CALCULUS 1, NOTE
DIFFERENTIATION

TANG-KAI LEE

3. **Differentiation**

We will now use the technique of taking limits to study derivatives of functions. From now on, if we say $f\colon I \to \mathbb{R}$ is a function, we usually implicitly mean that I is an open interval.

3.1. **Definition and example.**

Recall the tangent problem we mentioned as a motivation of studying limits. There, the limit that arises is

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0},$$

which, if exists, should be the slope we use to define the tangent of the graph of a function. Note that by viewing $x = x_0 + h$ as a small perturbation of x_0, we can write

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}.$$

When this limit exists, it is called the **derivative** of f at x_0. We give a formal definition as follows.

Definition 3.1. Let f be a real-valued function on an open interval $I \subseteq \mathbb{R}$. For $x \in I$, if the limit

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

exists, then we say f is **differentiable** at x and the limit is called the derivative of f at x. There are a few ways to denote it, and the most common ones are

$$f'(x) = \frac{df}{dx}(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}.$$

For a function $f\colon I \to \mathbb{R}$, if $f'(x)$ exists for all $x \in I$, we say f is a **differentiable function** on I. In this case, f' can also be viewed as a function on I.

Example 3.2. We see some examples.

(1) Let $f\colon \mathbb{R} \to \mathbb{R}$ be $f(x) = x$. Then for any $x \in \mathbb{R}$,

$$f'(x) = \lim_{h\to 0} \frac{(x+h)-x}{h} = \lim_{h\to 0} 1 = 1.$$

Thus, $f' = 1$ is a constant function.

(2) Let $f\colon \mathbb{R} \to \mathbb{R}$ be $f(x) = x^2$. Then for any $x \in \mathbb{R}$,

$$f'(x) = \lim_{h\to 0} \frac{(x+h)^2-x^2}{h} = \lim_{h\to 0} \frac{2xh+h^2}{h} = \lim_{h\to 0} (2x+h) = 2x.$$

Thus, $f'(x) = 2x$.

Date: March 3, 2026.

Recall that we define e to be the number such that the slope of the tangent of the graph of e^x is 1 at $x = 0$. This means that e is the number such that

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1. \tag{3.3}$$

We can use this to find out the derivative of e^x at any other point.

(3) Let $f\colon \mathbb{R} \to \mathbb{R}$ be $f(x) = e^x$. Then for any $x \in \mathbb{R}$,

$$f'(x) = \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} = \lim_{h \to 0} e^x \frac{e^h - 1}{h} = e^x \cdot \lim_{h \to 0} \frac{e^h - 1}{h} = e^x$$

by (3.3). Thus, $f'(x) = e^x$. That is, the derivative of e^x is itself!

Example 3.4. As a result of the previous examples, we can now find out the tangents of these graphs.

(1) We find the tangent of the graph of $f(x) = x^2$ at $x = 3$. By above, $f'(x) = 2x$, so $f'(3) = 6$. That is, the tangent passes through $(3, 9)$ and has slope 6. Thus, it is defined by

$$y - 9 = 6(x - 3),$$

which is $y = 6x - 9$.

(2) We find the tangent of the graph of $f(x) = e^x$ at $x = \log 2$. By above, $f'(x) = e^x$, so $f'(\log 2) = 2$. That is, the tangent passes through $(\log 2, 2)$ and has slope 2. Thus, it is defined by

$$y - 2 = 2(x - \log 2),$$

which is $y = 2x + 2 - \log 4$.

Derivatives appear in many problems. It usually arises when we are looking for the following concepts.

(1) Instantaneous change rate.

(2) First-order approximation.

We know some examples for rates of change. These include velocity and acceleration.

For first-order approximate, we will look into it in more detail later, but we can get an idea first. Given a function $f\colon I \to \mathbb{R}$ and $a \in I$, if $f'(a)$ exists, that roughly means that for x close to a, we get

$$f'(a) \approx \frac{f(x) - f(a)}{x - a},$$

which means

$$f(x) \approx f'(a)(x - a) + f(a).$$

This tells us that $f(x)$ behaves similarly to what a linear polynomial does when x is close enough to a. In fact, this linear polynomial is exactly the tangent of $f(x)$ at $(a, f(a))$. One can view this by magnifying the graph of $f(x)$, which will be closer and closer to a straight line (locally).

3.2. **Basic properties of differentiable functions.** Differentiability is a stronger property than continuity. In fact, the former implies the latter.

