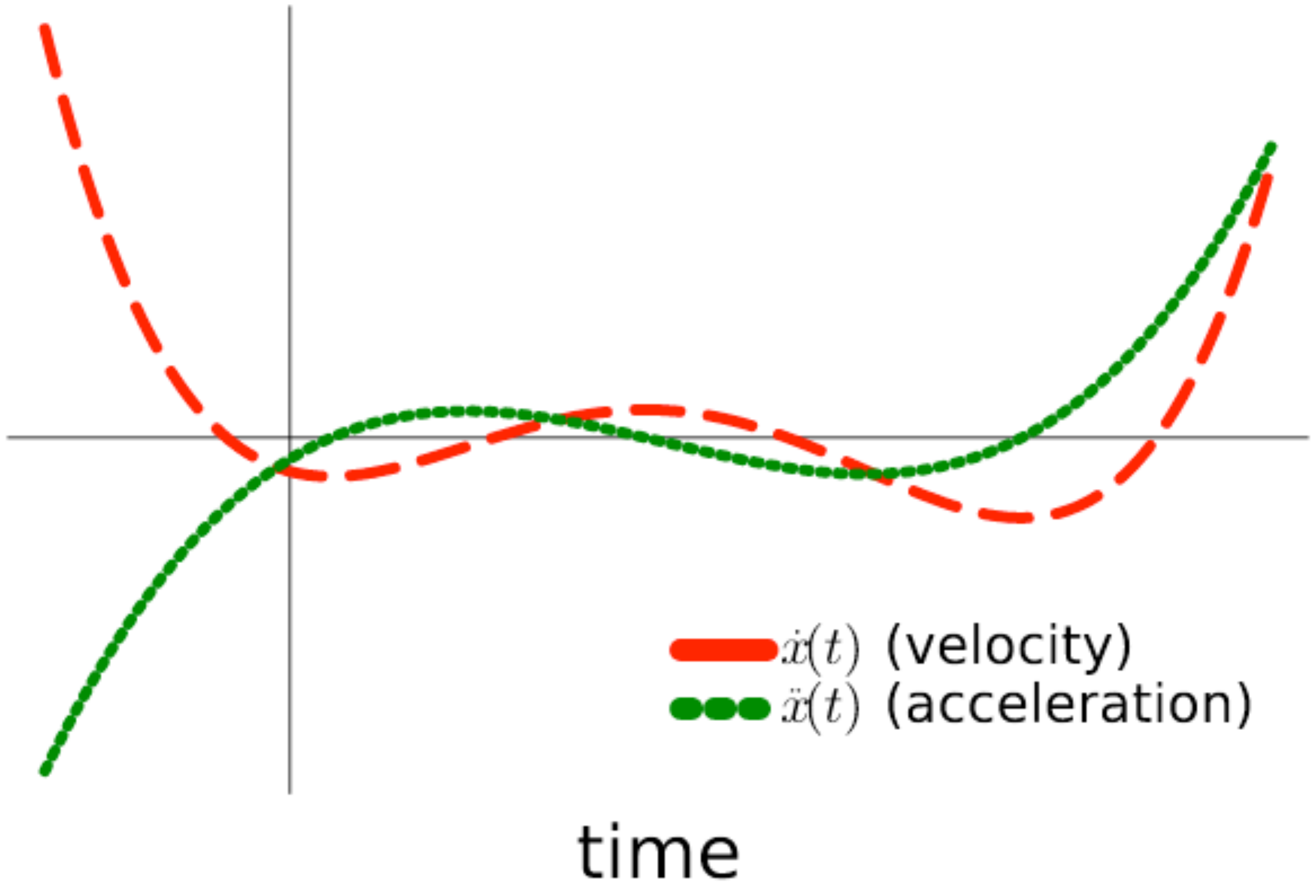
**Velocity**: rate of change of position … $x'(t)$

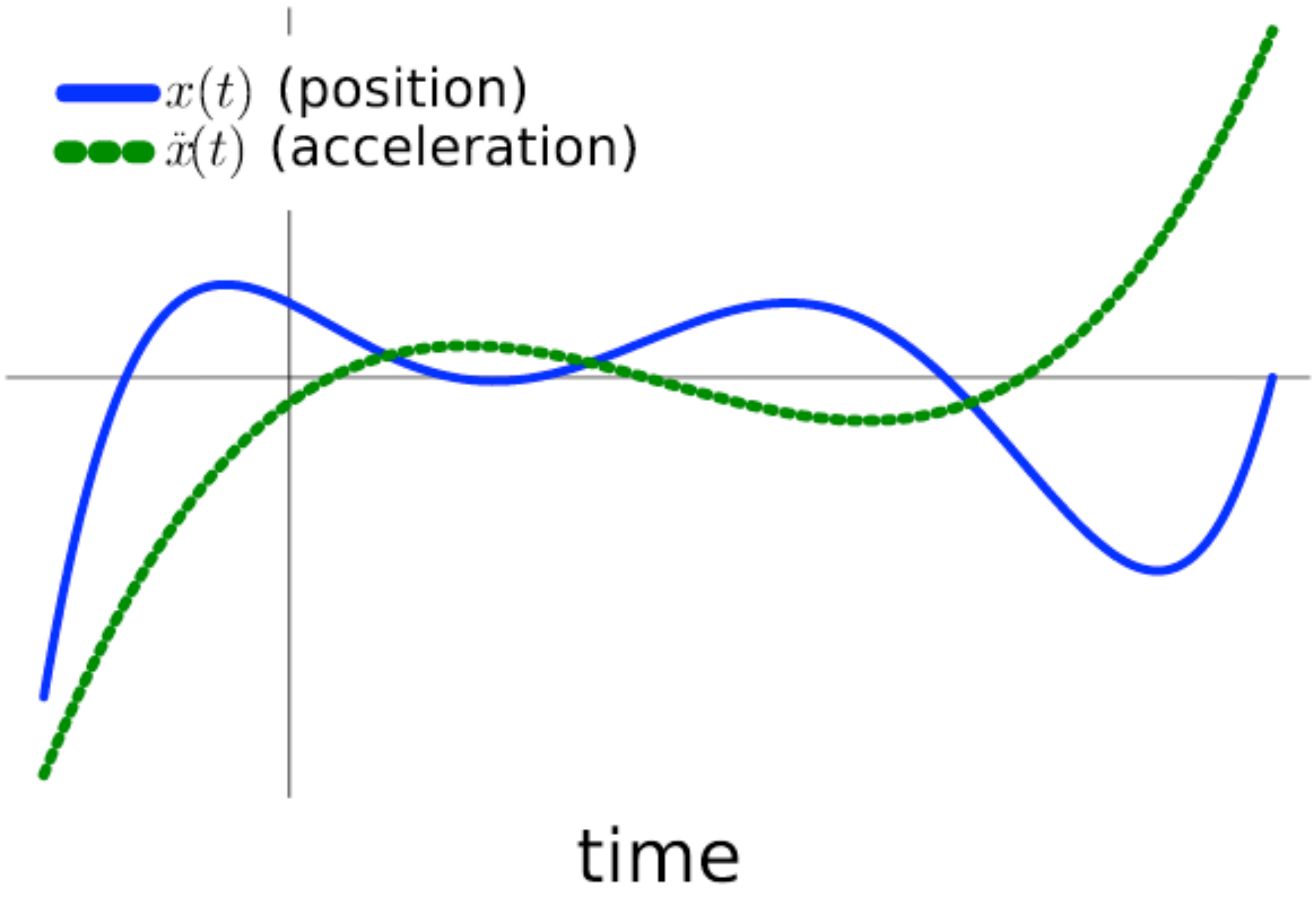
**Acceleration**: rate of change of velocity … $x''(t)$

