

Riemann-Stieltjes integral

ljp

October 9, 2023

Contents

1	Basic properties	2
2	Integration by parts	2
3	Change of variable	2
4	Reduction to Riemann integral	3
5	Step function integrator	4
5.1	Simple step function	4
5.2	General step function	5
6	Lower and upper integrals for increasing integrators	5
6.1	Lower and upper sums	6
6.2	Lower and upper integrals	6
6.3	Riemann's condition	7
6.4	Other properties when integrator is increasing	9
7	Integrators of bounded variation	10
8	Existence of Riemann-Stieltjes integrals	12
8.1	Sufficient conditions	12
8.2	Necessary conditions	13
9	Mean value theorems for integrals	13
10	Integral with variable upper limit	14
11	Second fundamental theorem of calculus	15
12	More on Riemann integral	16
12.1	Change of variable	16
12.2	Second MVT	17
13	Riemann-Stieltjes integral with parameter	17
13.1	Differentiation under the integral sign	18
13.2	Interchanging the order of integration	19
14	Lebesgue's criterion for Riemann integrability	20
14.1	Measure zero	20
14.2	Oscillation	20
14.3	Lebesgue's criterion	21

1 Basic properties

Theorem 1.1 (Linearity on integrand). *If $f, g \in \mathcal{R}(\alpha, [a, b])$, then $c_1 f + c_2 g \in \mathcal{R}(\alpha, [a, b])$, and*

$$\int_a^b (c_1 f + c_2 g) d\alpha = c_1 \int_a^b f d\alpha + c_2 \int_a^b g d\alpha.$$

Theorem 1.2 (Linearity on integrator). *If $f \in \mathcal{R}(\alpha, [a, b]) \cap \mathcal{R}(\beta, [a, b])$, then $f \in \mathcal{R}(c_1 \alpha + c_2 \beta, [a, b])$, and*

$$\int_a^b f d(c_1 \alpha + c_2 \beta) = c_1 \int_a^b f d\alpha + c_2 \int_a^b f d\beta.$$

2 Integration by parts

Theorem 2.1 (Integration by parts). *If on $[a, b]$ we have $f \in \mathcal{R}(\alpha)$ and $\alpha \in \mathcal{R}(f)$, then*

$$\int_a^b f d\alpha + \int_a^b \alpha df = [f(x)\alpha(x)]_a^b. \quad (1)$$

The key idea is that $[f\alpha]_a^b = \sum_{k=1}^n [f(x_k)\alpha(x_k) - f(x_{k-1})\alpha(x_{k-1})]$.

Proof. Since $f \in \mathcal{R}(\alpha)$, there is a partition $P_\varepsilon \in \mathcal{P}[a, b]$, every P finer than P_ε will satisfies

$$|S(P, f, \alpha) - \int_a^b f d\alpha| < \varepsilon.$$

For such P ,

$$\begin{aligned} S(P, \alpha, f) - [f\alpha]_a^b &= \sum_{k=1}^n \alpha(t_k) [f(x_k) - f(x_{k-1})] - \sum_{k=1}^n [f(x_k)\alpha(x_k) - f(x_{k-1})\alpha(x_{k-1})] \\ &= - \sum_{k=1}^n f(x_k) [\alpha(x_k) - \alpha(t_k)] - \sum_{k=1}^n f(x_{k-1}) [\alpha(t_k) - \alpha(x_{k-1})]. \end{aligned}$$

If we consider P' formed both partition points x_k of P and sample points t_k of P , then

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

And since P' is finer than P , we have

$$\left| -S(P, \alpha, f) + [f\alpha]_a^b - \int_a^b f d\alpha \right| < \varepsilon.$$

■

3 Change of variable

Theorem 3.1 (Change of variable). *If $f \in \mathcal{R}(\alpha)$ on $[a, b]$. Let g be a strictly monotonic continuous function (this implies that g is injective), and g maps endpoints of $[c, d]$ to endpoints $[a, b]$ (this implies that g is surjective). Then we have*

$$\int_a^b f d\alpha = \int_c^d (f \circ g) d(\alpha \circ g),$$

or more specifically, suppose $g(c) = a$, $g(d) = b$, then

$$\int_{g(c)}^{g(d)} f(x) d\alpha(x) = \int_c^d f(g(x)) d\alpha(g(x)).$$

Proof. Without loss of generality, assume g is strictly increasing, so $g(c) = a$ and $g(d) = b$. Let $\int_a^b f d\alpha = I$. By definition, for every $\varepsilon > 0$, there is a partition $P_\varepsilon \in \mathcal{P}[a, b]$, such that every finer partition $P \supseteq P_\varepsilon$ will satisfy

$$|S(P, f, \alpha) - I| < \varepsilon.$$

Function g in the condition above is bijective, so inverse function g^{-1} exists. Apply g^{-1} on P_ε , we get a partition on $[c, d]$, denote it with $Q_\varepsilon = g^{-1}(P_\varepsilon)$. Now for every partition Q which is finer than Q_ε , we have $g(Q) \supseteq g(Q_\varepsilon) = g(g^{-1}(P_\varepsilon)) = P_\varepsilon$ (last equal sign holds since g is surjective).