Theorem 3.5. Let $f\colon I \to \mathbb{R}$ be a function on an open interval and $x_0 \in I$. If f is differentiable at x_0, then f is continuous at x_0. In other words, if

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} \tag{3.6}$$

exists, then

$$\lim_{x\to x_0} f(x) = f(x_0). \tag{3.7}$$

Proof. We know the limit (3.6) exists. We also know the limit

$$\lim_{x\to x_0} (x-x_0) = 0$$

exists. Thus, we can apply the limit law to obtain

$$\lim_{x\to x_0} (f(x)-f(x_0)) = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} \cdot \lim_{x\to x_0} (x-x_0) = f'(x_0)\cdot 0 = 0,$$

so (3.7) follows. □

Example 3.8. Contrary to the theorem, there are many examples of continuous functions that are not differentiable.

(1) We look at $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$. We know that it is continuous, and now we check whether it is differentiable at $x = 0$. Since $|x|$ is defined by a piecewise function, we check the limit from both sides. We have
$$\lim_{x\to 0^+} \frac{|x|-|0|}{x-0} = \lim_{x\to 0^+} \frac{x}{x} = 1, \text{ while } \lim_{x\to 0^-} \frac{|x|-|0|}{x-0} = \lim_{x\to 0^+} \frac{-x}{x} = -1.$$
Thus, the limit does not exist, and hence f is not differentiable at $x = 0$.

(2) We look at a positive example of the theorem. Consider $f(x) = \begin{cases} 4x-4 & \text{if } x \le 2 \\ x^2 & \text{if } x > 2 \end{cases}$. Then we can find that $f'(x) = \begin{cases} 4 & \text{if } x \le 2 \\ 2x & \text{if } x > 2 \end{cases}$ is still continuous, so f must be continuous at 2. We remark that one can see that f' is not differentiable anymore at $x = 2$.

Another property we will commonly use is the linearity. We have seen this property many times, in both taking limits and continuity.

Proposition 3.9. Let f and g be real-valued functions. If f and g are both differentiable at $x_0 \in I$, then for any $c \in \mathbb{R}$, $f + cg$ is also differentiable at x_0 with derivative $f'(x_0) + cg'(x_0)$.

We remark that this includes the situation of $f + g$, $f - g$, and any multiple $c \cdot g$.

Proof. This follows from $\lim_{x\to x_0} \frac{(f+cg)(x)-(f+cg)(x_0)}{x-x_0} = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} + c \lim_{x\to x_0} \frac{g(x)-g(x_0)}{x-x_0}$ based on the limit law. □

The situation when we multiply a differentiable function by the other is a bit more complicated, and we will look at it later.

3.3. **Polynomials, products, and quotients.** Now, we start looking at the derivative of a general polynomial. In view of the linearity (Proposition 3.9), we essentially only need to look at a monomial x^n. We have seen that when $n = 0$,

$$\frac{d}{dx}(x^0) = \frac{d}{dx}(1) = 0;$$

when $n = 1$,

$$\frac{d}{dx}x = 1;$$

when $n = 2$,

$$\frac{d}{dx}(x^2) = 2x.$$

We now do the general case.

Proposition 3.10. For any $n \in \mathbb{N}$, $\dfrac{d}{dx}(x^n) = nx^{n-1}$.

Proof. We recall the binomial formula, which says that for any $a, b \in \mathbb{R}$, we have

$$\begin{aligned}(a+b)^n &= \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i \\ &= a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \cdots + \binom{n}{n-1} a b^{n-1} + b^n.\end{aligned}$$

With this, we can then compute

$$\begin{aligned}\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} &= \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} \\ &= \lim_{h\to 0} \frac{\left(x^n + \binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \cdots + \binom{n}{n-1}xh^{n-1} + h^n\right) - x^n}{h} \\ &= \lim_{h\to 0} \frac{nx^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \cdots + \binom{n}{n-1}xh^{n-1} + h^n}{h} \\ &= \lim_{h\to 0} \left(nx^{n-1} + \binom{n}{2}x^{n-2}h + \cdots + \binom{n}{n-1}xh^{n-2} + h^{n-1}\right) = nx^{n-1}.\end{aligned}$$

This is true for all $x \in \mathbb{R}$, and we are done. $\square$

Thus, combining Propositions 3.9 and 3.10, we know how to differentiate a general polynomial.

Example 3.11. If we look at $f(x) = 3x^3 - 5x^2 - 3x + 1$, we have

$$f'(x) = 9x^2 - 10x - 3.$$

We now know how to differential a polynomial, e^x, their constant multiples, and their sum. The next question is about their product.

Proposition 3.12. Let f and g be real-valued functions on an open interval I. If f and g are both differentiable at $x \in I$, then fg is also differentiable at x with derivative

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

This is sometimes called the product rule or the Leibniz rule.