“Accerelation is to velocity as velocity is to position”



# Acceleration




Geometrical intuition:

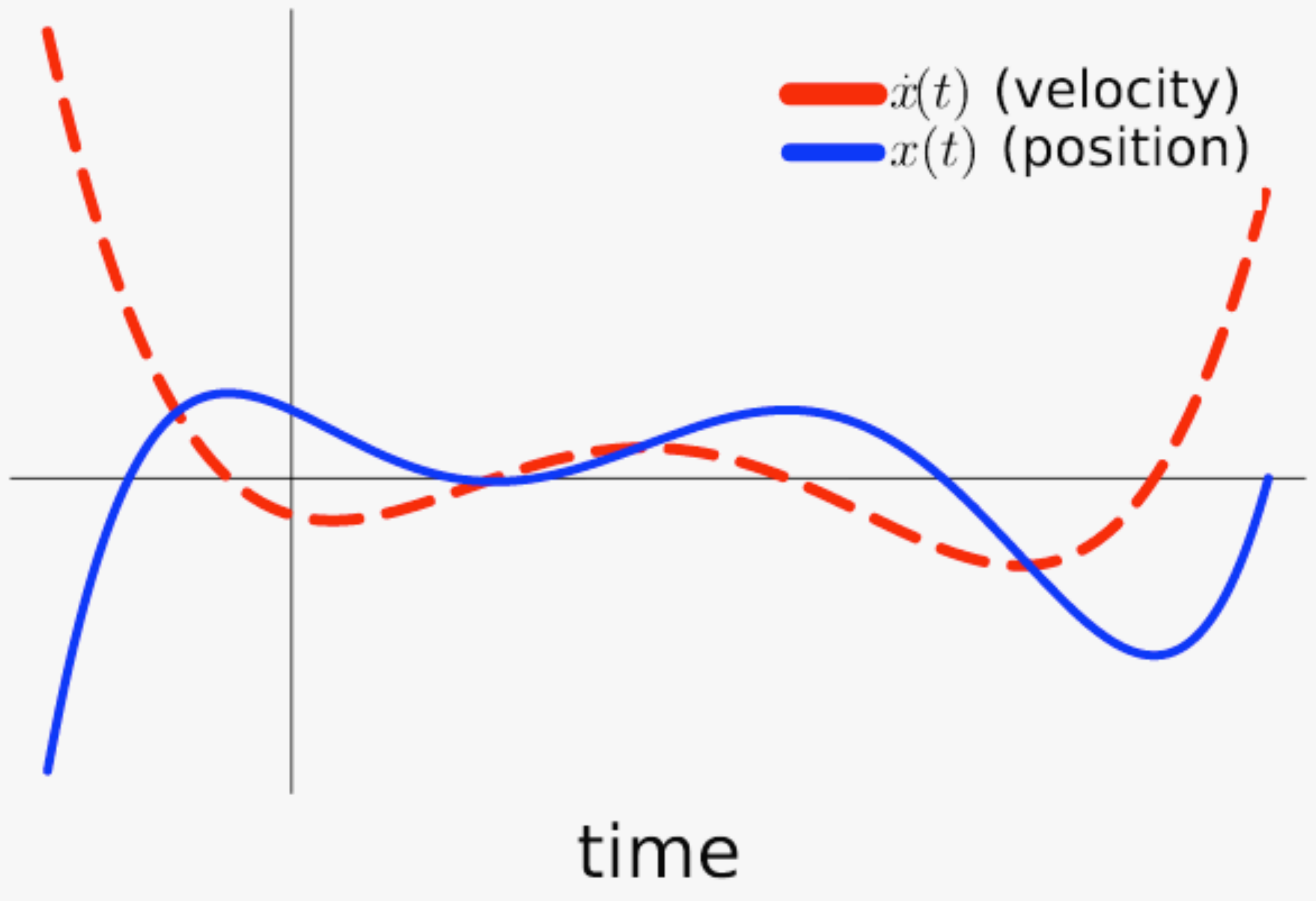
Derivative of velocity (= steepness)

Double derivative of position (= curviness)

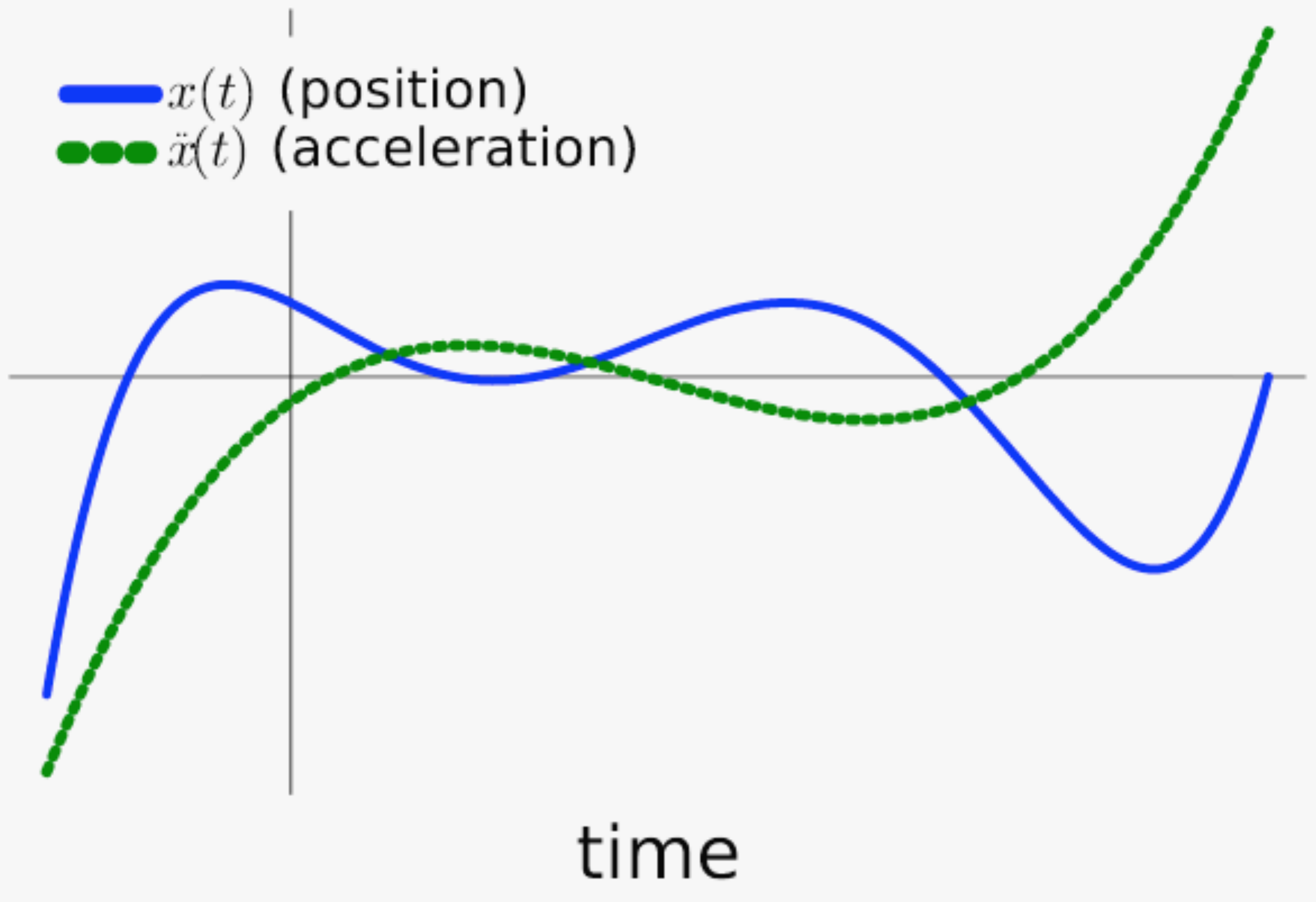
# Take home message

= steepness of steepness

First derivative $x'(t)$ is steepness of graph at $t$




Second derivative $x''(t)$ is curviness of graph at $t$



# Notation for higher-order derivatives

**Acceleration** is the second derivative of position $x(t)$

(**Jerk** is the third derivative)

(Then snap crackle and pop)

nth derivative?