Let sample point t'_k and partition point x'_k be $g(t_k)$ and $g(x_k)$ respectively, then

$$\begin{aligned} S(Q, f \circ g, \alpha \circ g) &= \sum_k f(g(t_k)) [\alpha(g(x_k)) - \alpha(g(x_{k-1}))] \\ &= \sum_k f(t'_k) [\alpha(x'_k) - \alpha(x'_{k-1})] \\ &= S(g(Q), f, \alpha). \end{aligned}$$

Since $g(Q)$ is a partition of $[a, b]$ finer than P_ε ,

$$|S(Q, f \circ g, \alpha \circ g) - I| = |S(g(Q), f, \alpha) - I| < \varepsilon.$$

■

4 Reduction to Riemann integral

Theorem 4.1. *If $\alpha \in C^1[a, b]$, then Riemann-Stieltjes integral can be reduced to Riemann integral:*

$$\int_a^b f(x) d\alpha(x) = \int_a^b f(x) \alpha'(x) dx.$$

Let $I = \int_a^b f d\alpha$ and $g(x) = f(x)\alpha'(x)$. The idea is to prove that the Riemann sum approaches I as partition goes finer. By definition we want to estimate $|S(P, g) - I|$ and show that $|S - I| < \varepsilon$ for all partition finer than some partition depending on ε . We'll utilize the Riemann-Stieltjes sum $S(P, f, \alpha)$ of the known integral to connect the Riemann sum $S(P, g)$ and I . This is done by using triangle inequality

$$|S(P, g) - I| \leq |S(P, g) - S(P, f, \alpha)| + |S(P, f, \alpha) - I| < \varepsilon.$$

So the key part is to show that $S(P, f, \alpha)$ and $S(P, g)$ will get arbitrarily closer as P goes finer.

Proof. For every $\varepsilon > 0$, there is a partition $P'_\varepsilon \in \mathcal{P}[a, b]$ such that for every finer partition $P \supseteq P'_\varepsilon$, we have

$$|S(P, f, \alpha) - I| < \varepsilon/2. \tag{2}$$

Since $\alpha \in C^1$, by MVT we have

$$S(P, f, \alpha) = \sum_k f(t_k) [\alpha(x_k) - \alpha(x_{k-1})] = \sum_k f(t_k) \alpha'(\xi_k) \Delta x_k.$$

Where $\xi_k \in (x_{k-1}, x_k)$ and $t_k \in [x_{k-1}, x_k]$. For the same partition and the same choice of inner points t_k , we have the Riemann sum

$$S(P, g) = \sum_k f(t_k) \alpha'(t_k) \Delta x_k.$$

Therefore the difference

$$S(P, g) - S(P, f, \alpha) = \sum_k f(t_k) [\alpha'(\xi_k) - \alpha'(t_k)] \Delta x_k.$$

To estimate this difference, note that f is bounded, so $|f| \leq M$ for some $M > 0$. α' is continuous on compact interval $[a, b]$, hence is uniformly continuous. So there is $\delta > 0$ such that $|x - y| < \delta$ implies $|\alpha'(x) - \alpha'(y)| < \varepsilon/(2M(b-a))$. We can find P' such that $\|P'\| < \delta$, therefore for every partition P finer than P' , we have $|\xi_k - x_k| < \delta$, causing $|\alpha'(\xi_k) - \alpha'(x_k)| < \varepsilon/(2M(b-a))$. For such P ,

$$|S(P, g) - S(P, f, \alpha)| \leq \sum_k |f(t_k)| |\alpha'(\xi_k) - \alpha'(t_k)| \Delta x_k < \varepsilon/2. \quad (3)$$

Let $P_\varepsilon = P'_\varepsilon \cup P'$, every P finer than P_ε will be finer than both P'_ε and P' , therefore we have both (2) and (3), that completes the proof. \blacksquare

5 Step function integrator

5.1 Simple step function

First consider the simple case where the step function has only two “steps”. Suppose α is defined on $[a, b]$, the point $c \in (a, b)$, and α has value

$$\alpha(x) = \begin{cases} \alpha(a) & x \in [a, c) \\ \alpha(b) & x \in [c, b] \end{cases}.$$

Theorem 5.1. *Suppose α is the step function defined above. At least one of f and α are left-continuous at c , and at least one of them is right-continuous at c . Then $f \in \mathcal{R}(\alpha, [a, b])$ and we have*

$$\int_a^b f d\alpha = f(c) [\alpha(c+) - \alpha(c-)].$$

In Apostol’s textbook, the proof was done with the whole interval $[a, b]$, so the left subinterval $[a, c]$ and right subinterval $[c, b]$ are considered at the same time. We consider instead to break the integral into two parts and use the additive property, so that we can deal with left and right subinterval separately and analogously. Nevertheless ideas behind the proofs are the same.

Proof. To prove the theorem, we only need to show two things:

1. $\int_a^c f d\alpha = f(c) [\alpha(c) - \alpha(c-)]$
2. $\int_c^b f d\alpha = f(c) [\alpha(c+) - \alpha(c)]$

Then $\int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha$ and the proof will be completed.

Consider the left subinterval $[a, c]$. The condition requires at least one of α and f is left-continuous at c .

(1) If α is left-continuous, then for arbitrary partition of $[a, c]$ we have $\Delta\alpha_k = 0$, and the Riemann-Stieltjes sum is 0. So $\int_a^c f d\alpha$ is indeed $f(c) [\alpha(c) - \alpha(c-)] = 0$.

