Mathematical Analysis

Author: Isaac FEI

Contents

Chapter 1 Limits and Continuity

1.1 Limit of a Function

> **Proposition 1.1.1**
>
> *Let f and g be complex-valued functions defined on a subset A of a metric space (X, d), and p a limit point of A. If $\lim_{x \to p} f(x) = 0$ and g is bounded, then*
> $$\lim_{x \to p} f(x)g(x) = 0$$

Proof ■

1.2 Continuous Functions

The next theorem states that the composite function of continuous functions is also continuous.

> **Theorem 1.2.1**
>
> *Let (X, d_X), (Y, d_Y) and (Z, d_Z) be metric spaces. Let $f : X \to Y$ and $g : Y \to Z$ be functions, and let $h = g \circ f$ be the composite function defined on X. If f is continuous at $p \in X$, and g is continuous at $f(p)$, then h is continuous at p.*

Proof Given $\varepsilon > 0$, since g is continuous at $f(p)$, there exists some $\delta' > 0$ such that
$$d_Y(y, f(p)) < \delta' \implies d_Z(g(y), g(f(p))) < \varepsilon \tag{1.1}$$
And since f is continuous at p, there exists some $\delta > 0$ such that
$$d_X(x, p) < \delta \implies d_Y(f(x), f(p)) < \delta' \tag{1.2}$$
Now, choosing $x \in X$ satisfying $d_X(x, p) < \delta$, and then applying (1.2) followed by (1.1), we have
$$d_Z(g(f(x)), g(f(p))) < \varepsilon$$
This is precisely $d_Z(h(x), h(p)) < \varepsilon$. Hence, h is continuous at p. ■

> **Theorem 1.2.2**
>
> *Let $f : X \to \mathbb{R}$ be a real-valued continuous function from a metric space X to $\mathbb{R}$. Let $K \subset X$ be a compact subset. Then there exist points $p, q \in K$ such that*
> $$f(p) = \sup f(K) \quad and \quad f(q) = \inf f(K)$$

1.3 Connectedness

> **Theorem 1.3.1 (Intermediate Value Theorem)**
>
> *Let $f : [a, b] \to \mathbb{R}$ be a real-valued continuous function. Suppose that $f(a) \neq f(b)$, and y is some number in between $f(a)$ and $f(b)$. Then there exists an interior point $c \in (a, b)$ such that $f(c) = y$.* ♡

Proof ∎

1.4 Monotonic Functions

> **Theorem 1.4.1**
>
> *If f is increasing on $[a, b]$, then $f(c+)$ and $f(c-)$ both exists for every $c \in (a, b)$, and we have*
> $$f(c-) \leq f(c) \leq f(c+)$$
> *As the endpoints, we have $f(a+)$ and $f(b-)$ both exists and*
> $$f(a) \leq f(a+) \quad and \quad f(b-) \leq f(b) \tag{1.3}$$
> *Analogous results exist for f being a decreasing function.* ♡

Proof We only prove this theorem for increasing functions. Let $c \in (a, b)$ be an interior point.

(Existence of $f(c-)$) Define a set
$$A = \{f(x) \mid a \leq x < c\}$$
Note that A is bounded above by $f(c)$ since f is increasing. Then A has the least upper bound, say $\alpha = \sup A$. We want to show that $f(c-)$ exists and equals α. Given $\varepsilon > 0$, by the property of α, there exists $x_\varepsilon = (c - \delta) \in [a, c)$ such that (δ depends on ε)
$$f(c - \delta) > \alpha - \varepsilon$$
Because f is increasing, we have
$$\alpha \geq f(x) \geq f(c - \delta) > \alpha - \varepsilon \quad \forall x \in (c - \delta, c)$$
This implies that
$$|f(x) - \alpha| < \varepsilon \quad \forall x \in (c - \delta, c)$$
Hence, α is the left-hand limit of f at $x = c$, i.e., $f(c-) = \alpha$. Recall A is bounded above by $f(c)$. We have $\alpha \leq f(c)$, i.e.,
$$f(c-) \leq f(c)$$

(Existence of $f(c+)$) Define
$$B = \{f(x) \mid c < x \leq b\}$$
Let β be the greatest lower bound of B. A similar argument shows that $f(c+)$, and $f(c+) = \beta$ with
$$f(c) \leq f(c+)$$

(Existence of $f(a+)$ and $f(b-)$) The existence of $f(a+)$ and $f(b-)$ along with (1.3) can be proved by the exact argument as before. ∎

Chapter 2   Differentiation

Let our exploration begins with real-valued functions defined on real numbers.

2.1 Definition of Derivative

Observe the slopes of secant lines (displayed as dashed lines) of the curve shown in Figure 2.1
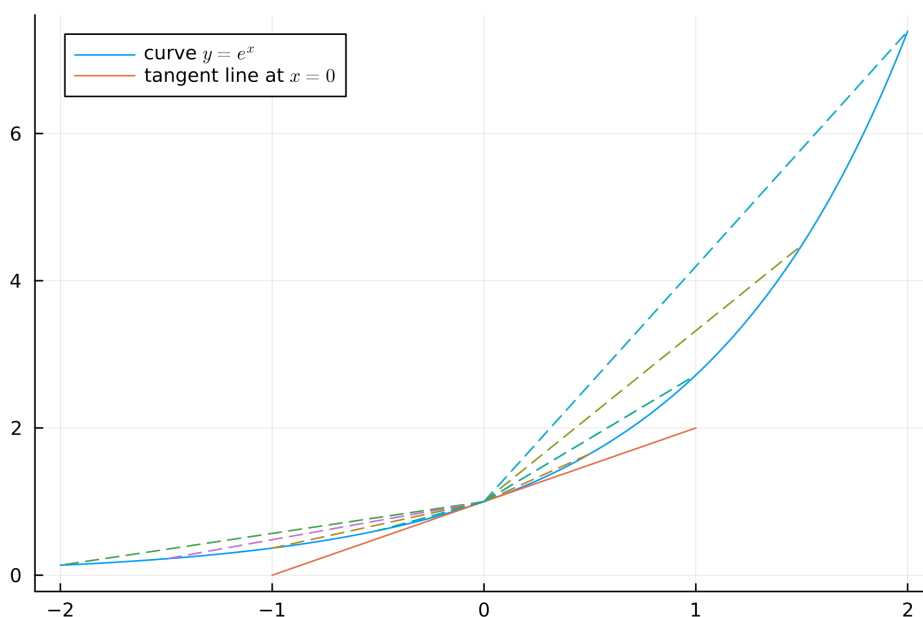

Figure 2.1: Secant lines and a tangent line of curve $y = e^x$.

All secant lines have a common point of intersection, $(0, 1)$. And as the other intersection point gets closer and closer to $(0, 1)$, the slope of the secant line tends to converge to some number. This number, the limiting slope, is precisely the **derivative** of this curve/function at $x = 0$, which we shall soon define. We can draw a straight line passing through the point $(0, 1)$ with the limiting slope. This line is then called the tangent of the curve at point $(0, 1)$, as illustrated in Figure 2.1.

> **Definition 2.1.1**
>
> *Let f be defined on an open interval (a, b), and let $c \in (a, b)$ an interior point. Then f is said to be **differentiable** at c if the limit*
> $$\lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$
> *exists. This limit, denoted by $f'(c)$, is called the **derivative** of f at point c.* ♣

If $f'(c)$ exists $\forall c \in I$ for some interval I, then we can define a function f' on I, which is called the derivative of f. Here the word *derivative* means a function instead of just a number.

The definition of derivatives we present here involves the limit obtained by letting one point approach the other one, $x \to c$. For computational convenience, we can treat the derivative of f at c as

the limit of the fraction $[f(c + h) - f(c)]/h$ as $h \to 0$. That is, the limit is reached when the distance between points x and c, $h = |x - c|$, is small. The following proposition states this idea formally.

Proposition 2.1.1

Let f be defined on (a, b). Then f is differentiable at $c \in (a, b)$ if and only if

$$\lim_{h \to 0} \frac{f(c + h) - f(c)}{h} \tag{2.1}$$

exists. In that case, $f'(c) = \lim_{h \to 0}[f(c + h) - f(c)]/h$.

Proof Suppose f is differentiable at c. For an arbitrary $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|x - c| < \delta \implies \left| \frac{f(x) - f(c)}{x - c} - f'(c) \right| < \varepsilon$$

Let number h be such that $|h| < \delta$. Then since $|(c + h) - c| = |h| < \delta$, we have

$$\left| \frac{f(c + h) - f(h)}{h} - f'(c) \right| = \left| \frac{f(c + h) - f(h)}{(c + h) - c} - f'(c) \right| < \varepsilon$$

This implies the limit (2.1) exists, and it equals $f'(c)$.

Conversely, suppose the limit (2.1) exists, say it equals L. Then for $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|h| < \delta \implies \left| \frac{f(c + h) - f(c)}{h} - L \right| < \varepsilon$$

Choose x such that $|x - c| < \delta$, we have

$$\left| \frac{f(x) - f(c)}{x - c} - L \right| = \left| \frac{f(c + (x - c)) - f(c)}{x - c} - L \right| < \varepsilon$$

Therefore, the limit $[f(x) - f(c)]/(x - c)$ exists, which is equal to L. By the definition of derivatives, f is differentiable at c, and $f'(c) = L$. ∎

Exercise 2.1 Calculate the derivative of e^x at $x = 0$.

[Hint: You may use the fact $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots$.]

Exercise 2.2 Calculate the derivative of

$$f(x) = x^2 \sin \frac{1}{x^2} \mathbb{1}\{x \neq 0\} = \begin{cases} x^2 \sin \frac{1}{x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

at $x = 0$.

Solution *We have*

$$\frac{f(0 + h) - f(0)}{h} = h \sin \frac{1}{h^2}$$

Note the limit of the above expression is zero as $h \to 0$ by Proposition 1.1.1 since $\lim_{h \to 0} h = 0$ and $\sin(1/h^2)$ is bounded by 1. Hence, $f'(0) = 0$ by Proposition 2.1.1.

The graph of $f(x)$ is shown in Figure 2.2. As we can see, $f(x)$ is clamped between curves $y = x^2$ and $y = -x^2$ with the same tangents $y = 0$ at $x = 0$. And we have shown $f'(0) = 0$, which means that the tangent of $f(x)$ is also $y = 0$ at $x = 0$. What conjecture can you make?