(Fourth order) $\longrightarrow$

**Common notations:**

$$\frac{\mathrm{d}^n x(t)}{\mathrm{d}t^n} \qquad \frac{\mathrm{d}^n x}{\mathrm{d}t^n}(t)$$

$$x''''(t) \qquad \ddot{\ddot{x}}(t)$$

# Independent and dependent variables

Denominator and numerator

**Position function**

$$x(t)$$

$$\text{Time} \longrightarrow \text{Position}$$

Input of function:
*"independent variable"*

Output of function:
*"dependent variable"*

**Velocity function**

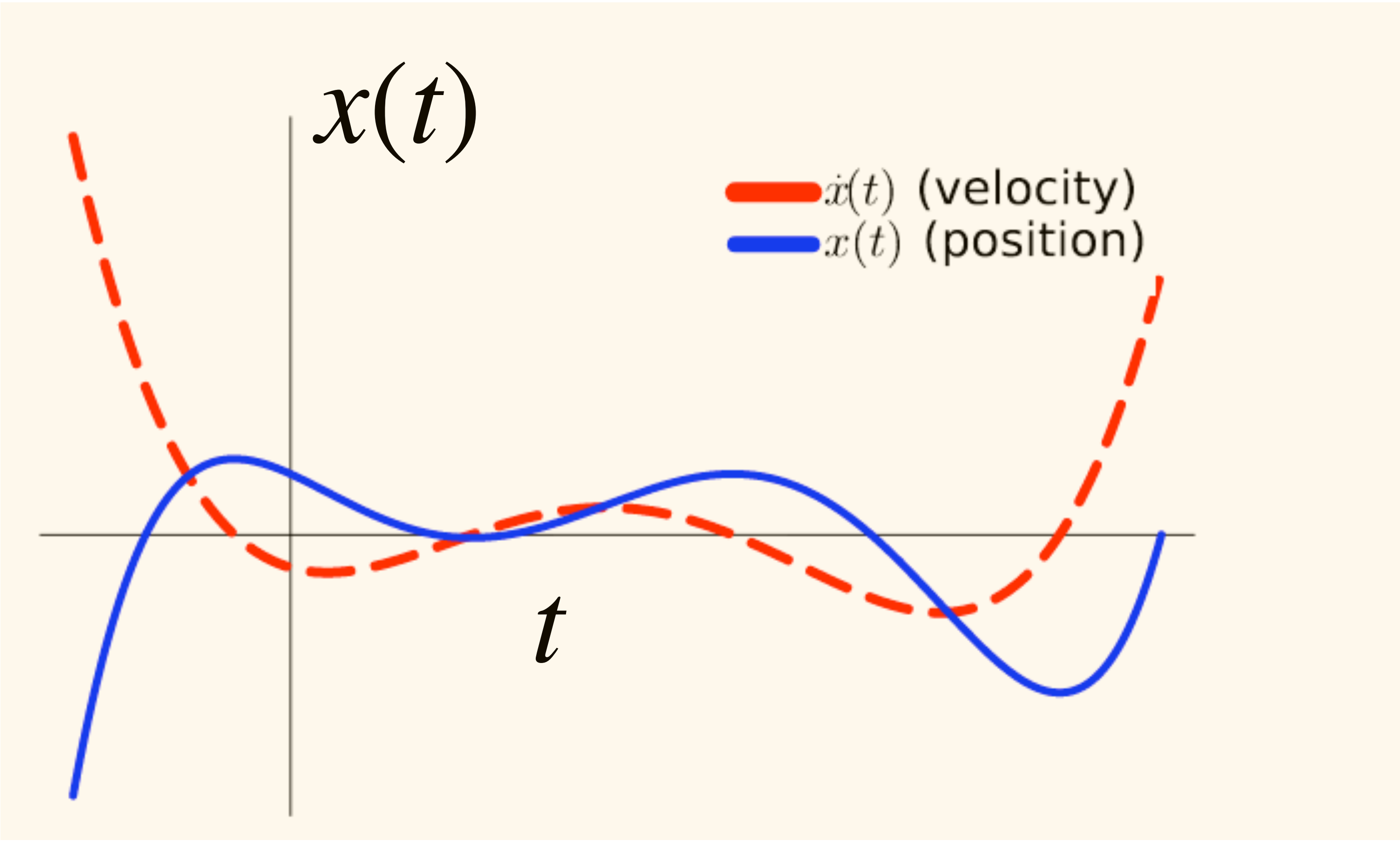
$$\frac{\mathrm{d}x}{\mathrm{d}t}(t)$$

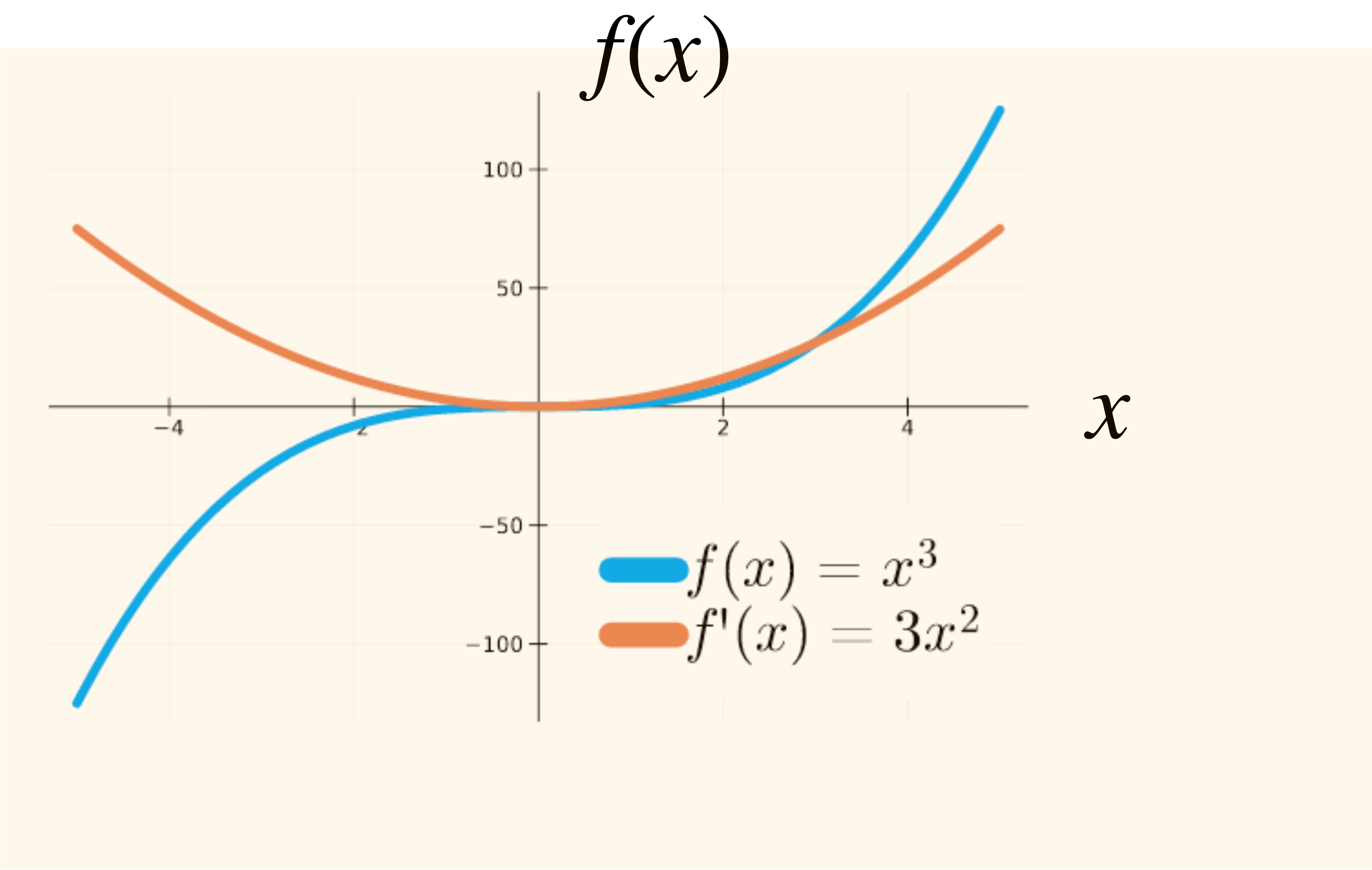
*How fast is it changing:*

$$\text{Time} \longrightarrow \text{Position}$$

Differentiated with respect to independent variable

# Two common (confusing) options




| Independent variable | Dependent variable | Derivatives |
|---|---|---|
| time $t$ | $x(t)$<br>(e.g. position) | $x'(t), x''(t), \ldots$ |
| space $x$ | $f(x)$<br>(Often called $y$) | $f'(x), f''(x), \ldots$ |

