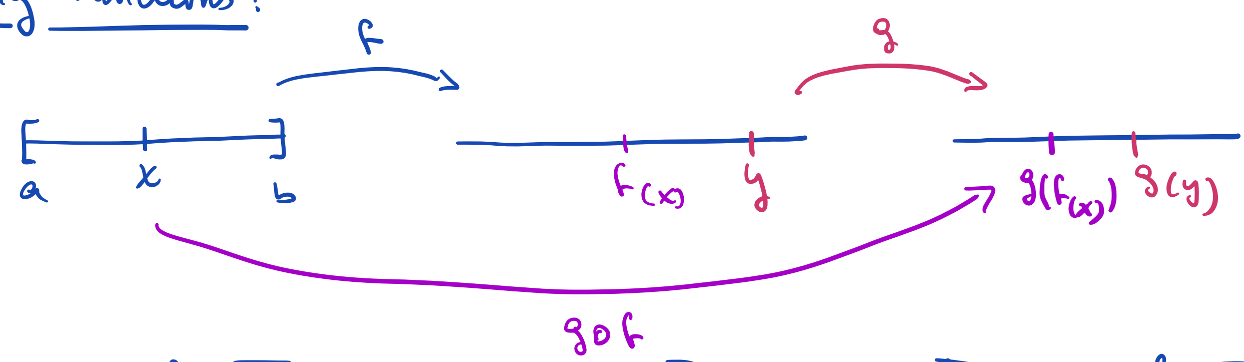
Lecture IX: §3.3 Composite Functions & the Chain Rule

§1 Composing Functions:

IDEA



We are given 2 functions $f:[a,b] \longrightarrow \mathbb{R}$. We create a new function
$g: \mathbb{R} \longrightarrow \mathbb{R}$

$g\circ f : [a,b] \longrightarrow \mathbb{R}$, called the composition of g with f. At each pt x it takes value $g\circ f(x) = g(f(x))$

Note: We first apply f and THEN apply g to $y=f(x)$

GOAL: We want to see how nice properties of f & g pass into $g\circ f$
(Main properties: continuity & differentiability)

Theorem 1: If f is continuous at x_0 & g is continuous at $y_0=f(x_0)$, then $g\circ f$ is continuous at x_0.

Theorem 2: (Chain Rule) If f is differentiable at x_0 & g is differentiable at $f(x_0)$, then $g\circ f$ is differentiable at x_0. Furthermore:

$$\frac{d(g\circ f)}{dx}\Big|_{x=x_0} = \frac{dg}{dy}\Big|_{y=f(x)} \cdot \frac{df}{dx}\Big|_{x=x_0}.$$

(Easy way to remember $\frac{d(g\circ f)}{dx} = \frac{dg}{df}\Big|_{f(x)} \cdot \frac{df}{dx}$.

Example: $f(x) = (x^3+4x)$ $\quad g(y) = y^{10} \rightsquigarrow g\circ f(x) = (x^3+4x)^{10}$

Clear: $g\circ f$ is a polynomial, so it's differentiable.

- $\frac{d\, g\circ f}{dx} = \frac{d\,y^{10}}{dy}\Big|_{y=f(x)} \cdot \frac{df}{dx} = 10y^9\Big|_{y=f(x)} \cdot (3x^2+4) = 10(x^3+4x)^9(3x^2+4)$

- Alternative way: expand $(x^3+4x)^{10}$ using the Binomial Theorem, and differentiate (this sounds terrible!)
- Composing f with g? $f\circ g_{(y)} = (x^3+4x)\Big|_{x=y^{10}} = y^{30}+4y^{10}$

So $\dfrac{d\, f\circ g}{dy} = 30y^{29}+40y^9$

Check this formula using the Chain Rule:

$$\frac{d\, f\circ g}{dy} = \frac{df}{dx}\Big|_{x=g(y)} \cdot \frac{dg}{dy} = (3x^2+4)\Big|_{x=y^{10}} \cdot 10y^9 = (3y^{20}+4)\,10y^9 = 30y^{29}+40y^9$$

General result: $h_{(x)} = f_{(x)}^n$ for n integer has derivatives:

$$\frac{dh}{dx} = n\left(f_{(x)}\right)^{n-1}\cdot\frac{df}{dx}.$$

§2. Why is Theorem 1 valid?