Moreover, we observe that $f(x)$ crosses its tangent line at $x = 0$ infinitely many times. This example shows that the tangent line does not have to touch the curve only at one point.
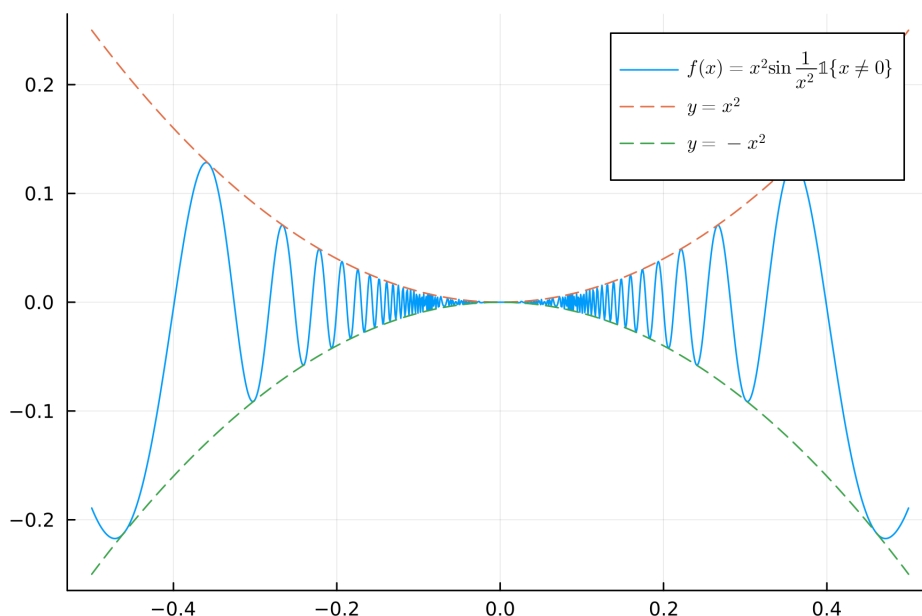
Figure 2.2: Graph of $f(x) = x^2 \sin \frac{1}{x^2} \mathbb{1}\{x \neq 0\}$.

2.1.1 One-Sided Derivatives and Infinite Derivatives

So far, the point at which the derivative is defined has to be an *interior* point. Sometimes, we are required to consider the derivatives at the endpoints of the interval. For example, as we shall see in more detail in Chapter 4, we need to take the derivative of function $F(x) = \int_a^x f(t)\mathrm{d}t$ at $x = a$. Hence, we are motivated to define the **one-sided derivatives**.

If we consider the derivative of f, f' as a function, then $f'(x)$ may tend to infinity as x approaches an endpoint. In addition, sometimes we need to interpret the meaning of vertical tangent lines. This leads to the definition of **infinite derivatives**.

> **Definition 2.1.2**
>
> *Let f be defined on a closed interval I. Suppose f is continuous at point $c \in I$. Then f is said to have a **right-hand derivative** at c if the right-hand limit*
> $$\lim_{x \to c^+} \frac{f(x) - f(c)}{x - c}$$
> *exists as a finite number, or the limit is ∞ or $-\infty$. This right-hand derivative shall be denoted by $f'_+(c)$. The **left-hand derivative**, denoted by $f'_-(c)$, is defined similarly. In addition, if c is an interior point, and $f'_+(c) = f'_-(c) = \infty$, then we write $f'(c) = \infty$. $f'(c) = -\infty$ is similarly defined.* ♣

Remark Note that though we write $f'(c) = \infty$ or $f'(c) = -\infty$, we do not say f is differentiable there.

It is not very intuitive to imagine, at first thought, a continuous function having infinite derivatives. But there are quite a lot of such examples.

Example 2.1 The first example is constructed by cutting a circle in the middle. If we cut a circle in half and stick the lower half to the right of the upper one, then we have a function with a vertical

tangent line in the middle. The explicit expression of this function is

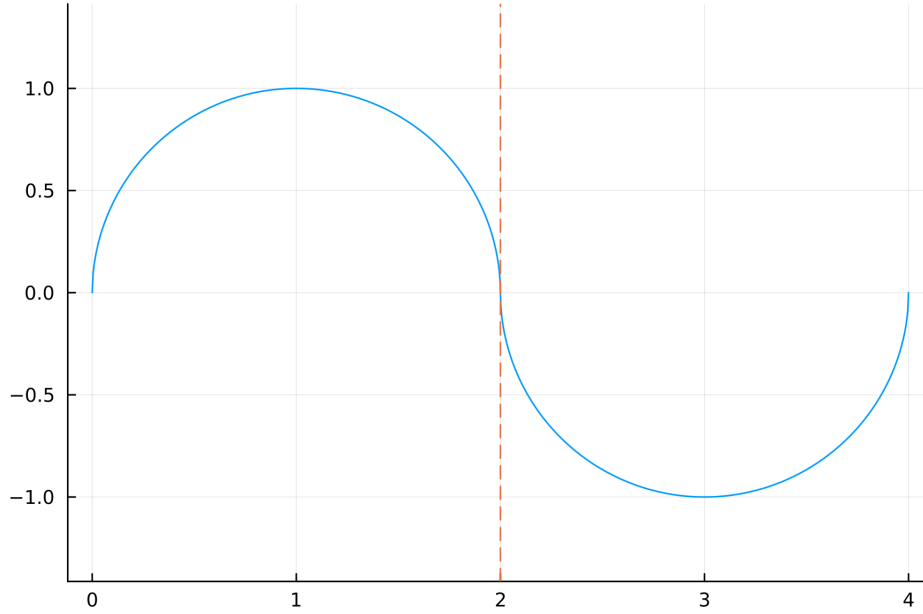$$f(x) = \begin{cases} \sqrt{1 - (x-1)^2} & 0 \le x \le 2 \\ -\sqrt{1 - (x-3)^2} & 2 < x \le 4 \end{cases}$$

Figure 2.3: Graph of $f(x) = \sqrt{1 - (x-1)^2}\mathbb{1}\{0 \le x \le 2\} - \sqrt{1 - (x-3)^2}\mathbb{1}\{2 < x \le 4\}$.

The function is continuous and one can show that $f'(2) = -\infty$.

Example 2.2 The following next example includes both positive and negative intuitive derivatives, and a point at which the right and left-hand side derivatives are ∞ and $-\infty$, respectively. Consider the function

$$f(x) = \sqrt[3]{x^2(x-1)(x-2)}$$

Its graph is shown in Figure 2.4.

It is an exercise to show $f'(1) = -\infty$ and $f'(2) = \infty$. We now consider the one-sided derivatives at $x = 0$. We have

$$\frac{f(x) - f(0)}{x - 0} = \sqrt[3]{\frac{(x-1)(x-2)}{x}}$$

Letting $x \to 0^+$ leads to $f'_+(0) = \infty$, while $x \to 0^-$ yields $f'_-(0) = -\infty$. Hence, we say the derivative of f does not exist at $x = 0$.

2.2 Derivatives and Continuity

The next theorem helps to reduce some theorems on derivatives to theorems on continuity.

Theorem 2.2.1

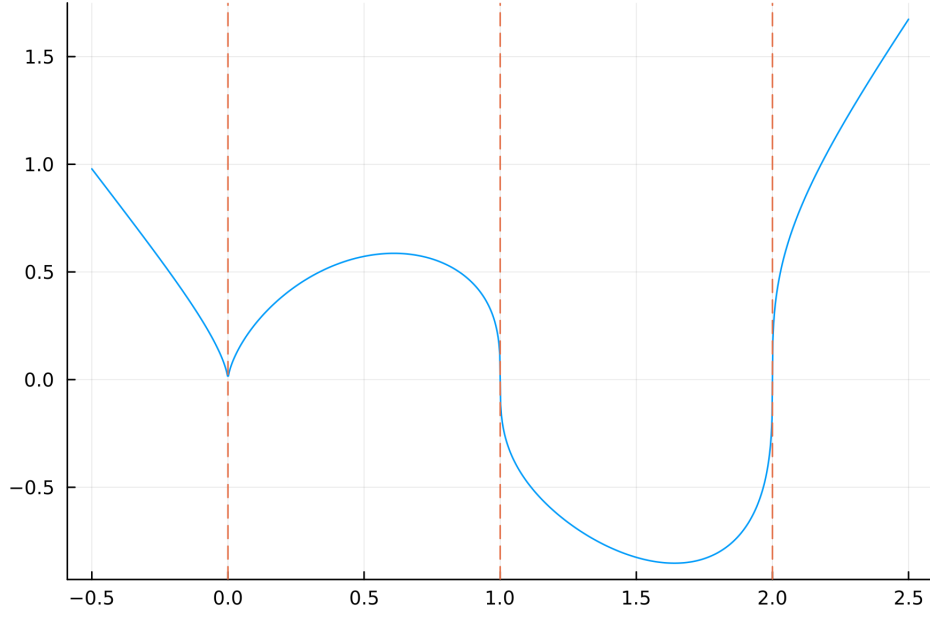*Let f be a function defined on (a, b), and $c \in (a, b)$ a fixed point in that interval. We have the following statements:*

Figure 2.4: Graph of $f(x) = \sqrt[3]{x^2(x-1)(x-2)}$.

> *1. If f is differentiable at c, then there exists a function ϕ on (a, b) which is continuous at c, and satisfies*
>
> $$f(x) - f(c) = (x - c)\phi(x) \quad \forall x \in (a, b) \tag{2.2}$$
>
> *And $f'(c) = \phi(c)$.*
>
> *2. Conversely, if there exists a function ϕ on (a, b) which is continuous at c, and satisfies (2.2), then f is differentiable at c with $f'(c) = \phi(c)$.*

The function ϕ is precisely the slope of function f's secant line.

Proof (Proof of 1) Suppose f is differentiable at c. Define

$$\phi(x) = \frac{f(x) - f(c)}{x - c}, \quad x \neq c \quad \text{and} \quad \phi(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

Note that $\phi(c)$ is well-defined since the limit (which equals $f'(c)$) indeed exists by the definition of the derivative of f at c. Notice also that by defining function ϕ in this way, it is automatically continuous at c. And of course, the equation (2.2) holds.

(Proof of 2) It follows from (2.2) that

$$\phi(x) = \frac{f(x) - f(c)}{x - c} \quad \forall x \neq c$$

Since ϕ is continuous at c, by sending $x \to c$, we have

$$\phi(c) = \lim_{x \to c} \phi(x) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

This implies f is differentiable at c, and $f'(c) = \phi(c)$. $\blacksquare$

With the help of Theorem 2.2.1, we can easily show that f must be continuous at c if it is differentiable there.

> **Theorem 2.2.2**
>
> *Let f be defined on (a,b). If f is differentiable at $c \in (a,b)$, then it is continuous at c.*

Proof   By Theorem 2.2.1, there exists a function ϕ on (a,b) such that ϕ is continuous at c, and

$$f(x) - f(c) = (x-c)\phi(x) \quad \forall x \in (a,b) \tag{2.3}$$

Since the limit of the right-hand side in (2.3) exists as $x \to c$, the limit of the left-hand side also exists, and

$$\lim_{x \to c} f(x) - f(c) = \lim_{x \to c}(x-c)\phi(x) = 0$$

This implies

$$\lim_{x \to c} f(x) = f(c)$$

Hence, f is continuous at c.   ∎

2.3 Computing Derivatives

It is difficult in practice to compute the derivative of any function only using the definition. In this section, we shall introduce some theorems and standard rules that are helpful with computation. If one knows the derivatives of elementary functions (e.g., x^a, e^x, $\ln x$, $\sin x$, $\cos x$, etc.), then by applying these theorems, one should be able to compute the derivatives of any functions that are built up out of elementary functions using summations, differences, products, quotients, and compositions.