# A bog-standard function

## *Mapping between values*

$$f : \mathbb{R} \to \mathbb{R}$$

$$y = f(x)$$

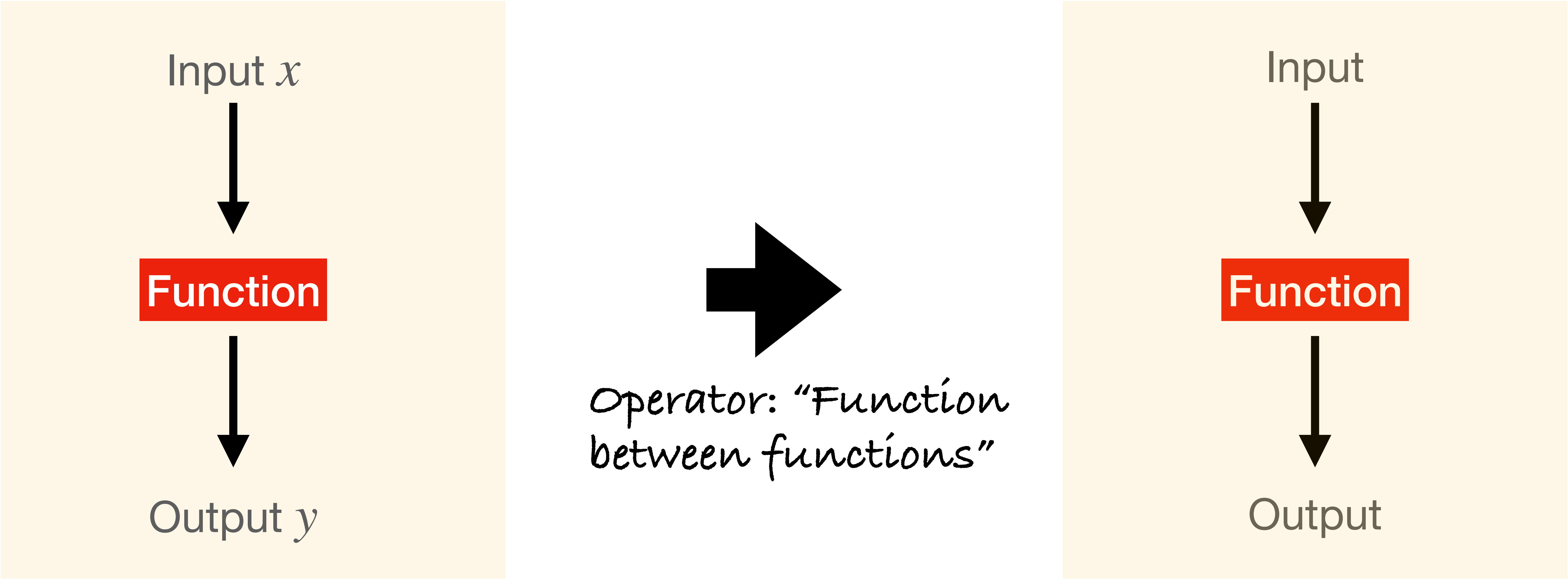
$y$ is the output .
$f$ is the mapping

*They are not the same*

$$x \xrightarrow{\quad f \quad} y$$

# Differentiation is an operator

## *Mapping between functions*




$D_x$ is the differential operator with respect to $x$

$\frac{\mathrm{d}f}{\mathrm{d}x}$ is the output of $D_x(f)$

*They are not the same*

$f$ $\xrightarrow{D_x}$ $\frac{\mathrm{d}f}{\mathrm{d}x}$

Input → Output

# Differentiation is an operator
## *Mapping between functions*

**What's the output here?**

Evaluating an operator at $f$:

$$D_x(f)$$

Function!

**What's the output here?**

Evaluating a function at $x$:

$$D_x(f)(x)$$

Number!

$$f \xrightarrow{D_x(f)} \frac{\mathrm{d}f}{\mathrm{d}x} \xrightarrow{D_x \circ D_x(f)} \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$$