Proof. We are going to evaluate

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

To express it in terms of f' and g', note that we can rewrite the numerator as

$$f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x).$$

Then we can calculate

$$\begin{aligned}
&\lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\
=&\lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x)}{h} \\
=&\lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h)}{h} + \lim_{h\to 0} \frac{f(x)g(x+h) - f(x)g(x)}{h} \\
=&\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \cdot \lim_{h\to 0} g(x+h) + f(x) \cdot \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} \\
=&f'(x)g(x) + f(x)g'(x).
\end{aligned}$$

Note that in the final step, we also use the continuity of g at x. □

Example 3.13. We see some examples.

(1) If we look at $f(x) = (3x^2 - 2)e^x$, then

$$f'(x) = 6x \cdot e^x + (3x^2 - 2)e^x = \left(3x^2 + 6x - 2\right) e^x.$$

(2) If we look at $f(x) = e^{2x}$, we can view it as $e^x \cdot e^x$, and hence

$$f'(x) = e^x \cdot e^x + e^x \cdot e^x = 2e^x \cdot e^x = 2e^{2x}.$$

(3) One can inductively prove that

$$\frac{d}{dx} e^{nx} = ne^{nx} \tag{3.14}$$

for any $n \in \mathbb{N}$.

Next, we look at the quotient of two functions.

Proposition 3.15. Let f and g be real-valued functions on an open interval I. If f and g are both differentiable at $x \in I$ and $g(x) \neq 0$, then f/g is also differentiable at x with derivative

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

This is sometimes called the quotient rule.

Proof. We first look at a simpler case when $f = 1$. That is,

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g(x)^2}. \tag{3.16}$$

For this, we need to calculate

$$\begin{aligned}\left(\frac{1}{g}\right)(x+h)-\left(\frac{1}{g}\right)(x) &= \frac{1}{g(x+h)}-\frac{1}{g(x)}\\ &= \frac{g(x)-g(x+h)}{g(x)g(x+h)}.\end{aligned}$$

Since g is particularly continuous at x, we can obtain

$$\begin{aligned}\left(\frac{1}{g}\right)'(x) &= \lim_{h\to 0}\frac{\left(\frac{1}{g}\right)(x+h)-\left(\frac{1}{g}\right)(x)}{h}\\ &= \lim_{h\to 0}\frac{g(x)-g(x+h)}{h\cdot g(x)g(x+h)}\\ &= -\lim_{h\to 0}\frac{g(x+h)-g(x)}{h}\cdot\lim_{h\to 0}\frac{1}{g(x)g(x+h)}\\ &= -g'(x)\cdot\frac{1}{g(x)^2},\end{aligned}$$

and (3.16) follows.

We can now combine (3.16) and the product rule to calculate the derivative of $f/g = f\cdot(1/g)$. In fact,

$$\begin{aligned}\left(\frac{f}{g}\right)'(x) &= \left(f\cdot\frac{1}{g}\right)'(x)\\ &= f'(x)\cdot\frac{1}{g}(x)+f(x)\cdot\left(\frac{1}{g}\right)'(x)\\ &= \frac{f'(x)}{g(x)}-\frac{f(x)g'(x)}{g(x)^2}\\ &= \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}.\end{aligned}$$

This finishes the proof. □

Example 3.17. We see some examples.

(1) We can calculate the derivative of $f(x)=x^{-n}$ for $n\in\mathbb{N}$. By viewing it as $1/x^n$, we get

$$f'(x) = -\frac{nx^{n-1}}{x^{2n}} = -nx^{-n-1}.$$

This tells us that Proposition 3.10 works for any integer n.

(2) We can calculate the derivative of $f(x)=e^{-x}$. By viewing it as $1/e^x$, we get

$$f'(x) = -\frac{e^x}{e^{2x}} = -e^{-x}.$$

(3) With (3.14), we can even get

$$\frac{d}{dx}e^{-nx} = \frac{d}{dx}\left(\frac{1}{e^{nx}}\right) = -\frac{ne^{nx}}{e^{2nx}} = -ne^{-nx} \tag{3.18}$$

for any $n\in\mathbb{N}$. This tells us that (3.14) is valid for any integer n.

3.4. **Trigonometric functions.** Next, we talk about derivatives of $\sin x, \cos x$, and $\tan x$. We recall that we evaluate the limit

$$\lim_{x\to 0} \frac{\sin x}{x} = 1. \tag{3.19}$$

With this, in fact, we already get the derivative of $\sin x$ when $x = 0$. If we let $f(x) = \sin x$, then

$$f'(0) = \lim_{x\to 0} \frac{\sin x - \sin 0}{x - 0} = \lim_{x\to 0} \frac{\sin x}{x} = 1.$$

As we have mentioned, this explains why the graph of $\sin x$ looks almost the same as that of x when we magnify the region near the origin.

Now, we look at the general case when we take the derivative of $f(x) = \sin x$. Before evaluating it formally, first, from the following figure, one can guess that its derivative might be $\cos x$.