Note   *In fact, these so-called elementary functions are not elementary at all. We shall present their definitions as well as their derivatives rigorously in later chapters.*

2.3.1 Algebra of Derivatives

> **Theorem 2.3.1**
>
> *Suppose f and g are defined on (a,b) and are both differentiable at $c \in (a,b)$. Then the functions $f+g$, $f-g$, fg are also differentiable at c. If $g(c) \neq 0$, then the quotient f/g is also differentiable at c. And the formulas of their derivatives are given by*
>   1. $(f \pm g)'(c) = f'(c) \pm g'(c)$,
>   2. $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$,
>   3. $(f/g)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{g^2(c)}$, *provided $g(c) \neq 0$.*

Remark One might be worried about the possibility that the quotient f/g may only be defined at point c if we only require $g(c) \neq 0$. Then how can we say f/g is differentiable at some point? However, the condition that g is differentiable at c implicitly implies that $g(x) \neq 0$ for x in some interval about c if we assume $g(c) \neq 0$. Take a few seconds and think about why this is true. Note that g is continuous at c by Theorem 2.2.2. Then, if $g(c) \neq 0$, by the continuity, g is also nonzero in some neighborhood of c. Hence, in this case, f/g is actually well-defined in an interval instead of at just one point.

The proof we present here is done by exploiting Theorem 2.2.1, and it may differ from a traditional proof that is solely based on the definition.

Proof   Since f and g are differentiable at c. It follows from Theorem 2.2.1 that there exists functions ϕ_1 and ϕ_2 on (a, b) that are continuous at c, and satisfy

$$f(x) - f(c) = (x - c)\phi_1'(x) \tag{2.4}$$

$$g(x) - g(c) = (x - c)\phi_2'(x) \tag{2.5}$$

with

$$\phi_1(c) = f'(c) \quad \text{and} \quad \phi_2(c) = g'(c) \tag{2.6}$$

(Proof of 1) Applying (2.4) and (2.5) to $(f \pm g)(x) - (f \pm g)(c)$, we have

$$(f \pm g)(x) - (f \pm g)(c) = (x - c)(\phi_1(x) \pm \phi_2(x))$$

Since $\phi_1(x) \pm \phi_2(x)$ is continuous at c, $(f \pm g)$ is differentiable at c by Theorem 2.2.1. Then by applying (2.6), we obtain

$$(f \pm g)'(c) = \phi_1(c) \pm \phi_2(c) = f'(c) \pm g'(c)$$

(Proof of 2) Again by applying (2.4) and (2.5), we have

$$(fg)(x) - (fg)(c) = (x - c)[\phi_1(x)g(c) + f(c)\phi_2(x) + (x - c)\phi_1(x)\phi_2(x)]$$

after a few steps of algebraic calculation. It then follows from Theorem 2.2.1 that fg is differentiable at c since the function

$$\phi_1(x)g(c) + f(c)\phi_2(x) + (x - c)\phi_1(x)\phi_2(x)$$

is continuous at c. And

$$(fg)'(c) = \phi_1(c)g(c) + f(c)\phi_2(c) + (c - c)\phi_1(c)\phi_2(c) = f'(c)g(c) + f(c)g'(c)$$

(Proof of 3) Note that the quotient (f/g) can be regarded as a product of f and $1/g$. Hence, we are going to show $1/g$ is differentiable at c, and then apply statement 2, which we have just proved, to conclude the proof. As explained in the remark of this theorem, $1/g$ is well-defined in a neighborhood of c. It follows from (2.5) that

$$\frac{1}{g(x)} - \frac{1}{g(c)} = (x - c)\frac{-\phi_2(x)}{g(x)g(c)}$$

Since $-\phi_2(x)/[g(x)g(c)]$ is continuous at c, we conclude from Theorem 2.2.1 that $1/g$ is differentiable at c, and

$$\left(\frac{1}{g}\right)'(c) = \frac{-\phi_2(c)}{g(c)g(c)} = \frac{-g'(c)}{g^2(c)}$$

Then by applying statement 2, we conclude that f/g is also differentiable at c with

$$(f/g)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{g^2(c)}$$

∎

2.3.2 The Chain Rule

Taking compositions is a more fundamental and natural way of combining two functions apart from summations, products, etc. The next result, the **chain rule**, provides a method of computing the derivative of a composite function.

> **Theorem 2.3.2**
>
> *Let f be a function defined on an open interval I, and g a function defined on $f(I)$. We consider the composite function $g \circ f$. If f is differentiable at $c \in I$, $f(c)$ is an interior of $f(I)$, and g is differentiable at $f(c)$, then $g \circ f$ is differentiable at c with*
>
> $$(g \circ f)'(c) = g'(f(c))f'(c) \tag{2.7}$$

With the help of Theorem 2.2.1, we can reduce the proof of this theorem to that of Theorem 1.2.1.

Proof Since f is differentiable at c, using Theorem 2.2.1, we have

$$f(x) - f(c) = (x - c)\phi_1(x) \quad \forall x \in I \tag{2.8}$$

where ϕ_1 is some function that is continuous at c with $\phi_1(c) = f'(c)$. Using the same argument, since g is differentiable at $f(c)$, we have

$$g(y) - g(f(c)) = (y - f(c))\phi_2(y) \quad \forall y \in f(I) \tag{2.9}$$

where ϕ_2 is some function that is continuous at $f(c)$ with $\phi_2(c) = g'(f(c))$. Replace y with $f(x)$ in (2.9) yields

$$g(f(x)) - g(f(c)) = (f(x) - f(c))\phi_2(f(x)) \quad \forall x \in I$$

Then by replacing $f(x) - f(c)$ with the right-hand side of (2.8), we have

$$g(f(x)) - g(f(c)) = (x - c)\phi_1(x)\phi_2(f(x)) \quad \forall x \in I$$

Note that $\phi_2(f(x))$ is continuous at c by Theorem 1.2.1. And the product $\phi_1(x)\phi_2(f(x))$ is also continuous. This implies $f \circ f$ is differentiable at c by Theorem 2.2.1. And its derivative at $x = c$ equals

$$(g \circ f)'(c) = \phi_1(c)\phi_2(f(c)) = f'(c)g'(f(c))$$

which is exactly (2.7). ∎

✍ **Exercise 2.3** Calculate the derivative of

$$g(x) = x \sin \frac{1}{x} \mathbb{1}\{x \neq 0\}$$

for $x \neq 0$. Does the derivative exist at $x = 0$?

✍ **Exercise 2.4** Calculate the derivative of

$$f(x) = x^2 \sin \frac{1}{x^2} \mathbb{1}\{x \neq 0\}$$

for all $x \in \mathbb{R}$. Recall we have already calculated its derivative at $x = 0$ in Exercise 2.2, which is $f'(0) = 0$.

[Hint: Make use of the result in Exercise 2.3 and the chain rule (Theorem 2.3.2).]

2.4 The Mean Value Theorem

2.4.1 Local Extrema

One may be familiar with the fact that if the derivative of function f is positive at some point, i.e., $f'(c) > 0$, then f is increasing near c, and if $f'(c) < 0$, it is decreasing. However, we can extend this result a little further for we have defined one-sided and infinite derivatives.

> **Theorem 2.4.1**
>
> *Let f be defined on a subset $A \subset \mathbb{R}$ and $c \in A$. If $f'_+(c) > 0$, possibly $f'_+(c) = \infty$, then there exists a half-open ball $(c, c + \delta) \subset A$ in which*
>
> $$x > c \implies f(x) > f(c) \tag{2.10}$$
>
> *Similarly, if $f'_-(c) > 0$, then $\exists (c - \delta, c) \subset A$ in which*
>
> $$x < c \implies f(x) < f(c)$$
>
> *Analogous results hold if we instead assume $f'_+(c) < 0$ or $f'_-(c) < 0$.* ♡

Proof We only prove (2.10) since the proofs of the rest of the results are similar. First suppose $f'_+(c) < \infty$. Because the right-hand limit exists, and it equals $f'_+(c)$, there exists $\delta > 0$ such that

$$\left| \frac{f(x) - f(c)}{x - c} - f'_+(c) \right| < \frac{f'_+(c)}{2} \quad \forall x \in (c, c + \delta)$$

This implies

$$\frac{f(x) - f(c)}{x - c} > \frac{f'_+(c)}{2} > 0 \quad \forall x \in (c, c + \delta)$$

Then $f(x) > f(c)$ provided $x > c$ since they have the same sign.

We also need to consider $f'_+(c) = \infty$. Because the limit is infinity, there exists $\delta > 0$ such that the following fraction exceeds any predetermined number, in particular,

$$\frac{f(x) - f(c)}{x - c} > 0 \quad \forall x \in (c, c + \delta)$$

One can instead choose any large number (non-negative of course, to prove this theorem). This completes the proof since both the numerator and denominator share the same sign. ∎

The following corollary is the version that is more familiar to the reader and is more common in practice.

> **Corollary 2.4.1**
>
> *Let f be defined on (a, b), and $c \in (a, b)$ an interior point. If $f'(c) > 0$, possibly $f'(c) = \infty$, then there exists an open ball $B_\delta(c) \subset (a, b)$ in which*
>
> $$x > c \implies f(x) > f(c) \quad and \quad x < c \implies f(x) < f(c) \tag{2.11}$$
>
> *An analogous result holds if $f'(c) < 0$.* ♡

✎ **Exercise 2.5** Prove Corollary 2.4.1 assuming $f'(c) < \infty$ using Theorem 2.2.1.

Solution *It follows from Theorem 2.2.1 that*

$$f(x) - f(c) = (x - c)\phi(x) \quad \forall x \in (a, b) \tag{2.12}$$

where $\phi(x)$ is continuous at c with $\phi(c) = f'(c) > 0$. Since $\phi(c) > 0$ and it is continuous there, there exists some neighborhood $B_\delta(c) \subset (a, b)$ such that $\phi(x) > 0 \ \forall x \in B_\delta(c)$. And then (2.11) follows from (2.12).

The **local extremum** of a function is the largest or smallest value within some *open ball*.

> **Definition 2.4.1**
>
> *Let f be a real-valued function defined on a subset A of a metric space (X, d), and suppose $p \in A$. Then f is said to have a **local maximum** at p if*
>
> $$f(x) \leq f(p) \quad \forall x \in B_\delta(p) \cap A$$
>
> *for some open ball $B_\delta(p)$. An analogous definition exists for **local minimum**.* ♣

The reader should be very familiar with the exercise of finding a local extremum by solving the equation in which the derivative vanishes. The following Theorem is the theory behind that.

> **Theorem 2.4.2**
>
> *Let f be defined on (a, b), and suppose f has a local extremum at an interior point $c \in (a, b)$. If $f'(c)$ exists as a finite or infinite number, then $f'(c) = 0$.* ♡

Remark Since the conclusion is $f'(c) = 0$, we may as well just assume f is differentiable at c. The reason why we suppose $f'(c)$ exists is that it makes the assumption somewhat looser.