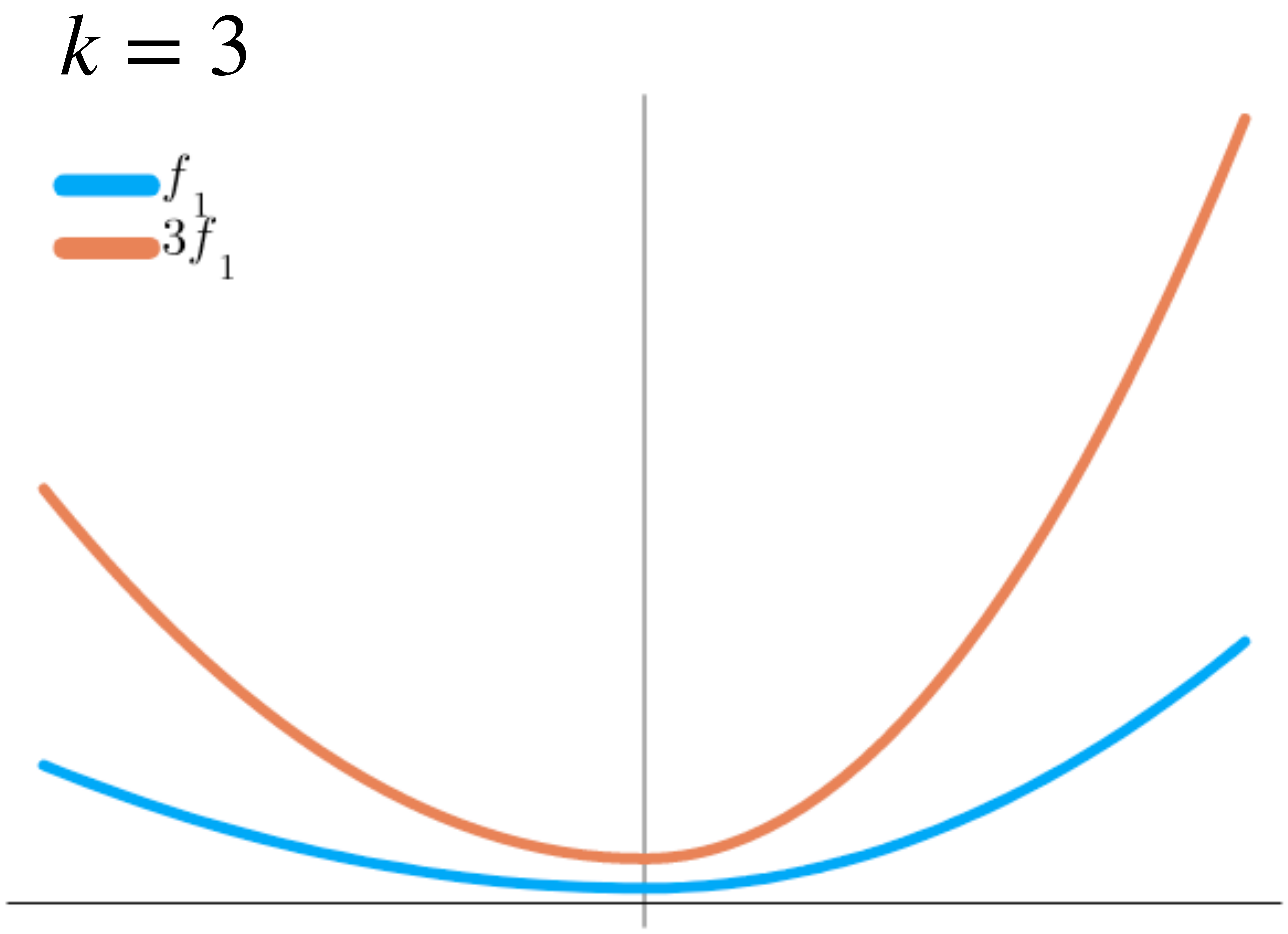
# Differentiation is a linear operator

$$kD_x(f) = D_x(kf)$$

Where $k \in \mathbb{R}$ and
$f$ is a function

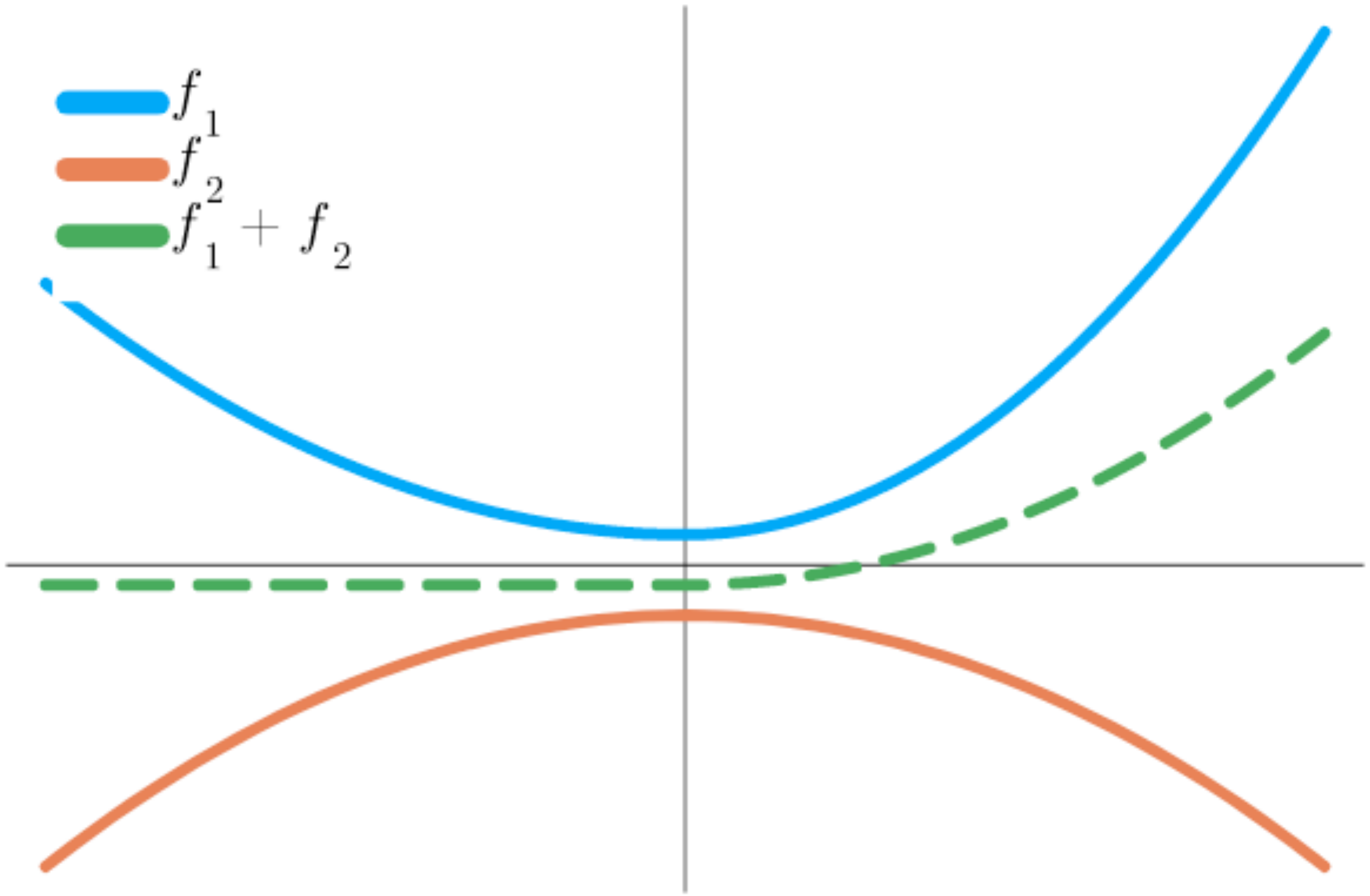
"Multiplying function by 3 scales the slope by 3!"

"Multiplying function by $k$ scales the slope by $k$!"



# Differentiation is a linear operator

$$D_x(f_1) + D_x(f_2) = D_x(f_1 + f_2)$$




"Adding two functions adds their slopes!"

# Differentiation is a linear operator

Overall:

$$\boxed{D_x\left(\sum_{i=1}^{n} a_i f_i\right) = \sum_{i=1}^{n} a_i D_x(f_i)}$$

*for any* $a_i \in \mathbb{R}$, $\quad n \in \mathbb{N}$,
*functions* $f$

# Example

$$f = f_1 + 3f_2 - 4f_3$$

$$f(x) = x^2 + 3x - 4$$

$$f_1(x) = x^2$$

$$f_2(x) = x$$

$$f_3(x) = 1$$

$$f_1'(x) = 2x$$

$$f_2'(x) = 1$$

$$f_3'(x) = 0$$

$$D_x(f) = \frac{\mathrm{d}f}{\mathrm{d}x} = D_x(f_1) + 3D_x(f_2) - 4D_x(f_3) \qquad = 2x + 3$$

# Standard derivatives

**Most calculus courses**

Calculating derivatives by hand

Needs lots of practice!