(2) If f is left-continuous at c , that is $f(c-) = f(c)$. Let $J = \alpha(c) - \alpha(c-)$. For every $\varepsilon > 0$, there is $\delta > 0$, such that $x \in (c - \delta, c)$ implies $|f(x) - f(c)| < \varepsilon/|J|$. Now we can find a partition P_ε with $\|P_\varepsilon\| < \delta$. For all partition P finer than P_ε , the Riemann-Stieltjes sum is

$$S(P, f, \alpha) = \sum_k f(t_k) [\alpha(x_k) - \alpha(x_{k-1})] = f(t_n) [\alpha(c) - \alpha(c-)] = f(t_n) J.$$

Since $t_n \in [x_{n-1}, c]$, so $|t_n - c| \leq \|P\| \leq \|P_\varepsilon\| < \delta$. Therefore $|f(t_n) - f(c)| < \varepsilon/|J|$, and we have

$$|S(P, f, \alpha) - f(c)J| = |f(t_n) - f(c)||J| < \varepsilon.$$

This means that $\int_a^c f d\alpha = f(c)J = f(c)[\alpha(c) - \alpha(c-)]$.

For right subinterval $[c, b]$, the proof is analogous. ■

This theorem tells us that the existence and the value Riemann-Stieltjes integral $\int_a^b f d\alpha$ may be affected if we change the value of f at a single point. However for Riemann integral $\int_a^b f(x) dx$, changing the value of f at finite points will not affect the existence and the value of this integral. To prove this, consider changing the value of f at a single point, denote this new function g , now show that $\int_a^b (g(x) - f(x)) dx = 0$, this is easy because $g - f$ is zero except at single point. Therefore $\int_a^b g(x) dx = \int_a^b f(x) dx$.

Example. Let $\alpha(x) = 0$ for $x \neq 0$ and $\alpha(x) = -1$. Let $f(x) = 1$ for $x \in [-1, 1]$. Then by theorem 5.1, the integral $\int_{-1}^1 f d\alpha = f(0)[\alpha(0+) - \alpha(0-)] = 0$. If we change the value of f at 0 to $f(0) = 2$, now the integral doesn't exist.

5.2 General step function

Now we consider the general case. The function α defined on $[a, b]$ is called the step function if there is a partition $P = \{x_0 < x_1 < \dots < x_n\} \in \mathcal{P}[a, b]$ such that α is constant on each subinterval (x_{k-1}, x_k) , $k = 1, \dots, n-1$. The number $\alpha(x_k+) - \alpha(x_k-)$ is called the jump at x_k .

Using step functions, Riemann-Stieltjes integral can be reduced to a finite sum.

Theorem 5.2 (Reduction to finite sum). *Suppose α is a step function with jumps at x_k , $k = 1, \dots, n$. If at each x_k , at least one of f and α is left-continuous and at least one of them is right-continuous, then*

$$\int_a^b f(x) d\alpha(x) = \sum_{k=1}^n f(x_k) [\alpha(x_k+) - \alpha(x_k-)].$$

Proof. Divide interval $[a, b]$ into n subintervals such that each contains only one jump. Apply theorem 5.1 and then add those integrals on subintervals. ■

Theorem 5.3 (Floor function integrator). *Every finite sum can be written as Riemann-Stieltjes integral with floor function as integrator. Define f on $[0, n]$ with $f(0) = 0$ and $f(x) = a_k$ for $x \in (k-1, k]$, $k = 1, \dots, n$, then*

$$\sum_{k=1}^n a_k = \int_0^n f(x) d[x].$$

Proof. $\int_0^n f(x) d[x] = \sum_{k=1}^n f(k)([k+] - [k-]) = \sum_{k=1}^n a_k(k - (k-1)) = \sum_{k=1}^n a_k$. ■

6 Lower and upper integrals for increasing integrators

When integrator is increasing, we have definitions of lower and upper sums, also lower and upper integrals. Therefore, Riemann integral (integrator $\alpha(x) = x$) is such type of integral and satisfies all statements in this subsection.

6.1 Lower and upper sums

Refinement of a partition increases the lower sum and decreases the upper sum. And lower sums are never greater than upper sums, no matter what partitions are chosen.

Theorem 6.1 (Lower sum, upper sum and refinement). *Assume $\alpha \uparrow [a, b]$, then:*

1. *If $P' \supseteq P$, then $U(P', f, \alpha) \leq U(P, f, \alpha)$ and $L(P', f, \alpha) \geq L(P, f, \alpha)$*
2. *For arbitrary P and Q , we have $L(P, f, \alpha) \leq U(Q, f, \alpha)$*

As a consequence, for $P' \supseteq P$ implies $U(P') - L(P') \leq U(P) - L(P)$.

6.2 Lower and upper integrals

Definition 6.1 (Lower and upper integral). *Assume $\alpha \uparrow [a, b]$. The upper Stieltjes integral is the infimum of all upper sums, and the lower Stieltjes integral is the supremum of all lower sums.*

$$\begin{aligned} \int_a^b f d\alpha &= \bar{I}(f, \alpha) = \inf_{P \in \mathcal{P}[a, b]} U(P, f, \alpha) \\ \int_a^b d\alpha &= \underline{I}(f, \alpha) = \sup_{P \in \mathcal{P}[a, b]} L(P, f, \alpha). \end{aligned}$$

Theorem 6.2. *Assume $\alpha \uparrow [a, b]$, then $\underline{I}(f, \alpha) \leq \bar{I}(f, \alpha)$.*

Proof. Let L and U be the set of all lower sums and all upper sums respectively, then $L \leq U$ by theorem 6.1, therefore $\sup L \leq \inf U$. ■