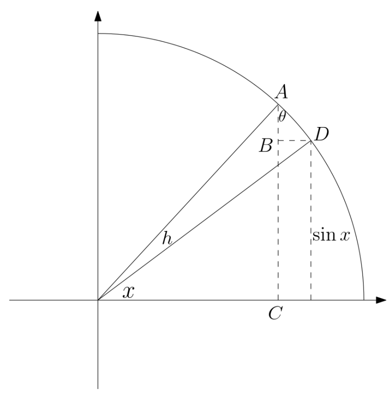

FIGURE 1. $\frac{\sin(x+h)-\sin x}{h} = \frac{\overline{AB}}{\text{arc}AD} \approx \frac{\overline{AB}}{\overline{AD}} = \cos\theta \approx \cos x$ when h is very small. In fact, when h is small, $\theta \approx x$ and $\overline{AD} \approx \text{arc}AD = h$.

We get

$$\begin{aligned} f'(x) = \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} &= \lim_{h\to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \lim_{h\to 0} \left(\cos x \cdot \frac{\sin h}{h} + \sin x \cdot \frac{\cos h - 1}{h} \right). \end{aligned}$$

We know the limit of the first term exists by (3.19). The problem is the second one, which we will evaluate now.

Before doing it rigorously, we observe that the limit also represent a derivative at a certain point. Note that

$$\lim_{h\to 0} \frac{\cos h - 1}{h} = \lim_{h\to 0} \frac{\cos h - \cos 0}{h - 0}$$

is the derivative of $\cos x$ when $x = 0$. If one looks at the graph of $\cos x$, one may then reasonably guess that this limit might be 0, or from Figure 1 again, it might be $\cos'(0) = -\sin 0 = 0$. That is,

$$\lim_{h\to 0} \frac{\cos h - 1}{h} = 0. \tag{3.20}$$

We now evaluate it using (3.19). The idea is that we can connect $\cos h - 1$ to $\sin h$ using the identity $\sin^2 h + \cos^2 h = 1$. With this in mind, we get

$$\begin{aligned}\lim_{h\to 0}\frac{\cos h - 1}{h} &= \lim_{h\to 0}\frac{(\cos h - 1)(\cos h + 1)}{h\,(\cos h + 1)}\\ &= \lim_{h\to 0}\frac{\cos^2 h - 1}{h\,(\cos h + 1)}\\ &= \lim_{h\to 0}\frac{-\sin^2 h}{h\,(\cos h + 1)}\\ &= -\lim_{h\to 0}\left(\frac{\sin h}{h}\cdot\frac{\sin h}{\cos h + 1}\right).\end{aligned}$$

Now, the limits of both of the terms exist, so we can use the limit law of products to get

$$\lim_{h\to 0}\frac{\cos h - 1}{h} = -\lim_{h\to 0}\left(\frac{\sin h}{h}\cdot\frac{\sin h}{\cos h + 1}\right) = -1\cdot\frac{0}{2} = 0,$$

which proves (3.20). Now, with both (3.19) and (3.20), we can differentiate $\sin x$ at a general point.

Theorem 3.21. $\dfrac{d}{dx}\sin x = \cos x$ and $\dfrac{d}{dx}\cos x = -\sin x$.

Proof. Continuing from the calculate at the beginning of the section, we get

$$\begin{aligned}\frac{d}{dx}\sin x = \lim_{h\to 0}\frac{\sin(x+h) - \sin x}{h} &= \lim_{h\to 0}\frac{\sin x\cos h + \cos x\sin h - \sin x}{h}\\ &= \lim_{h\to 0}\left(\cos x\cdot\frac{\sin h}{h} + \sin x\cdot\frac{\cos h - 1}{h}\right)\\ &= \cos x\cdot 1 + \sin x\cdot 0 = \cos x\end{aligned}$$

using (3.19) and (3.20). We can do the same thing to $\cos x$ and get

$$\begin{aligned}\frac{d}{dx}\cos x = \lim_{h\to 0}\frac{\cos(x+h) - \cos x}{h} &= \lim_{h\to 0}\frac{\cos x\cos h - \sin x\sin h - \cos x}{h}\\ &= \lim_{h\to 0}\left(\cos x\cdot\frac{\cos h - 1}{h} - \sin x\cdot\frac{\sin h}{h}\right)\\ &= \cos x\cdot 0 - \sin x\cdot 1 = -\sin x.\end{aligned}$$

This finishes the proof. $\square$

We can notice a pattern when we try to differentiate $\sin x$ more than once. If we let $f(x) = \sin x$, then we get

$$\begin{aligned}f'(x) &= \cos x,\\ f''(x) &= -\sin x,\\ f^{(3)}(x) &= -\cos x,\\ f^{(4)}(x) &= \sin x,\end{aligned}$$

and so on. One can then get $f^{(n)}(x)$ for any $n \in \mathbb{N}$ depending on

Next, we can then use these two derivatives to differentiate $\tan x$.