**MCMCS**

Use a computer!

| Function | Derivative |
|---|---|
| $y = x^n$ | $y' = nx^{n-1}$ |
| $y = \sin(x)$ | $y' = \cos(x)$ |
| $y = \cos(x)$ | $y' = -\sin(x)$ |
| $y = \ln(x)$ | $y' = 1/x$ |
| $y = e^x$ | $y' = e^x$ |

## Standard derivatives

| Function | Derivative |
|---|---|
| $y = x^n$ | $y' = nx^{n-1}$ |
| $y = \sin(x)$ | $y' = \cos(x)$ |
| $y = \cos(x)$ | $y' = -\sin(x)$ |
| $y = \ln(x)$ | $y' = 1/x$ |
| $y = e^x$ | $y' = e^x$ |

$$f(x) = 2x^3 + 3x^4 + 5\ln(x)$$

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = ? \qquad = 2 * (3x^2) + 3 * (4x^3) + \frac{5}{x}$$

$$= 6x^2 + 12x^3 + \frac{5}{x}$$

# Symbolic computation

$$f(x) = x^2 + 3x - 4$$

$$f'(x) = 2x + 3$$

[ 2 Items

0: $2x + 3$

1: $9$

]

```python
import sympy as sp

x = sp.symbols('x')
g = x**2 + 3*x + 4
g_prime = sp.diff(g, x)
(g_prime, g_prime.subs(x, 3))
```

Don’t use ChatGPT:
confidently incorrect

9

```python
g_prime_numeric = sp.lambdify(x, g_prime, 'numpy')
g_prime_numeric(3)
```

# Symbolic computation

Right-click to copy-paste Latex code!

```julia
begin
    using Symbolics
    @variables t # independent variable
    Dₜ = Differential(t)
    y = t^2 + 4sin(t) + log(t)
end
```

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\log\left(t\right) + 4\sin\left(t\right) + t^{2}\right)$$

```julia
Dₜ(y)
```

$$2t + \frac{1}{t} + 4\cos\left(t\right)$$

```julia
expand_derivatives(Dₜ(y))
```

# Symbolic computation

***…is slow***

$$f(x) = x^2 + 3x - 4$$

$$f'(x) = 2x + 3$$

Necessary if you need the formula $(2x + 3)$

Otherwise…?

# Automatic differentiation

$$f(x) = x^2 + 3x - 4$$

$$f'(x) = 2x + 3$$

```
9.0
```

```python
from autograd import grad
def f(x):
    return x**2 + 3*x + 4
dfdx = grad(f)
dfdx(3.0) #3 without decimal will give an error!
```

Grad means derivative! (Next lecture)

No formula but fast and accurate

Autograd is outdated but works on Marimo weblinks

Jax (modern version) works similarly

# Break: Alice in Wonderland

Lewis Carroll

"You need to be much, muchier.
You've lost your muchness"

"Alice: 'But this is impossible.' The
Mad Hatter: 'Only if you believe it is.'"

I could tell you my adventures — beginning from this
morning,' said Alice a little timidly: 'but it's no use going
back to yesterday, because I was a different person then.'"

# Break: Alice in Wonderland

Lewis Carroll

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f(x^*)}{\Delta x}$$

**19th Century**

Mathematical framework for infinitesimal quantities

➡

Lewis Carroll was a conservative mathematician: thought it was ridiculous

Alice in Wonderland is a pisstake of the mathematical foundations of calculus!

# Differential operator recap

What about multiplying functions?

$$fg(x) = f(x) \times g(x)$$

Product rule!

What about composing functions?

$$f \circ g(x) = f\big(g(x)\big)$$

Chain rule!

Linear operator

$$\boxed{D_x\left(\sum_{i=1}^{n} a_i f_i\right) = \sum_{i=1}^{n} a_i D_x(f_i)}$$

*for any* $a_i \in \mathbb{R}, \quad n \in \mathbb{N},$
*functions* $f$

Add functions -> add derivatives

Scale functions -> scale derivatives

# TLDR

**Product rule**

$$\boxed{D_x\left(fg\right) = gD_x(f) + fD_x(g)}$$

**Chain rule**

$$\boxed{D_x\left(f \circ g\right) = D_g(f) \times D_x(g)}$$

Not the same as $D_x(f)$!!!!

# Examples

**Product rule**

$$\boxed{D_x\left(fg\right) = gD_x(f) + fD_x(g)}$$

**Want to differentiate**

$$h(x) = x^3 \cos(x)$$

$$h(x) = f(x)g(x)$$

$$f(x) = x^3 \qquad g(x) = \cos(x)$$

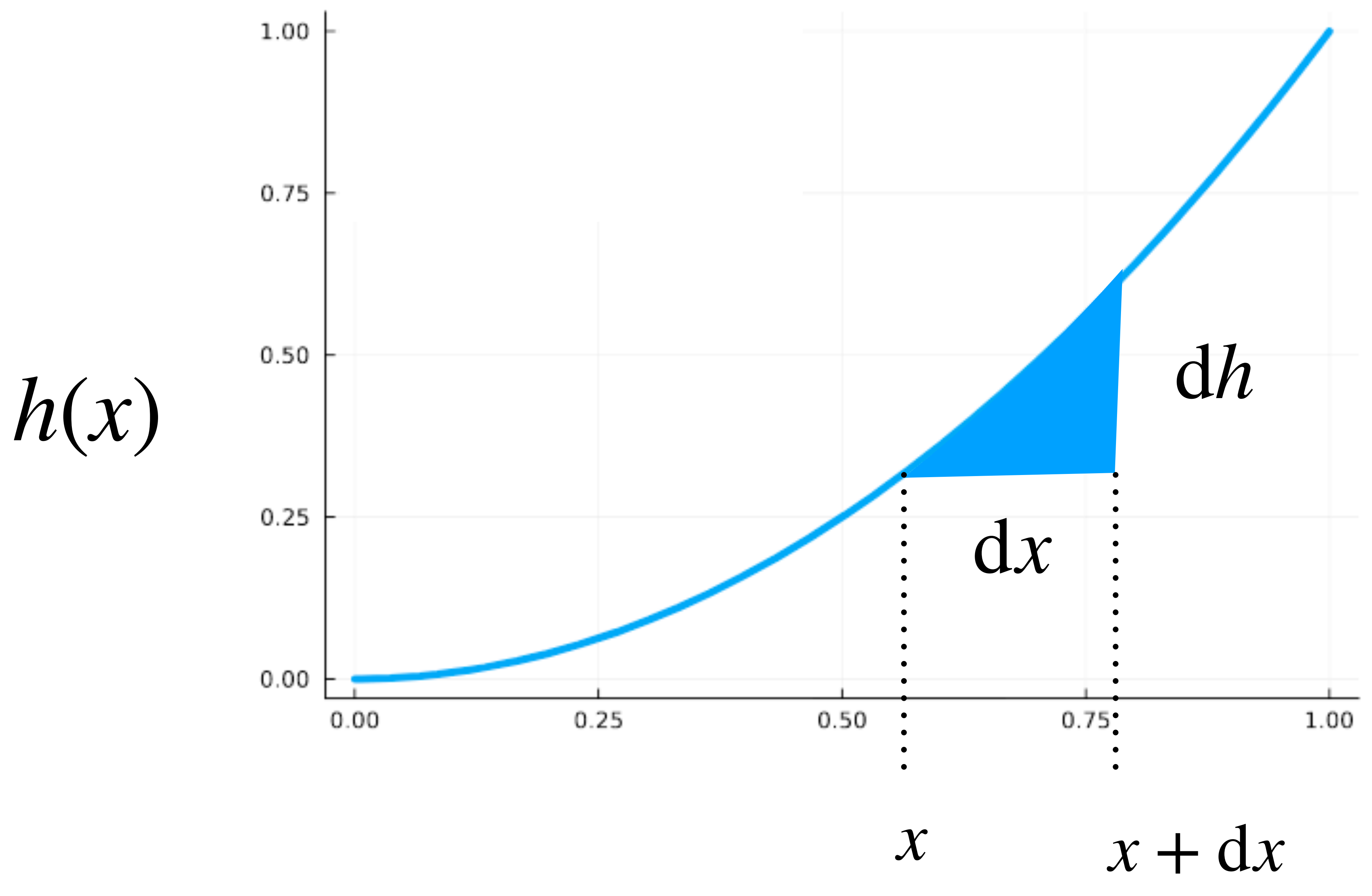
$$D_x(f) = 3x^2 \qquad D_x(g) = -\sin(x)$$