Proof We shall prove by contradiction. Assume $f'(c) \neq 0$, then either $f'(c) > 0$ or $f'(c) < 0$. (We do not exclude the possibilities that $f'(c) = \infty$ or $f'(c) = -\infty$.) If $f'(c) > 0$, then follows from Corollary 2.4.1 that there acts some open ball $B_\delta(c)$ in which $f(x) > f(c)$ for $x > c$ and $f(x) < f(c)$ for $x < c$. This contradicts the fact that $f(c)$ is a local extremum. Similarly, $f'(c) < 0$ will also lead to a contradiction. ∎

The converse of this Theorem is not true.

Example 2.3 Consider $f(x) = x^3$. Its derivative is $f'(x) = 3x^2$. We note that $f'(0) = 0$. But it is clear that $f(0) = 0$ is neither a local maximum nor a local minimum. In fact, $f(x)$ does not have any local extremum.

✍ **Exercise 2.6** Recall the function

$$f(x) = x^2 \sin \frac{1}{x^2} \mathbb{1}\{x \neq 0\}$$

defined in Exercise 2.2. Show that $f(0)$ is not a local extremum. Recall we have shown that $f'(0) = 0$. This is another example that the converse of Theorem 2.4.2 does not hold.

Moreover, it should be emphasized that we conclude in Theorem 2.4.2 that $f'(c) = 0$ under the condition that $f'(c)$ exists. $f(c)$ may be a local extremum without having a derivative there.

Example 2.4 The simplest example is $f(x) = |x|$. Note that $f(0) = 0$ is a local minimum. But the derivative does not exist at $x = 0$ since $f'_+(0) = 1 \neq -1 = f'_-(0)$.

2.4.2 Rolle's Theorem and Mean Value Theorem

If a function starts from point a and ends at point b with the same level of height as its initial position, then there must be a turning point somewhere in between.

> **Theorem 2.4.3 (Rolle)**
>
> *Let f be defined on $[a, b]$. Suppose $f'(x)$ exists (as finite or infinite number) for each $x \in (a, b)$, and f is continuous at the endpoints a and b. If $f(a) = f(b)$, then there exists $c \in (a, b)$ such that*
> $$f'(c) = 0$$

Proof If f is a constant function, i.e., $f(x) = f(a) = f(b) \; \forall x \in [a, b]$, then the conclusion is trivial since $f'(x) = 0 \; \forall x \in (a, b)$.

We now suppose f is non-constant. We know that f is continuous on $[a, b]$ since f' exists throughout (a, b) and f is continuous at endpoints. It then follows from Theorem 1.2.2 that f attains its maximum and minimum on $[a, b]$. Because we assume f is non-constant and $f(a) = f(b)$, at least one of the maximum and the minimum of f is reached at an interior point $c \in (a, b)$. Note that $f(c)$ is a global extremum, and of course also a local extremum. Then by applying Theorem 2.4.2, we conclude $f'(c) = 0$. ∎

Example 2.5 Review Example 2.1. Note that $f(0) = f(4) = 0$. And the derivative vanishes at $x = 1$ and $x = 3$.

The **mean value theorem** is a generalization of Rolle's theorem. Roughly speaking it says f has a derivative value that equals the slope of the secant line joining two endpoints of the graph of f. The mean value theorem itself can be treated as a special case of an even more generalized version, the **generalized mean value theorem** (Theorem 2.4.5).

> **Theorem 2.4.4 (Mean Value Theorem)**
>
> *Suppose f has a derivative (finite or infinite) at each point of (a, b), and suppose f is continuous at endpoints a and b. Then there exists $c \in (a, b)$ such that*
> $$f(b) - f(a) = f'(c)(b - a)$$

> **Theorem 2.4.5 (Generalized Mean Value Theorem)**
>
> *Let f and g be two functions, each having a derivative (finite or infinite) at each point in (a, b). Suppose f and g are both continuous at endpoints a and b, and $f'(x)$ and $g'(x)$ do not assume infinite values simultaneously. Then there exists $c \in (a, b)$ such that*
> $$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)] \tag{2.13}$$

Remark By taking $g(x) = x$, this theorem reduces to the standard mean value theorem, Theorem 2.4.4.

Though it is a generalized version of the mean value theorem, and hence Rolle's theorem, it can be proved easily using Rolle's theorem by constructing a function out of f and g.

Proof Define
$$\psi(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)]$$
It is clear that ψ is continuous on $[a, b]$. Since $f'(x)$ and $g'(x)$ can never both be infinity, the following equation holds:
$$\psi'(x) = f'(x)[g(b) - g(a)] - g'(x)[f(b) - f(a)] \quad \forall x \in (a, b)$$
Furthermore, we note that
$$\psi(a) = f(a)g(b) - f(b)g(a) = \psi(b)$$
Applying Rolle's theorem (Theorem 2.4.3) to ψ leads to $\psi'(c) = 0$ for some $c \in (a, b)$, which is exactly (2.13). ∎

Sometimes, functions f and g in Theorem 2.4.5 may not be defined at endpoints. But all the one-sided limits exist as finite values. In this case, we still wish to apply the theorem. The following is a slight extension of Theorem 2.4.5.

> **Theorem 2.4.6**
>
> *Let f and g be two functions, each having a derivative (finite or infinite) at each point in (a, b). Suppose that $f'(x)$ and $g'(x)$ do not assume infinite values simultaneously, and one-sided limits $f(a+)$, $f(b-)$, $g(a+)$ and $g(b-)$ all exist as finite values. Then there exists $c \in (a, b)$ such that*
> $$f'(c)[g(b-) - g(a+)] = g'(c)[f(b-) - f(a+)] \tag{2.14}$$

The proof is done by simply defining the function values at the endpoints.

Proof Define
$$\tilde{f}(x) = f(x) \quad \forall x \in (a, b), \quad \tilde{f}(a) = f(a+), \quad \text{and} \quad \tilde{f}(b) = f(b-)$$
$$\tilde{g}(x) = g(x) \quad \forall x \in (a, b), \quad \tilde{g}(a) = g(a+), \quad \text{and} \quad \tilde{g}(b) = g(b-)$$
It is evident that $\tilde{f}$ and $\tilde{g}$ are both continuous on $[a, b]$. Furthermore, the derivatives $\tilde{f}'(x) = f'(x)$ and $\tilde{g}'(x) = g'(x)$ for each $x \in (a, b)$. Applying Theorem 2.4.5 to functions $\tilde{f}$ and $\tilde{g}$ leads to (2.14). ∎

The following theorem is an immediate result of the mean value theorem, which provides a sufficient condition for determining strictly increasing and decreasing functions. It also says a function is constant if its derivative vanishes everywhere.

> **Theorem 2.4.7**
>
> *Suppose f has a derivative (finite or infinite) at each point in (a, b), and itself is continuous at the endpoints a and b.*
> 1. *If $f'(x) > 0 \ \forall x \in (a, b)$, then f is strictly increasing on $[a, b]$.*
> 2. *If $f'(x) < 0 \ \forall x \in (a, b)$, then f is strictly decreasing on $[a, b]$.*
> 3. *If $f'(x) = 0 \ \forall x \in (a, b)$, then f is constant on $[a, b]$.*

Proof (Proof of 1) Let $x_1, x_2 \in [a, b]$ with $x_1 < x_2$. Applying Theorem 2.4.4 to f on the interval $[x_1, x_2]$ leads to
$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$$

for some $c \in (x_1, x_2)$. Since $f'(c) > 0$ and $x_2 - x_1 > 0$, we have $f(x_2) > f(x_1)$. Because x_1 and x_2 are arbitrarily chosen, this implies f is strictly increasing on $[a, b]$.

(Proof of 2) Similarly, for any $x_1, x_2 \in [a, b]$ with $x_1 < x_2$, we have

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) < 0$$

(Proof of 3) For any $x_1, x_2 \in [a, b]$ with $x_1 < x_2$, by applying Theorem 2.4.4, we have

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) = 0$$

This implies f is a constant function. ■

If f and g have the same derivatives everywhere in (a, b), then they only differ by a constant.

> **Corollary 2.4.2**
>
> *If f and g are continuous on $[a, b]$, and $f'(x) = g'(x) \ \forall x \in (a, b)$, then $f - g$ is constant on $[a, b]$.*   ♡

Proof   Note that the difference $f - g$ is continuous on $[a, b]$ and $(f - g)'(x) = 0 \ \forall x \in (a, b)$. It then follows from Theorem 2.4.7 that $f - g$ is a constant. ■

2.5 Intermediate Value Theorem

In Exercise 2.4, we have shown that the derivative of function $f(x) = x^2 \sin(1/x^2)\mathbb{1}\{x \neq 0\}$ is given by

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

We plot the original function f along with its derivative f' side by side in Figure 2.5.
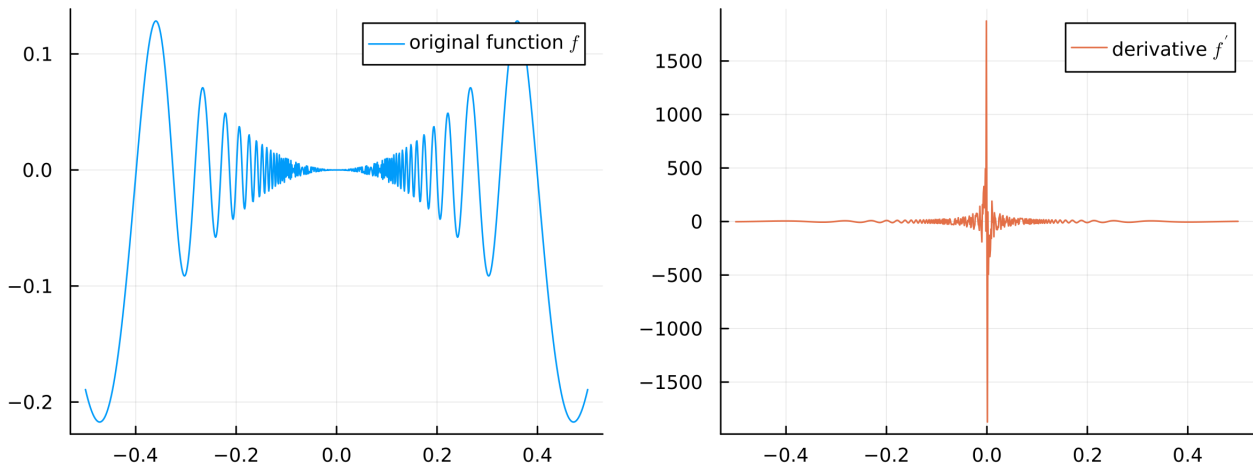

Figure 2.5: Left: original function. Right: derivative.

Note that the derivative f' is not continuous at $x = 0$ since the limit of $f'(x)$ does not exist as $x \to 0$. This example shows that the derivative of a function may not be continuous, and hence theorems for continuous functions do not apply to derivatives in general. However, the **intermediate**

value theorem is an exception. The intermediate value theorem for derivatives is also known as the **Darboux's theorem**.