Corollary 3.22. $\frac{d}{dx}\tan x = \frac{1}{\cos^2 x}$.

Proof. Using the quotient rule, we get

$$\frac{d}{dx}\tan x = \frac{d}{dx}\left(\frac{\sin x}{\cos x}\right) = \frac{\cos x \cdot \cos x - \sin x \cdot (-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x}. \tag{3.23}$$

$\square$

If we introduce the other three trigonometric functions

$$\begin{aligned} \cot x &= \frac{1}{\tan x}, \\ \sec x &= \frac{1}{\cos x}, \text{ and} \\ \csc x &= \frac{1}{\sin x} \end{aligned}$$

when they are well-defined (i.e., having non-zero denominators), then we can simply write

$$\frac{d}{dx}\tan x = \sec^2 x,$$

and one can find the derivatives of $\cot x, \sec x$, and $\csc x$ using the quotient rule.

Example 3.24. We see some examples.

(1) We can now differentiate $\sin 2x = 2\sin x\cos x$.

$$\frac{d}{dx}\sin 2x = 2\frac{d}{dx}(\sin x\cos x) = 2(\cos x\cos x + \sin x(-\sin x)) = 2\cos 2x.$$

(2) Similarly, we can differentiate $\cos 2x = \cos^2 x - \sin^2 x$.

$$\begin{aligned} \frac{d}{dx}\cos 2x &= \frac{d}{dx}(\cos x\cos x - \sin x\sin x) \\ &= -\sin x\cos x + \cos x(-\sin x) - (\cos x\sin x + \sin x\cos x) \\ &= -4\sin x\cos x \\ &= -2\sin 2x. \end{aligned}$$

(3) Combining above, we can do the same thing for $\sin 3x = \sin x\cos 2x + \cos x\sin 2x$.

$$\begin{aligned} \frac{d}{dx}\sin 3x &= \frac{d}{dx}(\sin x\cos 2x + \cos x\sin 2x) \\ &= \cos x\cos 2x + \sin x(-2\sin 2x) + (-\sin x)\sin 2x + \cos x \cdot 2\cos 2x \\ &= 3(\cos x\cos 2x - \sin x\sin 2x) \\ &= 3\cos 3x. \end{aligned}$$

We have seen this pattern many times. That is, once we know how to differentiate $f(x)$, then when we try to differentiate $f(nx)$ for some $n \in \mathbb{N}$, we always get

$$\frac{d}{dx}f(nx) = n \cdot f'(nx). \tag{3.25}$$

We did this when $f(x)$ is e^x, e^{-x} in (3.14) and (3.18), and now we also see this when $f(x)$ is $\sin x$. This is in fact a special case of the chain rule that we will see in the following section.

3.5. **Chain rule.** Many of the functions we deal with come from composition. The chain rule allows us to differentiate the composition of two functions.

Theorem 3.26. Let $f\colon I \to \mathbb{R}$ and $g\colon J \to \mathbb{R}$ be differentiable functions on open intervals I and J. Suppose $g(J) \subseteq I$ so that $f \circ g$ is well-defined. Then for $x \in J$,

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

This is the most important technique in differentiation. With it, we can now basically differentiate any differentiable function we can write down. A good way to think of the chain rule is writing the two functions by $z(y)$ and $y(x)$. Then z can be viewed as a function of x by the composition $z \circ y(x) = z(y(x))$, and the chain rule says

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}.$$

This looks like usual products of two quotients, but of course, it is not how it works. This explains why the notation $\frac{d}{dx}$ is reasonable.

Example 3.27. We will see a lot of examples to understand how it works.

(1) Let $f(x) = e^{3x}$. We can view it as the composition of $z = e^y$ and $y = 3x$, so

$$f'(x) = e^{3x} \cdot 3 = 3e^{3x}.$$

One can similarly derive (3.25) for any differentiable f and $n \in \mathbb{N}$.