GOAL: To show $\lim\limits_{x\to x_0} g\circ f_{(x)} = g\circ f_{(x_0)}$ via ε/δ

Given $\varepsilon>0$, we want to find $\delta>0$ so that if $0<|x-x_0|<\delta$, we always have $|g\circ f_{(x)} - g\circ f_{(x_0)}|<\varepsilon$.

We use the info on f & g.

intermediate step.

① Since g is continuous at $y_0=f(x_0)$ we can find $\alpha>0$ so that if $0<|y-y_0|<\alpha$, then $|g(y)-g(y_0)|<\varepsilon$

② We want to replace y by $f_{(x)}$. But to do so, we must first ensure that $0<|f_{(x)}-y_0|<\alpha$. But we can always ensure that this is true if x is closed enough to x_0.

Indeed pick $\varepsilon'=\alpha>0$. The continuity of f at x_0 ensures that we can find $\delta>0$ so that if $0<|x-x_0|<\alpha$, then $|f_{(x)}-f_{(x_0)}|<\alpha$.

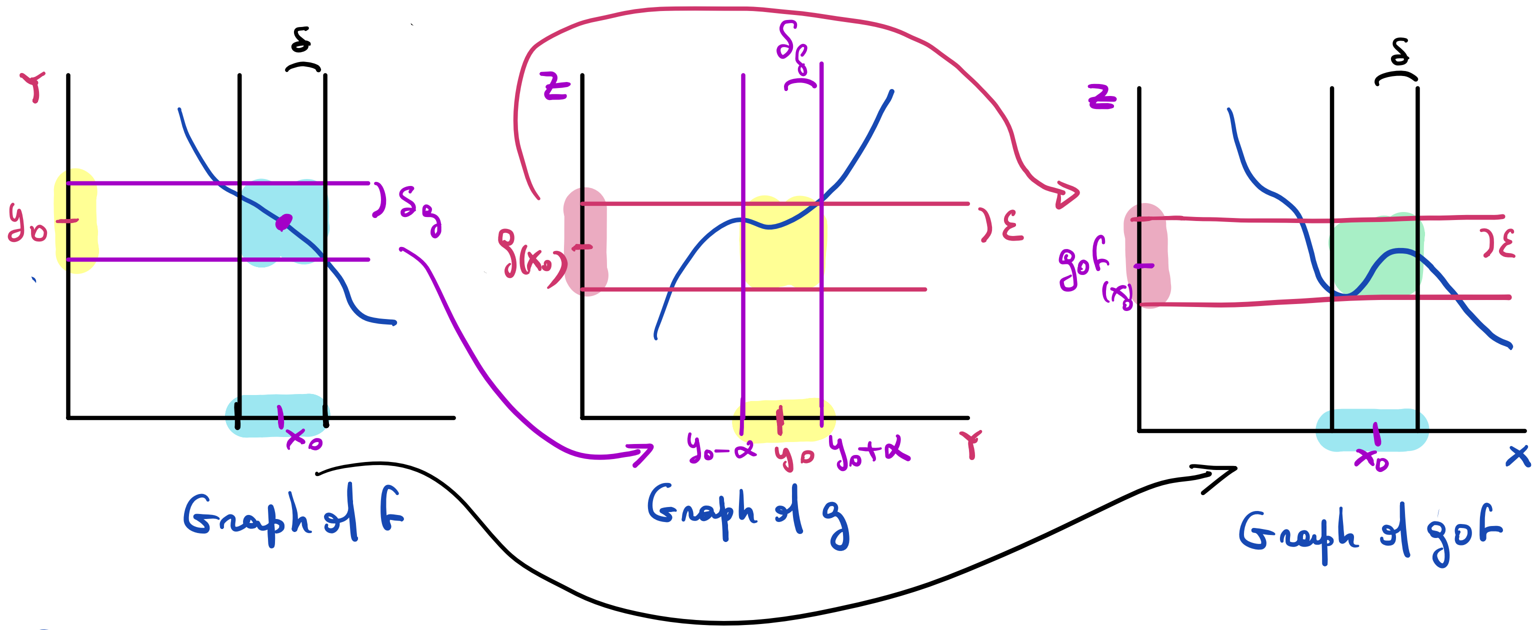
③ Now, we reverse engineer the process:

If $|x-x_0|<\delta$, by ② we get $|\overset{=y}{f(x)} - \overset{=y_0}{f(x_0)}| < \delta_g$

Now, by ① we get $|g(f(x)) - g(f(x_0))| < \varepsilon$.