> **Theorem 2.5.1 (Darboux)**
>
> *Suppose f is defined on $[a, b]$, and it has a derivative (finite or infinite) at each point in (a, b). Assume also that the one-sided derivatives $f'_+(a)$ and $f'_-(b)$ both exist as finite numbers, and $f'_+(a) \neq f'_-(b)$. If k is a number in between $f'_+(a)$ and $f'_-(b)$, then there exists an interior point $c \in (a, b)$ such that*
>
> $$f'(c) = k$$

Proof   Note that f is continuous on the entire closed interval $[a, b]$.

We first consider the case where $k = [f(b) - f(a)]/(b - a)$. Then by applying the mean value theorem (Theorem 2.4.4), there exists $c \in (a, b)$ such that $f'(c) = [f(b) - f(a)]/(b - a) = k$.

Now, we suppose $k \neq [f(b) - f(a)]/(b - a)$. We note that either $f'_+(a) \neq [f(b) - f(a)]/(b - a)$ or $f'_-(b) \neq [f(b) - f(a)]/(b - a)$ since $f'_+(a) \neq f'_-(b)$. In other words, k is either in between $f'_+(a)$ and $[f(b) - f(a)]/(b - a)$ to it is in between $f'_-(b)$ and $[f(b) - f(a)]/(b - a)$.

Without loss of generality, we may assume k is in between $f'_+(a)$ and $[f(b) - f(a)]/(b - a)$. Define a function ϕ on $[a, b]$ by

$$\phi(x) = \begin{cases} \frac{f(x) - f(a)}{x - a} & x \neq a \\ f'_+(a) & x = a \end{cases}$$

Note that ϕ is continuous on $[a, b]$, and k is in between $\phi(a)$ and $\phi(b)$. Hence, we may apply the intermediate values theorem for continuous functions (Theorem 1.3.1) to ϕ, there is a number $d \in (a, b)$ such that

$$k = \phi(d) = \frac{f(d) - f(a)}{d - a}$$

Then applying the mean value theorem (Theorem 2.4.4) to the fraction on the right-hand side of the above equation, we have

$$k = \frac{f(d) - f(a)}{d - a} = f'(c)$$

for some c in between a and d.

For the case where k lying in between $f'_-(b)$ and $[f(b) - f(a)]/(b - a)$, it can be proved with a similar argument, started by defining

$$\phi(x) = \begin{cases} \frac{f(x) - f(b)}{x - b} & x \neq b \\ f'_-(b) & x = b \end{cases}$$

∎

Essentially, Darboux's theorem tells us that though the derivative may not be continuous, it cannot have any *jump discontinuities*. The following is another classical example in addition to the one in Exercise 2.2.

Example 2.6 Consider the function

$$f(x) = x^2 \sin \frac{1}{x} \mathbb{1}\{x \neq 0\} = \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Its derivative is given by

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Note that the derivative f' is not continuous at $x = 0$. The graphs of f and f' are shown in Figure 2.6.
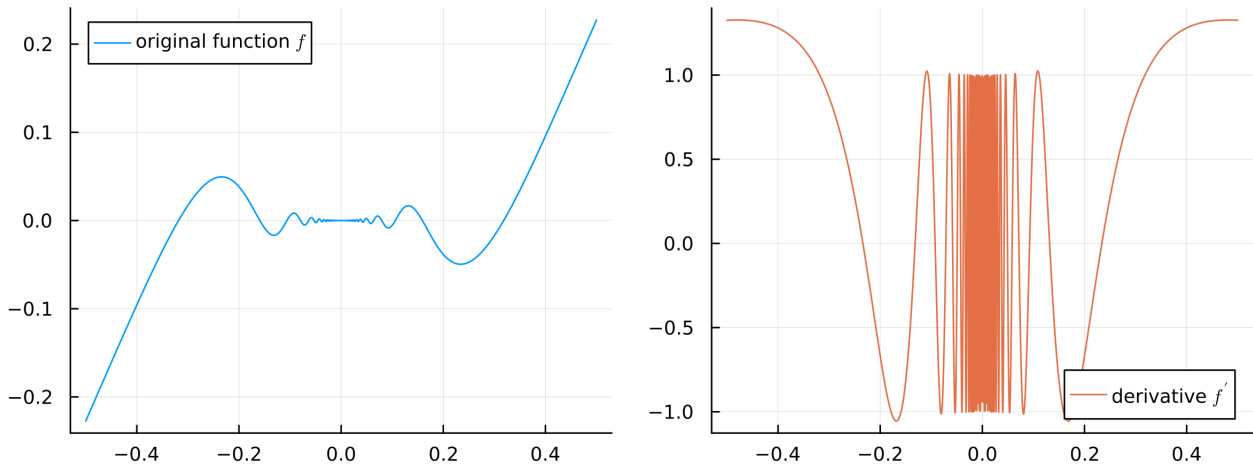

Figure 2.6: Left: original function. Right: derivative.

Actually, there is no need to require that $f'_+(a)$ and $f'_-(b)$ to be finite numbers in Theorem 2.5.1. The following theorem is an extension, and its proof is somewhat neater and more interesting than that of Theorem 2.5.1.

> **Theorem 2.5.2**
>
> *Suppose f is defined on $[a, b]$, and it has a derivative (finite or infinite) at each point in (a, b). Assume also that the one-sided derivatives $f'_+(a)$ and $f'_-(b)$ both exist (each of the two can either be finite or infinite), and $f'_+(a) \neq f'_-(b)$. If k is a number in between $f'_+(a)$ and $f'_-(b)$, then there exists an interior point $c \in (a, b)$ such that*
>
> $$f'(c) = k$$

Proof   Without loss of generality, we may assume $f'_+(a) < k < f'_-(b)$. Define a function g on $[a, b]$ by the equation

$$g(x) = f(x) - kx$$

Note that g is continuous on $[a, b]$, and the derivative of g exists at each point in (a, b). Furthermore, the one-sided derivatives of g, $g'_+(a)$ and $g'_-(b)$, both exist, which are given by

$$g'_+(a) = f'_+(a) - k \quad \text{and} \quad g'_-(b) = f'_-(b) - k \tag{2.15}$$

It should be emphasized that (2.15) also holds for infinite values.

Because we assume $f'_+(a) < k < f'_-(b)$, (2.15) yields

$$g'_+(a) < 0 \quad \text{and} \quad g'_-(b) > 0$$

It then follows from Theorem 2.4.1 that there exist points t_1 and t_2 at which

$$g(t_1) < g(a) \quad \text{and} \quad g(t_2) < g(b)$$

This implies that g must attain its minimum at some interior point $c \in (a, b)$ rather than at the endpoints. Then by applying Theorem 2.4.2, we have

$$g'(c) = 0$$

Because $g'(x) = f'(x) - k \ \forall x \in (a, b)$, it follows that

$$f'(c) = k$$

This completes the proof. ∎

Example 2.7 Consider a function defined on $[0, 1]$ given by

$$f(x) = \sqrt[3]{x(x - 1)}, \quad x \in [0, 1]$$

Note that the derivative exists as a finite number in $(0, 1)$. And it can be shown that $f'_+(0) = -\infty$ and $f'_-(1) = \infty$. Then Theorem 2.5.2 tells us f' takes all the real numbers in $(0, 1)$. See Figure 2.7.
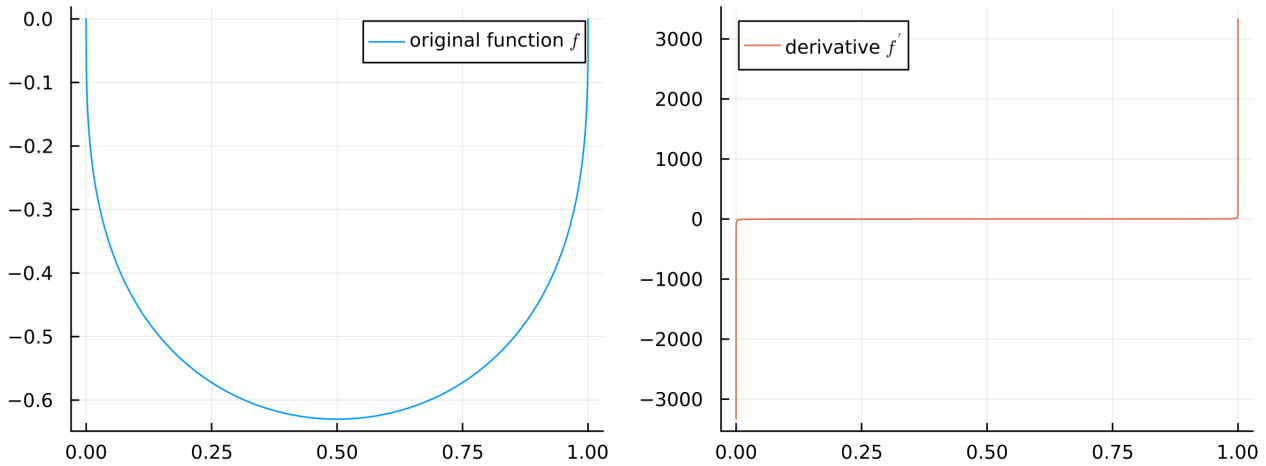

Figure 2.7: Left: original function. Right: derivative.

Note *Theoretically, the graph of f' in Figure 2.7 will tend to $-\infty$ and ∞ as x approaches 0 and 1, respectively. The reason why the absolute values of the derivatives near the endpoints shown in the graph are only about 3000 is because of the limited computational precision of the computer.*

By Theorem 2.4.7, we know a function is strictly increasing if its derivative is always positive, and it is strictly decreasing if its derivative is always negative. Suppose now we only know that f' is nonzero in some open interval, what can we conclude? Thanks to the intermediate value theorem, or Darboux's theorem, f' is either always positive or always negative. Hence, we can conclude that f must be strictly monotonic.

> **Theorem 2.5.3**
>
> *Suppose f has a derivative (finite or infinite) in (a, b), and is continuous at endpoints a and b. If $f'(x) \neq 0 \; \forall x \in (a, b)$, then f is strictly monotonic on $[a, b]$.*

Proof Pick a point $x_1 \in (a, b)$. Suppose $f'(x_0) > 0$. Let x_0 be fixed, and choose an arbitrary point $x \in (a, b)$ other than x_0. We claim that $f'(x) > 0$. Otherwise, if $f'(x) < 0$, then by Theorem 2.5.2, there exists some point c in between x_0 and x such that $f'(c) = 0$, which contradicts the given condition that f' is nonzero in (a, b). In this case, we have $f'(x) > 0 \; \forall x \in (a, b)$. It then follows from Theorem 2.4.7 that f is strictly increasing on $[a, b]$.

If $f'(x_0) < 0$, then a similar argument will show that f is strictly decreasing on $[a, b]$. ■

It is because the derivatives cannot have jump discontinuities that monotonic derivatives are necessarily continuous.