(2) Let $f(x) = \cos\left(x^2 + 3x\right)$. Then viewing $z = \cos y$ and $y = x^2 + 3x$, we get

$$f'(x) = -\sin\left(x^2 + 3x\right) \cdot (2x + 3).$$

We can do the chain rule more than once. For example, if $w = w(z)$, $z = z(y)$, and $y = y(x)$ are functions such that $w \circ z \circ y$ makes sense, then

$$\frac{dw}{dx} = \frac{dw}{dz}\frac{dz}{dy}\frac{dy}{dx}.$$

(3) Let $f(x) = e^{\sin 2x^2}$. Then

$$f'(x) = e^{\sin 2x^2} \cdot \cos 2x^2 \cdot 4x.$$

(4) Let $f(x) = \sin^2(2x + \tan e^x)$. Then

$$f'(x) = 2\sin(2x + \tan e^x) \cdot \cos(2x + \tan e^x) \cdot (2 + \sec e^x \cdot e^x).$$

Another application of the chain rule is that we can now differentiate all the good functions we have, including a^x and $\log_a x$. First, we do $\log x$.

Corollary 3.28. $\dfrac{d}{dx}\log x = \dfrac{1}{x}$.

Proof. We use the fact that $\log x$ is the inverse function of e^x. That is, for any $x > 0$, $e^{\log x} = x$. Differentiating both sides leads to

$$e^{\log x} \cdot \frac{d}{dx}\log x = 1.$$

This then implies

$$\frac{d}{dx}\log x = e^{-\log x} = \frac{1}{x},$$

which is what we want. $\square$

As a result, we can differentiate all a^x and $\log_a x$.

Corollary 3.29. Let $a > 0$. Then $\dfrac{d}{dx}a^x = a^x \log a$ and $\dfrac{d}{dx}\log_a x = \dfrac{1}{x \log a}$.

Proof. We can view $a^x = e^{\log a^x} = e^{x \log a}$. Then

$$\frac{d}{dx}a^x = \frac{d}{dx}e^{x\log a} = e^{x \log a} \cdot \log a = a^x \log a.$$

For $\log_a x$, we simply use the relation $\log_a x = \frac{\log x}{\log a}$ and derive

$$\frac{d}{dx}\log_a x = \frac{d}{dx}\left(\frac{\log x}{\log a}\right) = \frac{1}{x \log a}$$

since $\log a$ is just a constant. $\square$

A takeaway is that when we don't know what to do when differentiating a positive function $f(x)$, we can always write it as $e^{\log f(x)}$ and see if that makes things easier.

Corollary 3.28 also tells us another characterization of the number e. That is, we can get

$$e = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n. \tag{3.30}$$

To see this, we look at $f(x) = \log x$ and we know $f'(x) = 1/x$. On the other hand, we know from the definition of derivatives that

$$f'(1) = \lim_{h\to 0}\frac{\log(1+h) - \log 1}{h} = \lim_{h\to 0}\frac{\log(1+h)}{h} = \lim_{h\to 0}\log(1+h)^{\frac{1}{h}}.$$

As above, we know that $f'(1) = 1/1 = 1$. Then we can use the continuity of e^x and $\log(1+h)^{\frac{1}{h}}$ to get

$$e = e^1 = e^{\lim_{h\to 0}\log(1+h)^{\frac{1}{h}}} = \lim_{h\to 0} e^{\log(1+h)^{\frac{1}{h}}} = \lim_{h\to 0}(1+h)^{\frac{1}{h}}.$$

This particularly implies that

$$e = \lim_{h\to 0^+}(1+h)^{\frac{1}{h}},$$

so considering $n = 1/h$ leads to the relation (3.30).

We can also use the chain rule to generalize the power rule given in Proposition 3.10.

Corollary 3.31. For any $n \in \mathbb{R}$, $\dfrac{d}{dx}(x^n) = nx^{n-1}$.

Proof. We can write $x^n = e^{\log x^n} = e^{n \log x}$, so the chain rule implies

$$\frac{d}{dx}x^n = \frac{d}{dx}e^{n\log x} = e^{n \log x}\cdot\frac{n}{x} = x^n \cdot \frac{n}{x} = nx^{n-1}.$$

$\square$

The idea of Corollary 3.28 can be applied to any invertible function. In fact, when we did this for e^x, we should have checked that $\log x$ is differentiable to apply the chain rule. In general, if $f\colon I \to J \subseteq \mathbb{R}$ is invertible and differentiable, and if $f'(x_0) \neq 0$ for some $x_0 \in I$, then f^{-1} will also be differentiable at $f(x_0)$, and then we can use the chain rule to evaluate the derivative of f^{-1} at $f(x_0)$.

Example 3.32. We look at the inverse functions of trigonometric functions.