(by our choice of α). This is what we wanted.

Pictorial argument:



Summary: $\alpha>0$ was the middleman that allowed us to find $\delta>0$ given $\varepsilon>0$

§3 Why is Theorem 2 valid?

GOAL: Show $\dfrac{d\, g\circ f}{dx} = \dfrac{dg}{dy}\Big|_{y=f(x)} \cdot \dfrac{df}{dx}$ using the definition

$$\frac{d(g\circ f)}{dx} = \lim_{\Delta x\to 0} \frac{\Delta(g\circ f)}{\Delta x} = \lim_{\Delta x\to 0} \frac{g\circ f(x+\Delta x) - g\circ f(x)}{\Delta x}$$

Now, we want to see $f(x+\Delta x)$ as $f(x) + \Delta y$ ⇝ $\Delta y = f(x+\Delta x) - f(x)$

So
$$\frac{d(g\circ f)}{dx} = \lim_{\Delta x\to 0} \frac{g(f(x+\Delta x)) - g(f(x))}{\Delta x}$$

$$= \lim_{\Delta x\to 0} \frac{g(f(x)+\Delta y) - g(f(x))}{\Delta x} \overset{\to}{=}$$ multiply & divide by Δy. & rearrange.

$$⚠ = \lim_{\Delta x\to 0} \boxed{\frac{g(f(x)+\Delta y) - g(f(x))}{\Delta y}}\ \boxed{\frac{\Delta y}{\Delta x}} \qquad (*)$$

Q: What's this limit? has limit $\dfrac{df}{dx}$ as $\Delta x\to 0$

Claim: f is continuous a x because it's differentiable, so we have

$$\Delta y = f(x+\Delta x) - f(x) \xrightarrow[\Delta x \to 0]{} 0$$

Then: $$\frac{dg}{dy}\Big|_{y=f(x)} \lim_{\Delta y \to 0} \frac{g(f(x)+\Delta y) - g(f(x))}{\Delta y} \overset{\text{By our Claim}}{=} \lim_{\Delta x \to 0} \frac{g(f(x)+\Delta y) - g(f(x))}{\Delta x}$$

Conclusion: $$\frac{d(g\circ f)}{dx} = \frac{dg}{dy}\Big|_{y=f(x)} \cdot \frac{df}{dx}$$ by the product Rule for Limits.

⚠ There is only one issue: How do we know $\Delta y \neq 0$ if Δx is close enough to 0? (we divided by Δx in (*)!)

If we cannot, this means that we have Δx as close as 0 as we want with $\Delta y = f(x+\Delta x) - f(x) = 0$. But for these points we have

$$g(f(x)+\Delta y) - g(f(x)) = g(f(x)) - g(f(x)) = 0$$

so $$\frac{d\, g\circ f}{dx} = \lim_{\Delta x \to 0} \frac{\boxed{g(f(x)+\Delta y) - g(f(x))}}{\Delta x} = 0$$

and $$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{\boxed{f(x+\Delta x) - f(x)}}{\Delta x} = 0$$ as well.

- **Alternative argument** (due to Artin) that avoids dividing by Δy with $\Delta y = f(x+\Delta x) - f(x)$

Write $$\varepsilon(\Delta y) := \frac{g(y+\Delta y) - g(y)}{\Delta y} - g'(x)$$

Since $g'(y) = \lim_{\Delta y \to 0} \frac{g(y+\Delta y) - g(y)}{\Delta y}$, we know $\lim_{\Delta y \to 0} \varepsilon(\Delta y) = 0$.