> **Theorem 2.5.4**
>
> *Suppose f' exits, and is monotonic in (a, b). Then f' is continuous in (a, b).*

Proof Without loss of generality, suppose that f' is increasing. We shall prove by contradiction. Assume that f' is discontinuous at $x = c \in (a, b)$. By Theorem 1.4.1, we have $f'(c-)$ and $f'(c+)$ both exist, and

$$f'(c-) \leq f(c) \leq f'(c+)$$

But because we assume f' is not continuous at $x = c$, it holds that

$$f'(c-) < f'(c+)$$

Let $[c - \delta, c + \delta] \; (\delta > 0)$ be a closed interval contained in (a, b). We have

$$f'(c - \delta) \leq f'(c-) < f'(c+) \leq f'(c + \delta)$$

Let k be a number in between $f'(c-)$ and $f'(c+)$ other than $f'(c)$. That is,

$$f'(c-) < k < f'(c+) \quad \text{and} \quad k \neq f'(c)$$

Applying Theorem 2.5.2 to f' on $[c - \delta, c + \delta]$, we conclude that there exists $x_0 \in (c - \delta, c + \delta)$ such that $f'(x_0) = k$. But since f' is increasing, $f'(c-) < k = f'(x_0)$ and $k \neq f'(c)$, we have $x_0 > c$. On the other hand, because $f'(x_0) = k < f'(c+)$ and $k \neq f'(c)$, we have $x_0 < c$. This leads to a contradiction. ■

2.6 Taylor's Theorem

If f is differentiable at $x = c$, then it can be approximated by

$$f(x) \approx f(c) + f'(c)(x - c) \tag{2.16}$$

when $|x - c|$ is small. Geometrically, f is close to its tangent line at $x = c$ near the point of tangency. Another way to look at (2.16) is that f is approximated by a polynomial of degree one whose first-order derivative at $x = c$ is exactly $f'(c)$. And the value of this polynomial at $x = c$ is $f(c)$, that is, its zeroth-order derivative is $f(c)$. We can extend this idea by approximating f with higher degree

polynomials whose derivatives of all orders match with that of f at $x = c$.

Example 2.8 Suppose we know any order derivative of the function $f(x) = e^x$ is itself. That is,

$$f^{(k)}(x) = f(x) = e^x \quad \forall k = 0, 1, 2, \ldots$$

Consider the polynomial

$$p_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots + \frac{x^n}{n!}$$

One can check that

$$p_n^{(k)}(0) = f^{(k)}(0) \quad \forall k = 0, 1, \ldots, n$$

Hence, it is reasonable to approximate $f(x)$ with $p_n(x)$ near point $x = 0$. Figure 2.8 compares f with p_n with several choices of n.
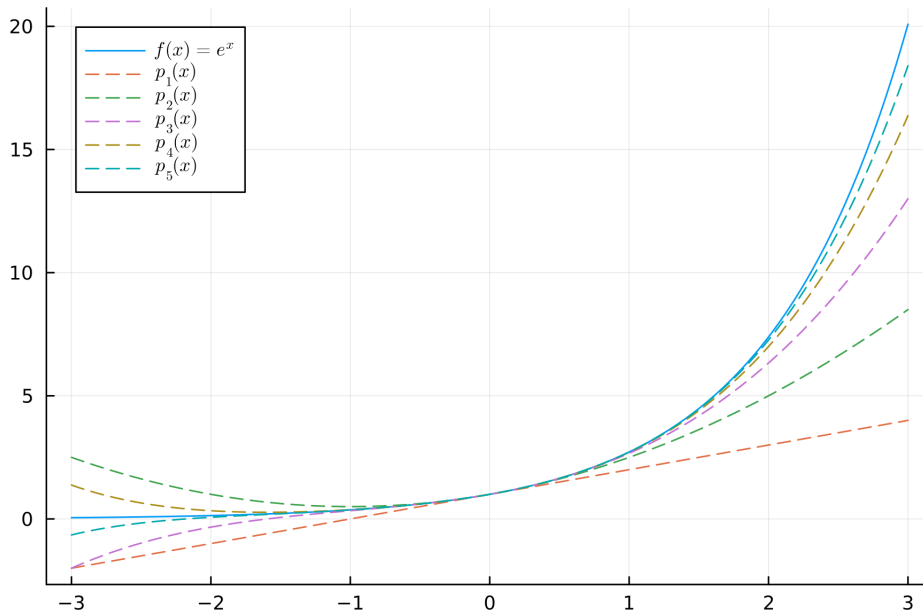

Figure 2.8: Approximations of f at $x = 0$.

In general, a function f can be approximated by the polynomial

$$\sum_{k=0}^{n-1} \frac{f^{(k)}(c)}{k}(x - c)^k = f(c) + f'(c)(x - c) + \cdots + \frac{f^{(n)}(c)}{n!}(x - c)^n$$

at point c provided that f has up to $(n + 1)$-th order derivative. The reason why we approximate f with a polynomial of n degree is that we have a remainder term, which makes use of the $(n + 1)$-th order derivative. The remainder tells us how accurate the approximation is. The formal statement is described in the following theorem, the **Taylor's theorem**.

> **Theorem 2.6.1 (Taylor)**
>
> *Let f be a function having $(n + 1)$-th order derivative everywhere in (a, b), and suppose $f^{(n)}$ is continuous on $[a, b]$. Let $c \in (a, b)$ be an interior point. Then for every $x \in [a, b]$ other than c,*

we have

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1} \tag{2.17}$$

where ξ is some number in between x and c.

This theorem can be obtained as an immediate corollary of the following extension of the mean value theorem.

Theorem 2.6.2

Let f and g be two functions each having $(n+1)$-th order derivative everywhere in (a,b). Suppose that $f^{(n)}$ and $g^{(n)}$ are both continuous on $[a,b]$. Let $c \in (a,b)$ be an interior point. Then for every $x \in [a,b]$ other than c, we have

$$\left[f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k \right] g^{(n+1)}(\xi) = f^{(n+1)}(\xi) \left[g(x) - \sum_{k=0}^{n} \frac{g^{(k)}(c)}{k!}(x-c)^k \right] \tag{2.18}$$

where ξ is some number in between x and c.

Remark Taking $g(x) = (x-c)^{n+1}$, we have

$$g^{(n+1)}(\xi) = (n+1)! \quad \text{and} \quad \sum_{k=0}^{n} \frac{g^{(k)}(c)}{k!}(x-c)^k = 0$$

Then (2.18) reduces to (2.17), and hence this theorem reduces to Theorem 2.6.1.

Proof Define functions $F(t)$ and $G(t)$ on $[a,b]$ by

$$F(t) = \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k \quad \text{and} \quad G(t) = \sum_{k=0}^{n} \frac{g^{(k)}(t)}{k!}(x-t)^k \quad \forall t \in [a,b]$$

Note that F and G are continuous on $[a,b]$, and they have derivatives everywhere in (a,b). Applying the generalized mean value theorem (Theorem 2.4.5) to $F(t)$ and $G(t)$ on the closed interval joining points x and c, we have

$$[F(x) - F(c)]G'(\xi) = F'(\xi)[G(x) - G(c)] \tag{2.19}$$

for some number ξ in between x and c.

After a few steps of computation, we obtain derivatives $F'(t)$ and $G'(t)$ as follows.

$$F'(t) = \frac{f^{(n+1)}(t)}{n!}(x-t)^n \quad \text{and} \quad G'(t) = \frac{g^{(n+1)}(t)}{n!}(x-t)^n \quad \forall t \in (a,b) \tag{2.20}$$

We also note that

$$F(x) = f(x), \quad F(c) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k, \quad G(x) = g(x), \quad \text{and} \quad G(c) = \sum_{k=0}^{n} \frac{g^{(k)}(c)}{k!}(x-c)^k \tag{2.21}$$

Plugging (2.20) and (2.21) into (2.19) leads to (2.18). ∎

We can rewrite (2.17) as

$$f(x) = p_n(x) + r_n(x)$$

where

$$p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x - c)^k$$

is the polynomial approximation, and

$$r_n(x) = \frac{f^{(n+1)}(\xi)}{(n + 1)!}(x - c)^{n+1}$$

is the remainder. The smaller the absolute value of $r_n(x)$ is, the more accurate the approximation will be.

✍ **Exercise 2.7** Use Taylor's theorem (Theorem 2.6.1) to approximate the value of $\ln(0.6)$ so that the absolute error is less than 0.1.

Solution *Putting $f(x) = \ln(x)$ in (2.17), we have*

$$p_n(x; c) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x - c)^k = \ln(c) + \sum_{k=1}^{n} \frac{(-1)^{k+1}}{kc^n}(x - c)^k$$

$$r_n(x; c) = \frac{(-1)^{n+2}}{(n + 1)\xi^{n+1}}(x - c)^{n+1}$$

Since the value of $\ln(x)$ at $x = 1$ is known, i.e., $\ln(1) = 0$, we can expand it about that point by putting $c = 1$. Then we have

$$p_n(x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}(x - 1)^k \quad and \quad r_n(x) = \frac{(-1)^{n+2}}{(n + 1)\xi^{n+1}}(x - 1)^{n+1}$$

To approximate $\ln(0.6)$ by $p_n(0.6)$ within the error bounds, we need to estimate the remainder $r_n(0.6)$. Note that ξ is now some number in between 0.6 and 1 since $x = 0.6$ and $c = 1$. The absolute value of $r_n(0.6)$ is bounded above by

$$|r_n(0.6)| = \frac{1}{n + 1}\left|\frac{0.6 - 1}{\xi}\right|^{n+1} < \frac{1}{n + 1}\left(\frac{0.4}{0.6}\right)^{n+1} = \frac{1}{n + 1}\left(\frac{2}{3}\right)^{n+1} \tag{2.22}$$

We then estimate $|r_n(0.6)|$ using (2.22) by trying the first few values of n. We have

$$r_1(0.6) = \frac{2}{9} > 0.1 \quad and \quad r_2(0.6) = \frac{8}{81} < 0.1$$

Hence, it is adequate to approximate $\ln(0.6)$ by $p_2(0.6)$. The approximated value is

$$\ln(0.6) \approx p_2(0.6) = (0.6 - 1) - \frac{1}{2} \times (0.6 - 1)^2 = -0.48$$

We can also visualize the result in Figure 2.9. As we can see, the approximated value is quite close to the true value. And the absolute error is actually much less than the required bound, 0.1.
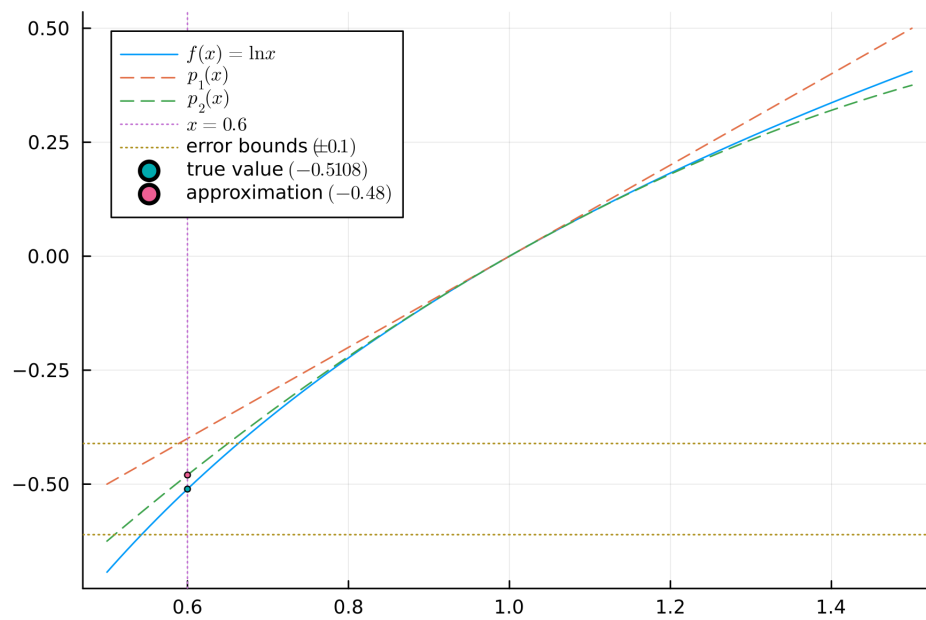
Figure 2.9: Approximation of $\ln(0.6)$.