(1) We look at $f\colon [-1,1] \to [-\pi/2, \pi/2]$ given by $f(x) = \arcsin x$. For $x \in (-1,1)$, we have $\sin f(x) = x$, so the chain rule implies
$$\cos f(x) \cdot f'(x) = 1.$$
Thus, we need to find out $\cos f(x)$. Using the relation $\sin^2\theta + \cos^2\theta = 1$, we know
$$\cos^2 f(x) = 1 - x^2.$$
Since $f(x) \in (-\pi/2, \pi/2)$, we have $\cos f(x) > 0$, so $\cos f(x) = \sqrt{1-x^2}$, and hence
$$f'(x) = \frac{1}{\sqrt{1-x^2}}.$$

(2) We similarly look at $f\colon [-1,1] \to [0,\pi]$ given by $f(x) = \arccos x$. For $x \in (-1,1)$, we have $\cos f(x) = x$, so the chain rule implies
$$-\sin f(x) \cdot f'(x) = 1.$$
Since $f(x) \in (0,\pi)$, $\sin x > 0$ so
$$f'(x) = -\frac{1}{\sin f(x)} = -\frac{1}{\sqrt{1-x^2}}.$$

(3) We then look at $f\colon \mathbb{R} \to (-\pi/2, \pi/2)$ given by $f(x) = \arctan x$. For $x \in \mathbb{R}$, we have $\tan f(x) = x$, so the chain rule implies
$$\sec^2 f(x) \cdot f'(x) = 1.$$
Since $f(x) \in (-\pi/2, \pi/2)$, $\cos f(x) > 0$. Thus,
$$f'(x) = \cos^2 f(x) = \frac{1}{1+x^2}.$$

Finally, we mention some ideas about the proof of the chain rule. We use the "linear approximation," which says that

$$\lim_{h\to 0} g(x+h) = \lim_{h\to 0} \left(g(x) + g'(x)h\right). \tag{3.33}$$

Then the continuity of f "suggests"

$$\lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{h} = \lim_{h\to 0} \frac{f\left(g(x) + g'(x)h\right) - f(g(x))}{h}. \tag{3.34}$$

If $g'(x) \neq 0$, then we can consider $H := g'(x)h$ and derive
$$\lim_{h\to 0} \frac{f\left(g(x) + g'(x)h\right) - f(g(x))}{h} = \lim_{H\to 0} \frac{f\left(g(x) + H\right) - f(g(x))}{H} \cdot g'(x) = f'(g(x)) \cdot g'(x).$$
If $g'(x) = 0$, then (3.34) is immediately zero. Thus, we are done.

The proof above is not completely rigorous. Later, we will use a stronger form of linear approximation (compared with (3.33)) to give a correct proof.

3.6. **Other methods of differentiation.** The chain rule essentially allows us to differentiate many differentiable functions. We will talk about two more consequences of it.

The first one is the **logarithmic differentiation**. In fact, we have seen this trick when differentiate a^x and x^n. In general, if $f(x)$ is a positive differentiable function, then we can differentiate it by

$$\frac{d}{dx}f(x) = \frac{d}{dx}e^{\log f(x)} = e^{\log f(x)} \cdot \frac{d}{dx}\log f(x).$$

If f involves many products or some exponents, taking logarithm usually makes the differentiation simpler.

Example 3.35. We see some examples.

(1) Suppose f, g, and h are positive differentiable functions on I and let $F = fgh$. Then

$$\begin{aligned}\frac{d}{dx}F(x) = \frac{d}{dx}e^{\log F(x)} &= e^{\log F(x)}\frac{d}{dx}\log\left(f(x)g(x)h(x)\right) \\ &= F(x)\cdot\frac{d}{dx}\left(\log f(x) + \log g(x) + \log h(x)\right) \\ &= f(x)g(x)h(x)\left(\frac{f'(x)}{f(x)} + \frac{g'(x)}{g(x)} + \frac{h'(x)}{h(x)}\right) \\ &= f'(x)g(x)h(x) + f(x)g'(x)h(x) + f(x)g(x)h'(x).\end{aligned}$$

One can use the same technique to write down a general product rule that works for a large product.

(2) The idea above also works for quotients. For example, using the same positive functions f, g, and h, we can get

$$\frac{d}{dx}\left(\frac{fg}{h}\right) = \frac{d}{dx}e^{\log f + \log g - \log h} = \frac{fg}{h}\cdot\left(\frac{f'}{f} + \frac{g'}{g} - \frac{h'}{h}\right).$$

(3) Consider $f(x) = x^x$ for $x > 0$. Then

$$\frac{d}{dx}f(x) = \frac{d}{dx}e^{x\log x} = e^{x\log x}\cdot\frac{d}{dx}(x\log x) = x^x\cdot\left(\log x + x\cdot\frac{1}{x}\right) = x^x\log x + x^x.$$

(4) Consider $f(x) = x^{\sin x}$ for $x > 0$. Then

$$\frac{d}{dx}f(x) = \frac{d}{dx}e^{\sin x\cdot\log x} = x^{\sin x}\cdot\left(\cos x\cdot\log x + \sin x\cdot\frac{1}{x}\right).$$

The next technique we are going to learn is the **implicit differentiation**. It is especially useful when we want to find the tangent to a curve that is not the graph of a single function. Generally speaking, the idea is that if an equation of functions holds, then we can differentiate the functions on both sides. If the two functions are the same on their domains, then their derivatives are also the same.