$$D_x(f) = 3x^2 \cos(x) - x^3 \sin(x)$$

# Why does it work?

$$h(x) = f(x)g(x)$$

$$h'(x) \quad \Longleftrightarrow \quad \mathrm{d}h(x) = h'(x)\ \mathrm{d}x$$




**Assume we've calculated:**

$f'(x) \qquad g'(x)$

$$\Longrightarrow \quad \begin{aligned} \mathrm{d}f &= f'(x)\ \mathrm{d}x \\ \mathrm{d}g &= g'(x)\ \mathrm{d}x \end{aligned}$$

$$\mathrm{d}h(x) = \left(\frac{\mathrm{d}h}{\mathrm{d}x}(x)\right)\mathrm{d}x$$

$h'(x)$

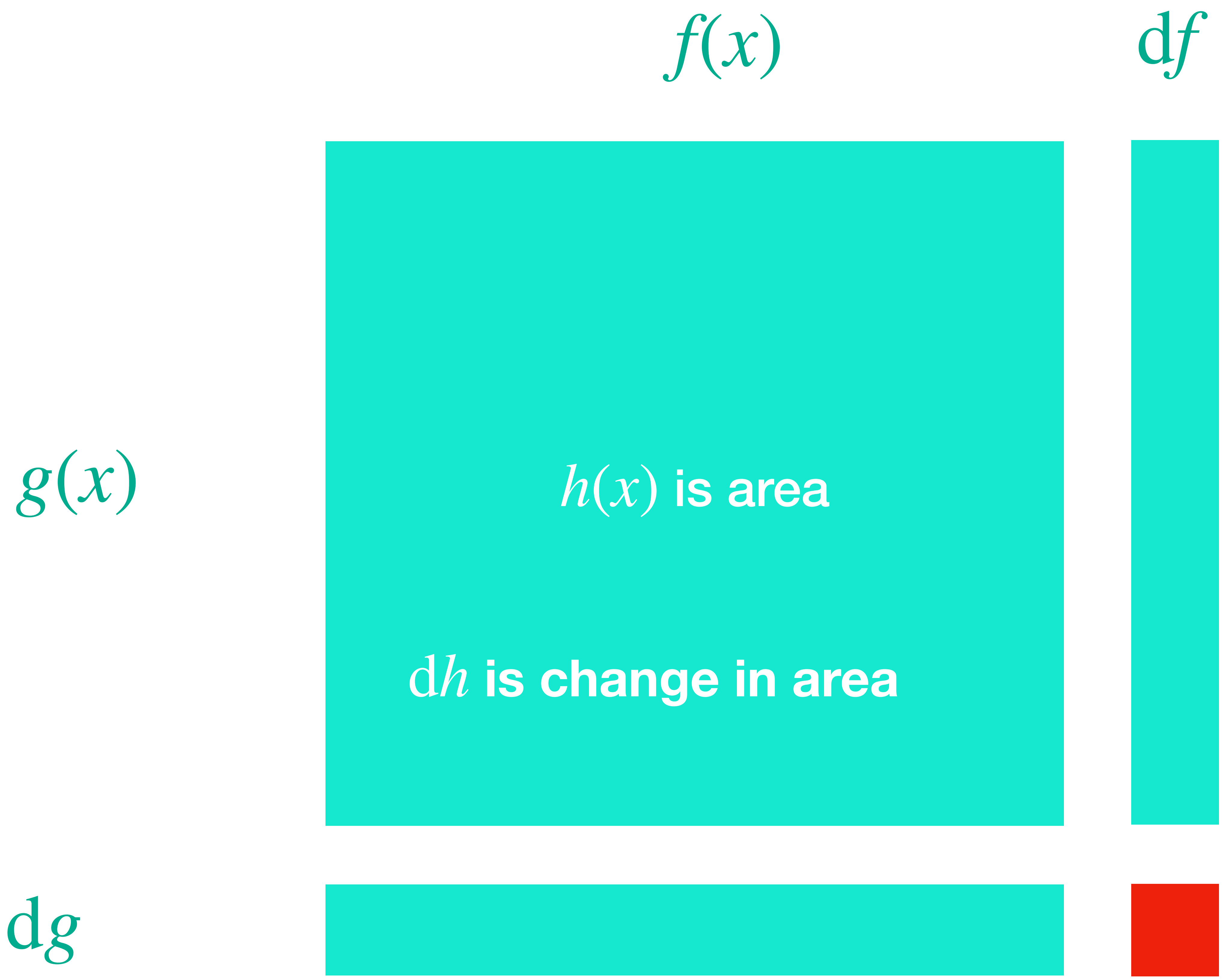
# Why does it work?

$$h(x) = f(x)g(x)$$

$h'(x)?$ $\qquad$ $\mathrm{d}h = \,?\, \mathrm{d}x$

$$x \to x + \mathrm{d}x$$

$$h(x) \to h(x) + \mathrm{d}h$$




**Adding new rectangles**

$$\mathrm{d}h = f(x)\mathrm{d}g + g(x)\mathrm{d}f + \ \mathrm{d}f\mathrm{d}g$$

$$\mathrm{d}h = f(x)g'(x)\mathrm{d}x + g(x)f'(x)\mathrm{d}x + \ \mathrm{d}f\mathrm{d}g$$

$\mathrm{d}g = g'(x)\mathrm{d}x$

$\mathrm{d}f = f'(x)\mathrm{d}x$

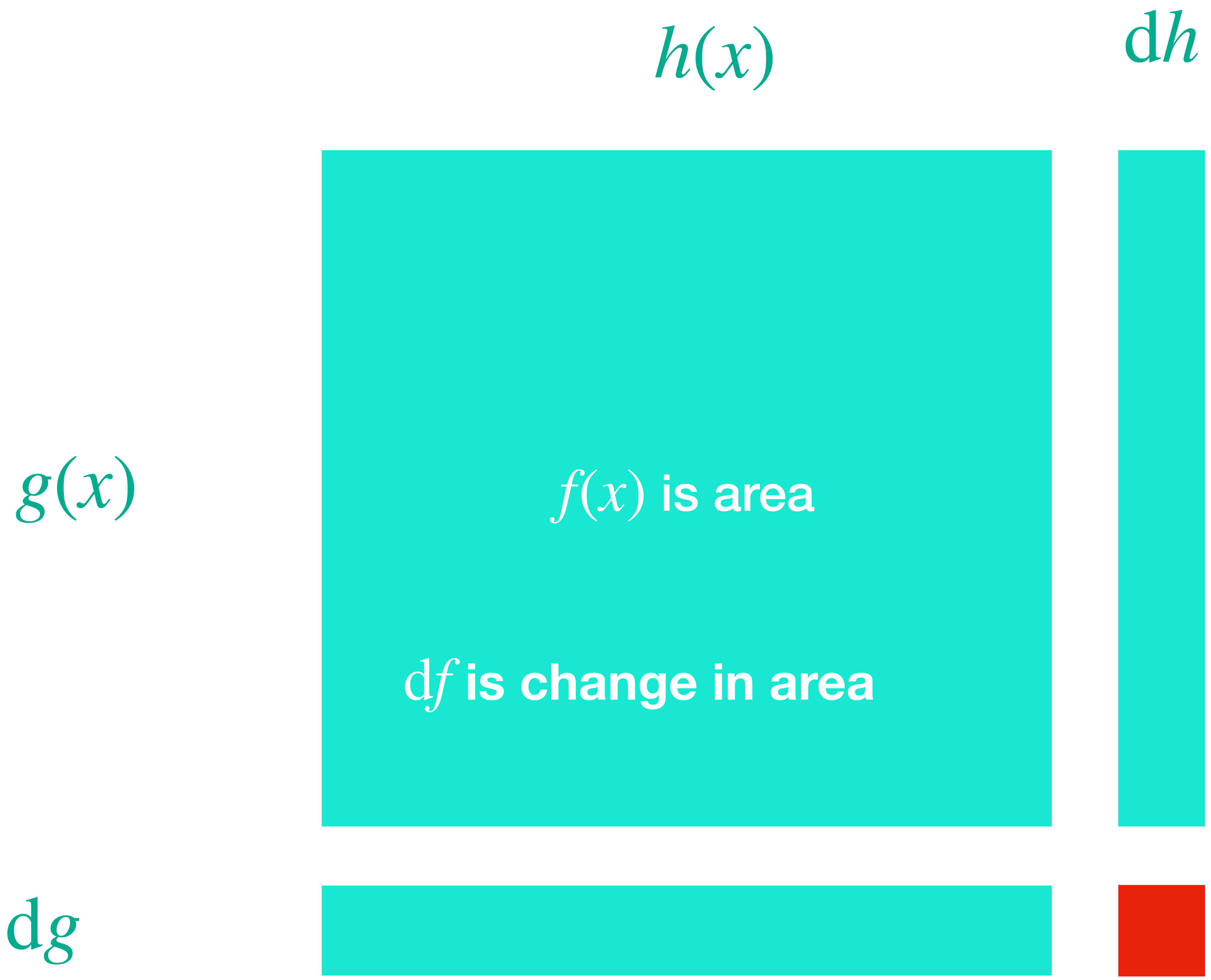
# Why does it work?