Equivalently: $$\boxed{g(y+\Delta y) - g(y) = \Delta y\, \varepsilon(\Delta y) + \Delta y\, g'(y)}$$

with $\varepsilon(\Delta y) \xrightarrow[\Delta y \to 0]{} 0$

- Now: $h = g\circ f(x)$ satisfies

$$\frac{h(x+\Delta x) - h(x)}{\Delta x} = \frac{g(y+\Delta y) - g(y)}{\Delta x} \text{ with } \begin{cases} y = f(x) \\ \Delta y = f(x+\Delta x) - f(x) \end{cases}$$

- Replace numerator by the right-hand side of the expression $\boxed{}$

We get $\dfrac{h(x+\Delta x)-h(x)}{\Delta x} = \dfrac{\Delta y\, \varepsilon(\Delta y) + \Delta y\, g'(y)}{\Delta x}$

$$\underset{\substack{\downarrow \\ \text{replace } \Delta y}}{=} \underbrace{\boxed{\dfrac{f(x+\Delta x)-f(x)}{\Delta x}}}_{\downarrow \Delta x \ \frac{df}{dx}} \cdot \big(\underbrace{\varepsilon(\Delta y)}_{\downarrow \Delta y \to 0 \ \ 0} + g'(y)\big)$$

& $\Delta y \to 0$ because $\Delta x \to 0$ & f is continuous.

<u>Conclude</u> $\dfrac{dh}{dx} = \lim\limits_{\Delta x \to 0} \dfrac{h(x+\Delta x)-h(x)}{\Delta x} = \dfrac{df}{dx} \cdot (0 + g'(y)) = \dfrac{df}{dx} \dfrac{dg}{dy}\Big|_{y=f(x)}$

§4 Examples:

① $h(x) = \sin^5(x)$ $\qquad h'(x) = ?$

Write h as a composite function:

$h(x) = g\circ f(x)$ with $g(y) = y^5$ $\quad f(x) = \sin x$

The chain rule gives $h'(x) = g'(f(x)) \cdot f'(x)$

$g'(y) = 5y^4 \rightsquigarrow g'(f(x)) = 5\, f(x)^4 = 5 \sin^4 x$

Assume that we know $\dfrac{d \sin(x)}{dx} = \cos x$ (we'll see it later in the course),

so $f'(x) = \cos x$

<u>Conclude</u>: $h'(x) = 5 \sin^4 x \cos x$

② $h(x) = \sin^5(x+1) = g\circ f(x)$ for $g(y) = \sin^5 y$ & $f(x) = x+1$

So $h'(x) = g'(f(x)) \cdot f'(x) = 5 \sin^4(x+1) \cos(x+1) \cdot 1$ because

- $g'(y) \underset{\text{Ex ①}}{=} 5 \sin^4 y \cos y$, so $g'(f(x)) = 5 \sin^4(x+1) \cos(x+1)$

- $f'(x) = \dfrac{d(x+1)}{dx} = 1$.

③ $h(x) = (5x^2+3)^{10} (x^4-1)$ $\qquad h'(x) = ?$

We use the Product Rule & the Chain Rule

Set $(5x^2+3)^{10} = h_1(x)$ & $(x^4-1) = h_2(x)$

Product Rule gives $h'(x) = (h_1 h_2)' = h_1' h_2 + h_1 h_2'$

$h_1'(x) = ?$ but $h_2'(x) = 4x^3$ (from Lecture VIII)

To get h_1' we will use the Chain Rule:

$h_1 = (5x^2+3)^{10} = g \circ f(x)$ for $g(y) = y^{10}$ $f(x) = 5x^2+3$

$g'(y) = 10y^9$ so $g'(f(x)) = 10(5x^2+3)^9$

$f'(x) = 10x$

So $h_1'(x) = g'(f(x)) \cdot f'(x) = 10(5x^2+3)^9 \cdot 10x = 100x(5x^2+3)^9$

Conclude: $h(x) = h_1' h_2 + h_1 h_2' = 100x(5x^2+3)^9(x^4-1) + (5x^2+3)^{10} \cdot 4x^3$

$= (5x^2+3)^9 x \,(100(x^4-1) + 4x^2(5x^2+3))$

$= (5x^2+3)^9 x \,(100x^4 - 100 + 20x^4 + 12x^2)$

$= 5x^2+3)^9 x \,(120x^4 + 12x^2 - 100)$