Example 3.36. We will see how the implicit differentiation works in the following example.

(1) Suppose the function $f\colon (0,1) \to (0,1)$ satisfies

$$x^2 + f(x)^2 = 1 \tag{3.37}$$

for all $x \in (0,1)$. This essentially means that the graph of $f(x)$ is part of the curve defined by

$$x^2 + y^2 = 1.$$

If we want to evaluate $f'(1/2)$, one way is to write

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

and then differentiate. Another way is to start differentiating (3.37) directly, which leads to

$$2x + 2f(x) \cdot f'(x) = 0.$$

Thus, at $x = 1/2$, we get

$$2 \cdot \frac{1}{2} + 2f\left(\frac{1}{2}\right) \cdot f'\left(\frac{1}{2}\right) = 0,$$

which implies

$$f'\left(\frac{1}{2}\right) = -\frac{1}{2f(1/2)}.$$

We can then go back to (3.37) again to find

$$\left(\frac{1}{2}\right)^2 + f\left(\frac{1}{2}\right)^2 = 1,$$

which implies $f(1/2) = \sqrt{3}/2$ since $f(x) > 0$. Thus, we can conclude

$$f'\left(\frac{1}{2}\right) = -\frac{1}{\sqrt{3}}.$$

Note that this tells us the tangent to the unit circle at $(1/2, \sqrt{3}/2)$ has slope $-1/\sqrt{3}$.

(2) In the first example, even if we don't like this indirect method, we can still try to write down an explicit formula of $f(x)$ and differentiate it. In some cases, this direct method does not work, and hence the implicit differentiation might be necessary. For example, suppose a differentiable function f satisfies

$$x^3 \sin f(x) + f(x) = x^2 - 3x. \tag{3.38}$$

That is, the graph of f is part of the curve defined by

$$x^3 \sin y + y = x^2 - 3x.$$

Thus, differentiating both sides of (3.38) leads to

$$3x^2 \sin f(x) + x^3 \cos f(x) \cdot f'(x) + f'(x) = 2x - 3,$$

which implies

$$f'(x) = \frac{2x - 3 - 3x^2 \sin f(x)}{1 + x^3 \cos f(x)}.$$

Now given any x, we can use $x, f(x)$, and this formula to calculate $f'(x)$. Note that it is not easy to figure out an explicit formula of $f(x)$ in terms of x using (3.38).

(3) We look at the curve defined by

$$y^3 + x^3 - 6xy = 0.$$

The point $(3,3)$ is on this curve. We can use the technique to find out the tangent to this curve at $(3,3)$. If near $(3,3)$ the curve is locally the graph of a function f, then the function f satisfies

$$f(x)^3 + x^3 - 6xf(x) = 0.$$

Thus, we can differentiate it to get

$$3f(x)^2 f'(x) + 3x^2 - 6\big(f(x) + xf'(x)\big) = 0,$$

so

$$f'(x) = \frac{6f(x) - 3x^2}{3f(x)^2 - 6x}.$$

Thus, at $(3,3)$, we get

$$f'(3) = \frac{18 - 27}{27 - 18} = -1.$$

As a result, the tangent can be described by

$$y = -(x-3) + 3 = -x + 6.$$

We end this section with some discussion on differentiable functions and continuous functions.

Remark 3.39. We know that differentiability is a stronger property than continuity. In fact, there are "much more" continuous functions than differentiable functions. In a way, we can "measure" the set of differentiable functions inside the (larger) set of continuous functions and reasonably show that it is a small set. Making this idea precise requires more advanced tools from real analysis, such as topology and measure theory. In fact, one can show that, in a precise mathematical sense, "most" continuous functions are nowhere differentiable. Although this result is far beyond the scope of Calculus I, it illustrates an important theme: differentiability is a very restrictive and delicate property. Continuity is common; smoothness is special.

Remark 3.40. One can in fact write down a continuous function that is **nowhere differentiable** on $\mathbb{R}$. Weierstrass constructed an example of the form

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n x)$$

for some carefully chosen positive numbers a and b. The idea is to build such a function by adding "infinitely many tiny waves" (given by cosine in this case). Each wave is smaller than the previous one, so the function stays continuous. But each wave oscillates faster and faster. The slopes become so wild that no tangent line exists anywhere.

For a long time, people mostly studied functions coming from geometry and physics, which were very smooth. Thus, it was widely expected that continuity should almost always imply differentiability, possibly except at a few exceptional points (like the absolute value function). The discovery of continuous nowhere differentiable functions was therefore surprising.