$$x \to x + \mathrm{d}x$$

$$f(x) \to f(x) + \mathrm{d}f$$

$$h(x) = f(x)g(x)$$

$h'(x)?$ $\qquad\qquad$ $\mathrm{d}h = \ ?\ \mathrm{d}x$




$$\mathrm{d}h = f(x)g'(x)\mathrm{d}x + g(x)f'(x)\mathrm{d}x + \mathrm{d}f\mathrm{d}g$$

$$\frac{\mathrm{d}h}{\mathrm{d}x} = f(x)\frac{\mathrm{d}g}{\mathrm{d}x} + g(x)\frac{\mathrm{d}f}{\mathrm{d}x} + \frac{\mathrm{d}f\mathrm{d}g}{\mathrm{d}x}$$

# Why does it work?

$$\frac{\cancel{\mathrm{d}f\mathrm{d}g}}{\mathrm{d}x} \to 0 \quad ?$$

$$\left|\frac{\mathrm{d}f\mathrm{d}g}{\mathrm{d}x}\right|$$

$$= \left| f'(x)g'(x)\frac{\mathrm{d}x^2}{\mathrm{d}x} \right| = |f'(x)g'(x)\mathrm{d}x|$$

$$\mathrm{d}g = g'(x)\mathrm{d}x$$

$$\mathrm{d}f = f'(x)\mathrm{d}x$$

$= 0$ as $\mathrm{d}x \to 0$

$$\frac{\mathrm{d}h}{\mathrm{d}x} = f(x)\frac{\mathrm{d}g}{\mathrm{d}x} + g(x)\frac{\mathrm{d}f}{\mathrm{d}x} + \frac{\mathrm{d}f\mathrm{d}g}{\mathrm{d}x}$$

# Product rule

$$\boxed{D_x\left(fg\right) = gD_x(f) + fD_x(g)}$$

$$\Updownarrow$$

$$\frac{\mathrm{d}h}{\mathrm{d}x} = f(x)\frac{\mathrm{d}g}{\mathrm{d}x} + g(x)\frac{\mathrm{d}f}{\mathrm{d}x}$$

# Chain rule

## Chain rule

$$\boxed{D_x\left(f \circ g\right) = D_g(f) \times D_x(g)}$$

$$h(x) = f \circ g(x)$$

$$g(x) = x^2$$

$$f(x) = \sin(x)$$

| $g \circ f?$ | $f \circ g?$ | $fg?$ | $gf?$ |
|---|---|---|---|
| $[sin(x)]^2$ | $\sin(x^2)$ | $x^2 \sin(x)$ | $x^2 \sin(x)$ |

# Chain rule

### Chain rule

$$\boxed{D_x\left(f \circ g\right) = D_g(f) \times D_x(g)}$$

$$h(x) = f \circ g(x)$$

**Example**

$$h(x) = \cos(x^3)$$

**Step 1: break into pieces**

$$f(u) = \cos(u)$$

$$u = g(x) = x^3$$

**Step 2: calculate derivatives**

$$\mathrm{d}f(u) = -\sin(u)\mathrm{d}u$$

$$\mathrm{d}u = 3x^2\mathrm{d}x$$

**Step 3: Put together**

$$\mathrm{d}f(x) = -\sin(x^3)3x^2\mathrm{d}x$$

# How does it work?

$$h(x) = f \circ g(x)$$

$$\boxed{D_x\left(f \circ g\right)?}$$

**Step 1: Rewrite**

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

$$\begin{tikzcd}
& \mathrm{d}u \arrow[dr] & \\
\mathrm{d}x \arrow[ur] \arrow[rr] & & \mathrm{d}h
\end{tikzcd}$$

# Chain rule

$$h(x) = f \circ g(x)$$

**Step 1: Rewrite**

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

**Step 2:**
**Individual derivatives**

$$\mathrm{d}u = g'(x)\mathrm{d}x$$

$$\mathrm{d}h = f'(u)\ \mathrm{d}u$$

**Step 3:**
**Put together**

$$\mathrm{d}h = f'(u)g'(x)\ \mathrm{d}x$$

$$\begin{tikzcd} & \mathrm{d}u \arrow[dr] & \\ \mathrm{d}x \arrow[ur] \arrow[rr] & & \mathrm{d}h \end{tikzcd}$$

# Summary

Already know how to add, multiply and compose functions

$$f(x) = x^2$$
$$g(x) = 4$$

| | |
|---|---|
| $h(x) = x^2 + 4$ | $h = f + g$ |
| $h(x) = 4x^2$ | $h = f \times g$ |
| $h(x) = 16$ | $h = f \circ g$ |

These have an algebra

$$f + g = g + f, \ldots$$
$$fg = gf, \ldots$$

(Actually, can make vector spaces of functions!)

# Summary

We've worked out an algebra for the differential operator

| | |
|---|---|
| **Operator** | **Derivative/differential** |
| Mapping between functions | Steepness of functions |

What happens to the steepness if I compose, add, multiply functions?

**Linearity**

$$\boxed{D_x\left(\sum_{i=1}^{n} a_i f_i\right) = \sum_{i=1}^{n} a_i D_x(f_i)}$$

**Product rule**

$$\boxed{D_x(fg) = gD_x(f) + fD_x(g)}$$

**Chain rule**

$$\boxed{D_x(f \circ g) = D_g(f) \times D_x(g)}$$

# Why?

## Differentiation operator tells us how (fast) things change

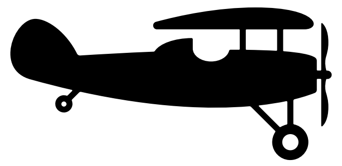
**Optimisation/machine learning**

Need differential operator to understand learning

**Dynamical systems**

Need differential operator to model processes that change over based on current state

*Electrical/mechanical system*



*Pandemic modelling!*

# It's difficult!

Will get lots of practice over next few weeks

Apply (maths + coding) helps with concepts

Equations are long, and need lots of memory-space. But make sense!