Example ($\bar{I} \neq \underline{I}$). *Consider Dirichlet function f , where $f(x) = 1$ for $x \in \mathbb{Q}$ and $f(x) = 0$ for $x \notin \mathbb{Q}$. Apparently, for every partition the Riemann sum $U(P, f) = 1$ and $L(P, f) = 0$. Therefore for Riemann integral ($\alpha(x) = x$) we have $\bar{I}(f) = 1$ and $\underline{I}(f) = 0$.*

Theorem 6.3. *Assume $\alpha \uparrow [a, b]$.*

1. $\bar{J}_a^b = \bar{J}_a^c + \bar{J}_c^b$, $\underline{J}_a^b = \underline{J}_a^c + \underline{J}_c^b$
2. $\bar{J}_a^b f + g d\alpha \leq \bar{J}_a^b f d\alpha + \bar{J}_a^b g d\alpha$
3. $\underline{J}_a^b f + g d\alpha \geq \underline{J}_a^b f d\alpha + \underline{J}_a^b g d\alpha$

Proof. (1) Consider arbitrary $P \in \mathcal{P}[a, b]$, let $P' = P \cup \{c\}$ and divide it into two sub-partition $P'_1 = P' \cap [a, c]$ and $P'_2 = P' \cap [c, b]$, so $P'_1 \in \mathcal{P}[a, c]$ and $P'_2 \in \mathcal{P}[c, b]$. Note that $P' \supseteq P$ and we have

$$U(P, f, \alpha) \geq U(P', f, \alpha) = U(P'_1, f, \alpha) + U(P'_2, f, \alpha) \geq \int_a^c f d\alpha + \int_c^b f d\alpha.$$

So $\bar{J}_a^c + \bar{J}_c^b$ is a lower bound of $U(P, f, \alpha)$ for all $P \in \mathcal{P}[a, b]$, this implies that

$$\inf_{P \in \mathcal{P}[a, b]} U(P, f, \alpha) = \int_a^b f d\alpha \geq \int_a^c f d\alpha + \int_c^b f d\alpha.$$

To prove the other direction, consider arbitrary $P_1 \in \mathcal{P}[a, c]$ and $P_2 \in \mathcal{P}[c, b]$. Then $P_1 \cup P_2 \in \mathcal{P}[a, b]$ and

$$U(P_1, f, \alpha) + U(P_2, f, \alpha) = U(P_1 \cup P_2, f, \alpha) \geq \int_a^b f d\alpha.$$

So we have

$$\begin{aligned}
\int_a^b f d\alpha &\leq \inf \{U(P_1, f, \alpha) + U(P_2, f, \alpha) : P_1 \in \mathcal{P}[a, c], P_2 \in \mathcal{P}[c, b]\} \\
&= \inf \{U(P_1, f, \alpha) : P_1 \in \mathcal{P}[a, c]\} + \inf \{U(P_2, f, \alpha) : P_2 \in \mathcal{P}[c, b]\} \\
&= \int_a^c f d\alpha + \int_c^b f d\alpha.
\end{aligned}$$

(2) For arbitrary $P \in \mathcal{P}[a, b]$, we have

$$\begin{aligned}
U(P, f + g, \alpha) &= \sum_k M_k(f + g) \Delta\alpha_k \\
&\leq \sum_k M_k(f) \Delta\alpha_k + \sum_k M_k(g) \Delta\alpha_k \\
&= U(P, f, \alpha) + U(P, g, \alpha) \\
&\leq \bar{I}(f) + \bar{I}(g).
\end{aligned}$$

Take inf of both sides,

$$\bar{I}(f + g) = \inf_{P \in \mathcal{P}[a, b]} U(P, f + g, \alpha) \leq \bar{I}(f) + \bar{I}(g).$$

(3) Similar to (2). ■

6.3 Riemann's condition

Again we consider integrator $\alpha \uparrow [a, b]$. For a bounded function f , suppose $m \leq f \leq M$ on $[a, b]$. It is immediate that for every partition $P \in \mathcal{P}[a, b]$,

$$m(b - a) \leq L(P, f, \alpha) \leq U(P, f, \alpha) \leq M(b - a).$$

This means that lower and upper sums for bounded functions are always bounded. Therefore for a bounded function f , we can always find the lower and upper integrals (\underline{I} and \bar{I}) for it.

When $\underline{I} = \bar{I}$, the integral exists and from the inequality above we can conclude that $U(P, f, \alpha) - L(P, f, \alpha)$ must be able to get arbitrarily small. This fact is shown in the next theorem.

Definition 6.2 (Riemann's condition). *Function f is said to satisfy Riemann's condition with respect to α on $[a, b]$, if for every $\varepsilon > 0$, there is a partition $P_\varepsilon \in \mathcal{P}[a, b]$, such that $P \supseteq P_\varepsilon$ implies*

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon.$$

Note that if $P' \supseteq P$, then $U(P') - L(P') \leq U(P) - L(P)$. So if we can find one partition P such that $U(P) - L(P) < \varepsilon$, every finer partition will also satisfy the inequality, bringing the Riemann's condition.