Chapter 3   Functions of Bounded Variation

3.1   Characterization of Functions of Bounded Variation

> **Theorem 3.1.1 (Characterization of Functions of Bounded Variation)**
>
> *Let f be a real-valued function on $[a, b]$. Then, the following statements are equivalent.*
>
> 1. *f is of bounded variation on $[a, b]$.*
> 2. *There exist two increasing functions g and h on $[a, b]$ such that $f = g - h$.*
> 3. *There exist two **strictly** increasing functions g and h on $[a, b]$ such that $f = g - h$.*

Chapter 4 Riemann-Stieltjes Integral

4.1 Definition of the Riemann-Stieltjes Integral

> **Definition 4.1.1**
>
> *Suppose that f and α are **real-valued bounded** functions on $[a,b]$. Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition on $[a,b]$ and t_k a point in the sub-interval $[x_{k-1}, x_k]$. A sum of the form*
> $$S(P, f, \alpha) = \sum_{k=1}^{n} f(t_k)\Delta\alpha_k$$
> *is called a Riemann-Stieltjes sum of f with respect to α. We say f is Riemann-integrable with respect to α, and write $f \in \mathfrak{R}(\alpha)$ on $[a,b]$ if there exists a number A having the following property: for any given $\varepsilon > 0$, there exists a partition P_ε such that*
> $$|S(P, f, \alpha) - A| < \varepsilon$$
> *for any refinement P of P_ε and for any choice of points t_k. (Note that $S(P, f, \alpha)$ depends on t_k.) Moreover, the number A is **uniquely** determined if it exists (this is proved in Proposition 4.1.1) and is denoted by*
> $$\int_a^b f \, \mathrm{d}\alpha$$
> ♣

Note *To ease the nation, we will sometimes write*
$$\sum_k f(t_k)\delta\alpha_k$$
without specifying the lower and upper limits of k if no confusion will be caused.

Recall Definition 4.1.1 only requires the existence of A. We now show that such number A is also unique.

> **Proposition 4.1.1**
>
> Let $S(P, f, \alpha)$ be as in Definition 4.1.1. If A and A' both satisfy the property stated in Definition 4.1.1, then $A = A'$.
> ♠

Proof Given $\varepsilon > 0$, by the property in Definition 4.1.1, there exist partitions P_1 and P_2 such that
$$|S(P_1, f, \alpha) - A| < \varepsilon/2 \qquad\qquad |S(P_2, f, \alpha) - A'| < \varepsilon/2$$
Let $P = P_1 \cup P_2$. We have
$$|S(P, f, \alpha) - A| < \varepsilon/2 \qquad\qquad |S(P, f, \alpha) - A'| < \varepsilon/2$$
It then follows that
$$|A - A'| \leq |S(P, f, \alpha) - A| + |S(P, f, \alpha) - A'| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$
Since $\varepsilon > 0$ is arbitrary, we must have $A = A'$. ∎

4.2 Linear Properties

The integral is linear in the integrand. In other words, the integral of a linear combination of functions is equal to the linear combination of integrals of each function.

> **Theorem 4.2.1**
>
> *Suppose that $f, g \in \mathfrak{R}(\alpha)$ on $[a, b]$. Then $c_1 f + c_2 g \in \mathfrak{R}(\alpha)$ on $[a, b]$ where c_1 and c_2 are constants. In that case,*
> $$\int_a^b c_1 f + c_2 g \ \mathrm{d}\alpha = c_1 \int_a^b f \ \mathrm{d}\alpha + c_2 \int_a^b g \ \mathrm{d}\alpha$$
> ♡

Proof Let P be a partition on $[a, b]$. The Riemann-Stieltjes sum of $c_1 f + c_2 g$ can be written as

$$S(P, c_1 f + c_2 g, \alpha) = \sum_k c_1 f(t_k) + c_2 g(t_k) \Delta \alpha_k$$

$$= c_1 \sum_k f(t_k) \Delta \alpha_k + c_2 \sum_k g(t_k) \Delta \alpha_k$$

$$= c_1 S(P, f, \alpha) + c_2 S(P, g, \alpha)$$

Given $\varepsilon > 0$. Since $f \in \mathfrak{R}(\alpha)$ on $[a, b]$ then there exists a partition P_ε' such that

$$\left| S(P, f, \alpha) - \int_a^b f \ \mathrm{d}\alpha \right| < \frac{\varepsilon/2}{1 + |c_1|} \quad \forall P \supset P_\varepsilon'$$

Similarly, since $g \in \mathfrak{R}(\alpha)$, there exists a partition P_ε'' such that

$$\left| S(P, g, \alpha) - \int_a^b g \ \mathrm{d}\alpha \right| < \frac{\varepsilon/2}{1 + |c_2|} \quad \forall P \supset P_\varepsilon''$$

Let P_ε be the refinement of P_ε' and P_ε'', i.e., $P = P_\varepsilon' \cup P_\varepsilon''$. Then for any $P \supset P_\varepsilon$, we have

$$\left| S(P, c_1 f + c_2 g, \alpha) - c_1 \int_a^b f \ \mathrm{d}\alpha - c_2 \int_a^b g \ \mathrm{d}\alpha \right|$$

$$\leq |c_1| \left| S(P, f, \alpha) - \int_a^b f \ \mathrm{d}\alpha \right| + |c_2| \left| S(P, g, \alpha) - \int_a^b g \ \mathrm{d}\alpha \right|$$

$$< |c_1| \frac{\varepsilon/2}{1 + |c_1|} + |c_2| \frac{\varepsilon/2}{1 + |c_2|}$$

$$< \varepsilon$$

This completes the proof. ∎

The integral is also linear in the integrator.

> **Theorem 4.2.2**
>
> *If $f \in \mathfrak{R}(\alpha)$ and $f \in \mathfrak{R}(\beta)$ on $[a, b]$, then $f \in \mathfrak{R}(c_1 \alpha + c_2 \beta)$ where c_1 and c_2 are constants. In that case,*
> $$\int_a^b f \ \mathrm{d}(c_1 \alpha + c_2 \beta) = c_1 \int_a^b f \ \mathrm{d}\alpha + c_2 \int_a^b f \ \mathrm{d}\beta$$
> ♡

Proof Let P be a partition on $[a, b]$. We have

$$S(P, f, c_1\alpha + c_2\beta) = \sum_{k=0} f(t_k) + \Delta(c_1\alpha + c_2\beta)_k$$

$$= c_1 \sum_{k=0} f(t_k)\Delta\alpha_k + c_2 \sum_{k=0} f(t_k)\Delta\beta_k$$

$$= c_1 S(P, f, \alpha) + c_2 S(P, f, \beta)$$

Given $\varepsilon > 0$. Since $f \in \mathfrak{R}(\alpha)$ on $[a, b]$ then there exists a partition P'_ε such that

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \frac{\varepsilon/2}{1 + |c_1|} \quad \forall P \supset P'_\varepsilon$$

Similarly, since $f \in \mathfrak{R}(\beta)$, there exists a partition P''_ε such that

$$\left| S(P, f, \beta) - \int_a^b f \, d\beta \right| < \frac{\varepsilon/2}{1 + |c_2|} \quad \forall P \supset P''_\varepsilon$$

Let $P = P'_\varepsilon \cup P''_\varepsilon$. Then for any $P \supset P_\varepsilon$, we have

$$\left| S(P, f, c_1\alpha + c_2\beta) - c_1 \int_a^b f \, d\alpha - c_2 \int_a^b f \, d\beta \right|$$

$$\leq |c_1| \left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| + |c_2| \left| S(P, f, \beta) - \int_a^b f \, d\beta \right|$$

$$< |c_1| \frac{\varepsilon/2}{1 + |c_1|} + |c_2| \frac{\varepsilon/2}{1 + |c_2|}$$

$$< \varepsilon$$

∎

If we divide the interval $[a, b]$ into two parts with some point $c \in (a, b)$ in the middle, then the integral over the entire interval is the sum of the integrals on these two sub-intervals. This is also a kind of linearity of integrals considering the interval of integration.

> **Lemma 4.2.1**
>
> *Suppose $c \in (a, b)$. We have*
>
> $$\int_a^b f \, d\alpha = \int_a^c f \, d\alpha + \int_c^b f \, d\alpha \qquad (4.1)$$
>
> *The existence of two integrals in (4.1) will imply the existence of the third one.* ♡

Proof We have

$$S(P, f, \alpha) = S(P', f, \alpha) + S(P'', f, \alpha) \quad \forall P = P' \cup P'' \qquad (4.2)$$

where P, P' and P'' are partitions on $[a, b]$, $[a, c]$ and $[c, b]$, respectively. Let $\varepsilon > 0$ be chosen arbitrarily.

(Proof of existence of $\int_a^b f \, d\alpha$) Assume $\int_a^c f \, d\alpha$ and $\int_c^b f \, d\alpha$ exist. Then

$$\left| S(P', f, \alpha) - \int_a^c f \, d\alpha \right| < \varepsilon/2 \quad \forall P' \supset P'_\varepsilon$$

for some P'_ε on $[a, c]$. And

$$\left| S(P'', f, \alpha) - \int_a^c f \, d\alpha \right| < \varepsilon/2 \quad \forall P'' \supset P''_\varepsilon$$

for some P_ε'' on $[c, b]$. Let

$$P_\varepsilon = P_\varepsilon' \cup P_\varepsilon''$$

(Note that $c \in P_\varepsilon$.) Let

$$P \supset P_\varepsilon \qquad\qquad P' = P \cap [a, c] \qquad\qquad P'' = P \cap [c, b]$$

Observe that

$$P' \supset P_\varepsilon' \qquad\qquad\qquad\qquad P'' \supset P_\varepsilon''$$

It then follows from (4.2) that

$$\left| S(P, f, \alpha) - \int_a^c f \, d\alpha - \int_c^b f \, d\alpha \right|$$
$$\leq \left| S(P', f, \alpha) - \int_a^c f \, d\alpha \right| + \left| S(P'', f, \alpha) - \int_c^b f \, d\alpha \right|$$
$$< \varepsilon/2 + \varepsilon/2$$
$$= \varepsilon$$

Therefore, $f \in \mathfrak{R}(\alpha)$ on $[a, b]$ and (4.1) holds.