Theorem 6.4. *Assume $\alpha \uparrow [a, b]$. The following statements are equivalent:*

1. $f \in \mathcal{R}(\alpha, [a, b])$
2. f satisfies Riemann's condition with respect to α on $[a, b]$
3. $\int_a^b f d\alpha = \int_a^b f d\alpha$

(1 \Rightarrow 2) The idea is to bound $U(P, f, \alpha) - L(P, f, \alpha)$ from above with the difference $S(P, f, \alpha) - S'(P, f, \alpha)$, here S and S' have different choices of $t_k \in [x_{k-1}, x_k]$ and $t'_k \in [x_{k-1}, x_k]$. This leads to bounding $M_k - m_k$ with $f(t_k) - f(t'_k)$ from above, which can be done using approximation property of sup. Using M_k and m_k only as upper and lower

bounds will lead to $M_k - m_k > f(t_k) - f(t'_k)$, not the direction we want. But using approximation property will give us less-than symbol: $M_k - m_k - \varepsilon < f(t_k) - f(t'_k)$, the arbitrary small ε can be easily ignored in the end because $x \leq y + \varepsilon$ or $x - \varepsilon \leq y$ both imply $x \leq y$.

This “ ε -trick” can be handy when we want to prove some inequalities between a variable and its limit-type value (for example sup or integral), in a direction that is hard to be obtained directly. We manually vary from the limit by small amount ε , and it won't affect the direction in the end, but can reverse the relation between the variable and the limit.

This trick works because the nature of limit involves some “error” (ε). To prove $x = y$, one can simply show $|x - y| < \varepsilon$. And to prove $x \leq y$, one can show $x - y \leq \varepsilon$.

Proof. ($1 \Rightarrow 2$) Let

$$S(P, f, \alpha) = \sum_k f(t_k) \Delta \alpha_k, \quad S'(P, f, \alpha) = \sum_k f(t'_k) \Delta \alpha_k,$$

$$U(P, f, \alpha) = \sum_k M_k \Delta \alpha_k, \quad L(P, f, \alpha) = \sum_k m_k \Delta \alpha_k.$$

Here $M_k = \sup f([x_{k-1}, x_k])$ and $m_k = \inf f([x_{k-1}, x_k])$. For every subinterval, we have $M_k - m_k = \sup\{f(x) - f(y) : x, y \in [x_{k-1}, x_k]\}$, hence for every $h > 0$, we can find $t_k, t'_k \in [x_{k-1}, x_k]$, such that

$$M_k - m_k - h < f(t_k) - f(t'_k).$$

Then for sums we have

$$\sum_k (M_k - m_k - h) \Delta \alpha_k < \sum_k [f(t_k) - f(t'_k)] \Delta \alpha_k.$$

that is

$$U(P, f, \alpha) - L(P, f, \alpha) < h[\alpha(b) - \alpha(a)] + S(P, f, \alpha) - S'(P, f, \alpha).$$

Since $f \in \mathcal{R}$, for every $\varepsilon > 0$, there is partition P_ε , such that for every finer P , we have

$$\left| S(P, f, \alpha) - \int_a^b f d\alpha \right| < \varepsilon/3, \quad \left| S'(P, f, \alpha) - \int_a^b f d\alpha \right| < \varepsilon/3.$$

So $|S(P, f, \alpha) - S'(P, f, \alpha)| < 2\varepsilon/3$. If $\alpha(b) - \alpha(a) = 0$, the proof is done. If $\alpha(b) - \alpha(a) > 0$, in order to control $h[\alpha(b) - \alpha(a)]$, choose $h = \varepsilon/3[\alpha(b) - \alpha(a)]$. After that we have $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ for all P finer than P_ε .

($2 \Rightarrow 3$) By Riemann's condition, we have, for every $\varepsilon > 0$, a corresponding $P_\varepsilon \in \mathcal{P}[a, b]$, such that $P \supseteq P_\varepsilon$ implies $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$. That means $U(P, f, \alpha) < L(P, f, \alpha) + \varepsilon$. So $\bar{I} \leq U(P, f, \alpha) < L(P, f, \alpha) + \varepsilon \leq \underline{I} + \varepsilon$. Since ε is arbitrary, we have $\bar{I} \leq \underline{I}$. We already proved that $\underline{I} \leq \bar{I}$, hence $\underline{I} = \bar{I}$.

($3 \Rightarrow 1$) Let $I = \bar{I} = \underline{I}$. For every $\varepsilon > 0$, there is P' and P'' such that

$$L(P', f, \alpha) > \underline{I} - \varepsilon, \quad U(P'', f, \alpha) < \bar{I} + \varepsilon.$$

Let $P_\varepsilon = P' \cup P''$, for every P finer than P_ε , we have $P \supseteq P', P''$, therefore

$$I - \varepsilon < L(P', f, \alpha) \leq L(P, f, \alpha) \leq S(P, f, \alpha) \leq U(P, f, \alpha) \leq U(P'', f, \alpha) < I + \varepsilon.$$

That is $|S(P, f, \alpha) - I| < \varepsilon$. ■

6.4 Other properties when integrator is increasing

Theorem 6.5 (Comparison theorem). Assume $\alpha \uparrow [a, b]$. If $f, g \in \mathcal{R}(\alpha, [a, b])$ and $f \leq g$ on $[a, b]$, then

$$\int_a^b f d\alpha \leq \int_a^b g d\alpha.$$

Theorem 6.6. Assume $\alpha \uparrow [a, b]$. If $f \in \mathcal{R}$ then $|f| \in \mathcal{R}$ and

$$\left| \int_a^b f d\alpha \right| \leq \int_a^b |f| d\alpha.$$

Proof. On k -th subinterval $[x_{k-1}, x_k]$, we can find t_k and t'_k such that for every $h > 0$

$$\sup|f| - \inf|f| - h < |f(t_k)| - |f(t'_k)| \leq |f(t_k) - f(t'_k)| \leq \sup f - \inf g.$$

This implies

$$U(P, |f|, \alpha) - L(P, |f|, \alpha) \leq U(P, f, \alpha) - L(P, f, \alpha).$$

Given arbitrary $\varepsilon > 0$, we can find P_ε such that every finer P will satisfies

$$U(P, |f|, \alpha) - L(P, |f|, \alpha) < \varepsilon.$$

That is $|f|$ satisfies Riemann's condition, and therefore $|f| \in \mathcal{R}$. Since $-|f| \leq f \leq |f|$, applying comparison theorem, we have

$$-\int_a^b |f| d\alpha \leq \int_a^b f d\alpha \leq \int_a^b |f| d\alpha.$$

This is the inequality we want. ■

Theorem 6.7. Assume $\alpha \uparrow [a, b]$. If $f, g \in \mathcal{R}(\alpha, [a, b])$ then the product $fg \in \mathcal{R}(\alpha, [a, b])$.

Proof. Consider a partition $P \in \mathcal{P}$, on the k -th subinterval $[x_{k-1}, x_k]$, for every $h > 0$, we can find some $t_k, t'_k \in [x_{k-1}, x_k]$ such that

$$\sup fg - \inf fg - h < (fg)(t_k) - (fg)(t'_k).$$

Since $f, g \in \mathcal{R}(\alpha, [a, b])$ as a normal Riemann-Stieltjes integral, they must be bounded, say $|f|, |g| \leq M$ without loss of generality. Now

$$\begin{aligned} (fg)(t_k) - (fg)(t'_k) &= [f(t_k) - f(t'_k)]g(t_k) + f(t'_k)[g(t_k) - g(t'_k)] \\ &\leq [\sup f - \inf f]M + M[\sup g - \inf g]. \end{aligned}$$

Since h is arbitrary, for every subinterval $[x_{k-1}, x_k]$ we have

$$\sup fg - \inf fg \leq M(\sup f - \inf f) + M(\sup g - \inf g).$$

Therefore

$$U(P, fg, \alpha) - L(P, fg, \alpha) \leq M[U(P, f, \alpha) - L(P, f, \alpha)] + M[U(P, g, \alpha) - L(P, g, \alpha)].$$

For every $\varepsilon > 0$, we can find P'_ε and P''_ε such that every $P' \supseteq P'_\varepsilon$ and every $P'' \supseteq P''_\varepsilon$ will satisfy

$$U(P', f, \alpha) - L(P', f, \alpha) < \varepsilon/2M, \quad U(P'', g, \alpha) - L(P'', g, \alpha) < \varepsilon/2M.$$

Let $P_\varepsilon = P'_\varepsilon \cup P''_\varepsilon$, then for every P finer than P_ε , we have $P \supseteq P'_\varepsilon$ and $P \supseteq P''_\varepsilon$, so

$$U(P, fg, \alpha) - L(P, fg, \alpha) < M \frac{\varepsilon}{2M} + M \frac{\varepsilon}{2M} = \varepsilon.$$

This means that fg satisfies Riemann's condition. Hence $fg \in \mathcal{R}(\alpha, [a, b])$. ■

As a direct corollary, $f \in \mathcal{R}$ implies $f^2 \in \mathcal{R}$.

7 Integrators of bounded variation

Given α as integrator on $[a, b]$. Define the function of total variation with respect to right endpoint $V(x) = V_a^x(\alpha)$ and $V(a) = 0$. Recall that $V \uparrow [a, b]$. Also for a partition $P = \{x_0, \dots, x_n\}$ if we use $\sum_P |\Delta f|$ to denote the sum $\sum_{k=1}^n |\Delta f_k|$ for this partition, then adding a point to P will only increase this sum. This implies that for $Q \supseteq P$, we have $\sum_Q |\Delta f| \geq \sum_P |\Delta f|$.

If $\alpha \in \text{BV}$, then there is a decomposition $\alpha = \alpha_1 - \alpha_2$ where α_1, α_2 are increasing. So if $f \in \mathcal{R}(\alpha_1) \cap \mathcal{R}(\alpha_2)$, then by linearity $f \in \mathcal{R}(\alpha)$. Conversely, if $f \in \mathcal{R}(\alpha)$ and we have decomposition $\alpha = \alpha_1 - \alpha_2$, then there is a chance that both $f \notin \mathcal{R}(\alpha_1)$ and $f \notin \mathcal{R}(\alpha_2)$. But as the following theorem tells us, $f \in \mathcal{R}(\alpha)$ always implies $f \in \mathcal{R}(V)$. So at least there is a decomposition $\alpha = V - (V - \alpha)$, such that $f \in \mathcal{R}(V)$ and $f \in \mathcal{R}(V - \alpha)$. Moreover V and $V - \alpha$ are increasing, so the original Riemann-Stieltjes integral with BV integrator is now decomposed to two integrals with increasing integrators, and Riemann's condition can now be applied to the decomposed integrals.