(Proof of existence of $\int_c^b f \, d\alpha$) Assume $\int_a^b f \, d\alpha$ and $\int_a^c f \, d\alpha$ exist. Then

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon/2 \quad \forall P \supset P_\varepsilon$$

for some P_ε on $[a, b]$. And

$$\left| S(P', f, \alpha) - \int_a^c f \, d\alpha \right| < \varepsilon/2 \quad \forall P' \supset P_\varepsilon'$$

for some P_ε' on $[a, c]$. Let

$$P_\varepsilon'' = (P_\varepsilon \cup P_\varepsilon') \cap [c, b]$$

Let

$$P'' \supset P_\varepsilon'' \qquad\qquad P' \supset (P_\varepsilon \cup P_\varepsilon') \cap [a, c] \qquad\qquad P = P' \cup P''$$

Observe that

$$P' \supset P_\varepsilon' \qquad\qquad P = P' \cup P'' \supset (P_\varepsilon \cap [a, c]) \cup (P_\varepsilon \cap [c, b]) = P_\varepsilon$$

It then follows from (4.2) that

$$\left| S(P'', f, \alpha) - \int_a^b f \, d\alpha + \int_a^c f \, d\alpha \right|$$
$$\leq \left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| + \left| S(P', f, \alpha) - \int_a^c f \, d\alpha \right|$$
$$< \varepsilon/2 + \varepsilon/2$$
$$= \varepsilon$$

Therefore, $f \in \mathfrak{R}(\alpha)$ on $[c, b]$ and (4.1) holds.

(Proof of existence of $\int_a^c f \, d\alpha$) Assume $\int_a^b f \, d\alpha$ and $\int_c^b f \, d\alpha$ exist. Then

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon/2 \quad \forall P \supset P_\varepsilon$$

for some P_ε on $[a, b]$. And

$$\left| S(P'', f, \alpha) - \int_c^b f \, \mathrm{d}\alpha \right| < \varepsilon/2 \quad \forall P'' \supset P_\varepsilon''$$

for some P_ε'' on $[c, b]$. Let

$$P_\varepsilon' = (P_\varepsilon \cup P_\varepsilon'') \cap [a, c]$$

Let

$$P' \supset P_\varepsilon' \qquad\qquad P'' \supset (P_\varepsilon \cup P_\varepsilon'') \cap [c, b] \qquad\qquad P = P' \cup P''$$

Observe that

$$P'' \supset P_\varepsilon'' \qquad\qquad P = P' \cup P'' \supset (P_\varepsilon \cap [a, c]) \cup (P_\varepsilon \cap [c, b]) = P_\varepsilon$$

It then follows from (4.2) that

$$\left| S(P', f, \alpha) - \int_a^b f \, \mathrm{d}\alpha + \int_c^b f \, \mathrm{d}\alpha \right|$$
$$\leq \left| S(P, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| + \left| S(P'', f, \alpha) - \int_c^b f \, \mathrm{d}\alpha \right|$$
$$< \varepsilon/2 + \varepsilon/2$$
$$= \varepsilon$$

Therefore, $f \in \mathfrak{R}(\alpha)$ on $[a, c]$ and (4.1) holds.

∎

4.3 Integration by Parts

> **Theorem 4.3.1 (Integration by Parts)**
>
> *If $f \in \mathfrak{R}(\alpha)$ on $[a, b]$, then $\alpha \in \mathfrak{R}(f)$ on $[a, b]$, and*
> $$\int_a^b f \, \mathrm{d}\alpha + \int_a^b \alpha \, \mathrm{d}f = f(b)\alpha(b) - f(a)\alpha(a)$$
> ♡

Remark This can be treated as the *reciprocity law* for integrals.

Proof Given $\varepsilon > 0$, since $f \in \mathfrak{R}(\alpha)$ on $[a, b]$, there exists a partition P_ε such that

$$\left| S(P, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| < \varepsilon \quad \forall P \supset P_\varepsilon \tag{4.3}$$

Let $P = \{x_0, x_1, \ldots, x_n\} \supset P_\varepsilon$ be any refinement of P_ε. The Riemann-Stieltjes sum of α with respect to f is

$$S(P, \alpha, f) = \sum_{k=1}^n \alpha(t_k)(f(x_k) - f(x_{k-1})) = \sum_{k=1}^n \alpha(t_k)f(x_k) - \sum_{k=1}^n \alpha(t_k)f(x_{k-1}) \tag{4.4}$$

Let

$$P^* = P \cup \{t_k \mid 1 \leq k \leq n\} \tag{4.5}$$

Denote by A the value

$$A = f(b)\alpha(b) - f(a)\alpha(a)$$

Note that A can be written as

$$A = \sum_{k=1}^{n} f(x_k)\alpha(x_k) - \sum_{k=1}^{n} f(x_{k-1})\alpha(x_{k-1}) \tag{4.6}$$

Subtracting (4.4) from (4.6), we obtain

$$A - S(P, \alpha, f) = \sum_{k=1}^{n} f(x_k)(\alpha(x_k) - \alpha(t_k)) + \sum_{k=1}^{n} f(x_{k-1})(\alpha(t_k) - \alpha(x_{k-1}))$$

By recalling the construction of P^* in (4.5), we observe that the right-hand side of the above equation is precisely the Riemann-Stieltjes sum $S(P^*, f, \alpha)$. That is,

$$A - S(P, \alpha, f) = S(P^*, f, \alpha)$$

Since $P^* \supset P \supset P_\varepsilon$, it follows from (4.3) that

$$\left| A - S(P, \alpha, f) - \int_a^b f \ \mathrm{d}\alpha \right| = \left| S(P^*, f, \alpha) - \int_a^b f \ \mathrm{d}\alpha \right| < \varepsilon$$

Recall $A = f(b)\alpha(b) - f(a)\alpha(a)$, we have

$$\left| S(P, \alpha, f) + \int_a^b f \ \mathrm{d}\alpha - f(b)\alpha(b) + f(a)\alpha(a) \right| < \varepsilon \quad \forall P \supset P_\varepsilon$$

This implies that $\alpha \in \mathfrak{R}(f)$ on $[a, b]$, and

$$\int_a^b f \ \mathrm{d}\alpha = - \int_a^b f \ \mathrm{d}\alpha + f(b)\alpha(b) - f(a)\alpha(a)$$

∎

4.4 Change of Variables

> **Theorem 4.4.1**
>
> *Suppose $f \in \mathfrak{R}(\alpha)$ on $[a, b]$, and ϕ is a **strictly monotonic continuous** function on a closed interval I with endpoints c and d. (I is either $[c, d]$ or $[d, c]$.) Assume*
>
> $$a = \phi(c) \qquad\qquad b = \phi(d)$$
>
> *Define two composite functions:*
>
> $$h = f(\phi(x)) \qquad\qquad \beta = \alpha(\phi(x))$$
>
> *Then $h \in \mathfrak{R}(\beta)$ on I, and $\int_a^b f \ \mathrm{d}\alpha = \int_c^d h \ \mathrm{d}\beta$, i.e.,*
>
> $$\int_a^b f(x) \ \mathrm{d}\alpha(x) = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(x)) \ \mathrm{d}\alpha(\phi(x))$$
>
> ♡

Remark The reason why we assume that ϕ is strictly monotonic and continuous is to ensure that it is a bijective function. It is equivalent to assuming ϕ is strictly monotonic and injective.

The main idea of this proof is exploiting the one-to-one relation between the partitions on $[a, b]$ and $[c, d]$.

Proof (One-To-One Relation of Partitions) Without loss of generality, we may assume that g is strictly *increasing* and continuous. Then $I = [c, d]$. From the conditions of g, we can immediately conclude that it has a bijective inverse function

$$\phi^{-1} : [a, b] \to [c, d]$$

For any partition $P' = \{x_0, \ldots, x_n\}$ on $[a, b]$, we can associate it with a partition P on $[c, d]$, which is given by

$$P := \phi^{-1}(P') := \left\{\phi^{-1}(x_0), \ldots, \phi^{-1}(x_n)\right\}$$

On the other hand, for any partition $P = \{y_0, \ldots, y_n\}$ on $[c, d]$, we can define a partition P' on $[a, b]$ by

$$P' := \phi(P) := \{\phi(y_0), \ldots, \phi(y_n)\}$$

(Existence of the Integral) Given $\varepsilon > 0$, since $f \in \mathfrak{R}(\alpha)$ on $[a, b]$, there exists a partition P'_ε on $[a, b]$ such that

$$\left| S(P', f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| < \varepsilon \quad \forall P' \supset P'_\varepsilon \tag{4.7}$$

Let partition P_ε on $[c, d]$ be given by $P_\varepsilon = \phi^{-1}(P'_\varepsilon)$. For any refinement $P \supset P_\varepsilon$, we have

$$S(P, h, \beta) = \sum h(s_i)\Delta\beta_i = \sum h(s_i)(\beta(s_i) - \beta(s_{i-1}))$$

For each point s_i, we can map it to $[a, b]$ by $t_i = \phi(s_i)$. It then follows that

$$S(P, h, \beta) = \sum h(\phi^{-1}(t_i))(\beta(\phi^{-1}(t_i)) - \beta(\phi^{-1}(t_{i-1})))$$
$$= \sum f(t_i)(\alpha(t_i) - \alpha(t_{i-1}))$$
$$= S(P', f, \alpha)$$

where $P' = \phi(P)$. In summary,

$$S(P, h, \beta) = S(P', \phi, \alpha) \tag{4.8}$$

What is left to show is that $P' \supset P'_\varepsilon$. For any point $x \in P'_\varepsilon$, we have $\phi^{-1}(x) \in P_\varepsilon$ since $P_\varepsilon = \phi^{-1}(P_\varepsilon)$. Recall that $P \supset P_\varepsilon$. Thus, $\phi^{-1}(x) \in P$. And because $P' = \phi(P)$, we have $x = \phi(\phi^{-1}(x)) \in P'$. Therefore, indeed $P' \supset P'_\varepsilon$. It then follows from (4.7) and (4.8) that

$$\left| S(P, h, \beta) - \int_a^b f \, \mathrm{d}\alpha \right| < \varepsilon \quad \forall P \supset P_\varepsilon$$

This completes the proof. ∎

4.5 Monotonically Increasing Integrators

It is much easier to study and prove some properties of the Riemann-Stieltjes integrals if we require that α is monotonically increasing.

Thanks to Theorem 3.1.1, which states that every function of bounded variation can express as a difference of two increasing functions, almost all the properties introduced in this section can be extended with ease to integrators of bounded variation.

Index