Theorem 7.1. *If f is bounded, $f \in \mathcal{R}(\alpha, [a, b])$ and $\alpha \in \text{BV}[a, b]$, then $f \in \mathcal{R}(V, [a, b])$.*

Proof. If $V(b) = 0$, then the proof is trivial. Suppose $V(b) > 0$ and $|f| \leq M$ on $[a, b]$. Since $V \uparrow$, the idea is to utilize Riemann's condition to prove $f \in \mathcal{R}(V)$. This involves estimating $U(P, f, V) - L(P, f, V)$ of partition P which can be split into two parts

$$\begin{aligned} \sum_k [M_k(f) - m_k(f)] \Delta V_k &= \sum_k [M_k(f) - m_k(f)] [\Delta V_k - |\Delta \alpha_k|] \\ &\quad + \sum_k [M_k(f) - m_k(f)] |\Delta \alpha_k|. \end{aligned}$$

The idea to bring in the $|\Delta \alpha_k|$ term is that the definition of BV function and total variation involves the sum $\sum_P |\Delta \alpha|$.

To estimate the first sum, we have

$$\begin{aligned} \sum_k [M_k(f) - m_k(f)] [\Delta V_k - |\Delta \alpha_k|] &\leq 2M \sum_k [\Delta V_k - |\Delta \alpha_k|] \\ &= 2M \left(V(b) - \sum_k |\Delta \alpha_k| \right). \end{aligned}$$

For every $h > 0$, we can find P_ε such that $V(b) - h < \sum_{P_\varepsilon} |\Delta \alpha|$. Refinement of P_ε will only increase the sum on the right, thus for every finer P , we have

$$V(b) < \sum_P |\Delta \alpha| + h.$$

So the first sum can be controlled by h chosen by us, and can be made arbitrarily small. This is what we intended.

$$\sum_k [M_k(f) - m_k(f)] [\Delta V_k - |\Delta \alpha_k|] \leq 2Mh. \quad (4)$$

Now to estimate the second sum, let $A(P) = \{k: \Delta \alpha_k \geq 0\}$ and $B(P) = \{k: \Delta \alpha_k < 0\}$. If we choose t_k and t'_k so that

$$\begin{aligned} M_k(f) - m_k(f) - h' &< f(t_k) - f(t'_k), \quad \text{if } k \in A(P), \\ M_k(f) - m_k(f) - h' &< f(t'_k) - f(t_k), \quad \text{if } k \in B(P). \end{aligned}$$

The second sum is now bounded by

$$\begin{aligned}
\sum_k [M_k(f) - m_k(f)] |\Delta\alpha_k| &\leq \left(\sum_{k \in A(P)} + \sum_{k \in B(P)} \right) [M_k(f) - m_k(f)] |\Delta\alpha_k| \\
&= \sum_{k \in A(P)} [f(t_k) - f(t'_k)] \Delta\alpha_k + h' \sum_{k \in A(P)} \Delta\alpha_k \\
&\quad - \sum_{k \in B(P)} [f(t'_k) - f(t_k)] \Delta\alpha_k - h' \sum_{k \in P(B)} \Delta\alpha_k.
\end{aligned}$$

After combining terms on the right, we have

$$\sum_k [M_k(f) - m_k(f)] |\Delta\alpha_k| \leq \sum_k [f(t_k) - f(t'_k)] \Delta\alpha_k + h' \sum_k |\Delta\alpha_k|. \quad (5)$$

Now given arbitrary $\varepsilon > 0$, first choose $h = \varepsilon/2M$, and find P_ε so that (4) holds for every finer partition.

Since $\alpha \in \text{BV}[a, b]$, $\sum_k |\Delta\alpha_k| \leq V(b)$. And since $f \in \mathcal{R}(\alpha, [a, b])$ we can find P'_ε so that for finer P and every choice of t_k, t'_k will satisfy

$$\left| \sum_k [f(t_k) - f(t'_k)] \Delta\alpha_k \right| < \varepsilon/4.$$

Now choose $h' = \varepsilon/2V(b)$. Find P'_ε and corresponding choice of t_k, t'_k so that (5) holds. Therefore for every P that is finer than both P_ε and P'_ε , inequalities (4) and (5) both hold. For such P

$$\begin{aligned}
\sum_k [M_k(f) - m_k(f)] \Delta V_k &\leq 2Mh + \sum_k [f(t_k) - f(t'_k)] \Delta\alpha_k + h'V(b) \\
&< 2M \frac{\varepsilon}{2M} + \frac{\varepsilon}{4} + \frac{\varepsilon}{2V(b)} V(b) = \varepsilon.
\end{aligned}$$

This means that $\int_a^b f dV$ satisfies Riemann's condition. ■

Theorem 7.2. Assume $\alpha \in \text{BV}[a, b]$ and $f \in \mathcal{R}(\alpha)$ on $[a, b]$, then $f \in \mathcal{R}(\alpha)$ on every subinterval $[c, d]$.

Proof. We can decompose the integral with BV integrator α to two integrals with increasing integrators V and $V - \alpha$:

$$\int_a^b f d\alpha = \int_a^b f dV - \int_a^b f d(V - \alpha).$$

Therefore it suffices to prove that: given $\alpha \uparrow [a, b]$, if $f \in \mathcal{R}(\alpha, [a, b])$, then $f \in \mathcal{R}(\alpha, [c, d])$. By additivity, we only need to show that $f \in \mathcal{R}(\alpha)$ on both $[a, c]$ and $[a, d]$.

Consider arbitrary $x \in [a, b]$. Since $f \in \mathcal{R}(\alpha, [a, b])$, there is a partition $P'_\varepsilon \in \mathcal{P}[a, b]$, such that every finer partition P will satisfy

$$U(P, [a, b]) - L(P, [a, b]) < \varepsilon.$$

Here we explicitly include the interval which the partition applies to. Let $P_\varepsilon = P'_\varepsilon \cup \{x\}$, $P_1 = P_\varepsilon \cap [a, x]$ and $P_2 = P_\varepsilon \cap [x, b]$. For every $P \in \mathcal{P}[a, x]$ finer than P_1 , we have $P \cup P_2 \supseteq P_1 \cup P_2 = P_\varepsilon$. We choose the same t_k in the subintervals of P for $U(P, [a, x]) - L(P, [a, x])$ and $U(P \cup P_2, [a, b]) - L(P \cup P_2, [a, b])$, then

$$U(P, [a, x]) - L(P, [a, x]) \leq U(P \cup P_2, [a, b]) - L(P \cup P_2, [a, b])$$

since the righthand side contains all terms of $U(P, [a, x]) - L(P, [a, x])$ and more terms corresponding to P_2 . And furthermore,

$$U(P \cup P_2, [a, b]) - L(P \cup P_2, [a, b]) \leq U(P_\varepsilon, [a, b]) - L(P_\varepsilon, [a, b]) < \varepsilon.$$

This implies that for every $P \in \mathcal{P}[a, x]$ finer than P_1 ,

$$U(P, [a, x]) - L(P, [a, x]) < \varepsilon.$$

So f satisfies Riemann's condition on $[a, x]$ and therefore is integrable. Hence $f \in \mathcal{R}(\alpha, [c, d])$ for arbitrary $[c, d] \subseteq [a, b]$. ■

8 Existence of Riemann-Stieltjes integrals

8.1 Sufficient conditions

Theorem 8.1 (Riemann-Stieltjes). *If $f \in C$ and $\alpha \in BV$, then $f \in \mathcal{R}(\alpha)$.*

Proof. Since $\alpha \in BV$, we can decompose $\int_a^b f d\alpha$ into two integrals with increasing integrators. It suffices to prove this theorem under conditions $f \in C[a, b]$ and $\alpha \uparrow [a, b]$. Also when $\alpha(a) = \alpha(b)$ the proof is trivial. So suppose $\alpha(a) < \alpha(b)$.

Since $f \in C[a, b]$, then f must be uniformly continuous on every subinterval $[x_{k-1}, x_k]$ of arbitrary partition. Given $\varepsilon > 0$, we can find $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon/2[\alpha(b) - \alpha(a)]$.

There is $P_\varepsilon \in \mathcal{P}[a, b]$ with $\|P_\varepsilon\| < \delta$, for every P finer than P_ε , we have $\|P\| \leq \|P_\varepsilon\| < \delta$. Therefore for every subinterval $[x_{k-1}, x_k]$, every $x, y \in [x_{k-1}, x_k]$ will satisfies

$$f(x) - f(y) \leq |f(x) - f(y)| < \frac{\varepsilon}{2[\alpha(b) - \alpha(a)]}.$$

Take sup of both sides. For every subinterval, we have

$$M_k(f) - m_k(f) = \sup [f(x) - f(y)] \leq \frac{\varepsilon}{2[\alpha(b) - \alpha(a)]}.$$

Then for every $P \supseteq P_\varepsilon$,

$$U(P, f, \alpha) - L(P, f, \alpha) = \sum_k [M_k(f) - m_k(f)] \Delta\alpha_k \leq \frac{\varepsilon}{2} < \varepsilon.$$

That is f satisfies Riemann's condition, and therefore $f \in \mathcal{R}(\alpha, [a, b])$. ■

Riemann integral is a special type of Riemann-Stieltjes integral where $\alpha(x) = x$, so

Theorem 8.2 (Riemann). *Sufficient conditions for existence of $\int_a^b f(x) dx$:*

1. $f \in C[a, b]$
2. $f \in BV[a, b]$

Proof. (1) Since $\alpha(x) = x$ is of bounded variation.

(2) Integration by parts: $\int_a^b f d\alpha = [f\alpha]_a^b - \int_a^b \alpha df$. ■

8.2 Necessary conditions

When $\alpha \in \text{BV}$, the continuity of f is sufficient for existence of the integral. But for the integral to exist, continuity of f is not necessary.

Example (Integral exists, but f is not continuous). *Consider a simple 2-step function α which has jump discontinuity at single point $c \in (a, b)$. As long as f is continuous at c , then by theorem 5.1 the integral must exist:*

$$\int_a^b f d\alpha = f(c)[f(c+) - f(c-)],$$

even if f is discontinuous elsewhere.

At least one of f and α is left-continuous and at least one of them is right-continuous on $[a, b]$ is the necessary condition for the integral to exist.

Theorem 8.3. *Assume $\alpha \uparrow [a, b]$ and $c \in (a, b)$. If f and α are both right-discontinuous at c or both left-discontinuous, then $\int_a^b f d\alpha$ does not exist.*

Proof. Suppose both f and α are right-discontinuous at $c \in (a, b)$. Proof for left-discontinuous case is similar.

Let P be arbitrary partition of $[a, b]$, and $P' = P \cup \{c\} = \{x'_0, x'_1, \dots, x'_n\}$. Suppose $c = x'_{i-1}$ in P' . Fix some $\varepsilon > 0$, since α is right-discontinuous at c , then let $\delta' < x'_i - x'_{i-1}$, we can find $y \in (c, c + \delta')$ such that

$$|\alpha(y) - \alpha(c)| \geq \varepsilon.$$

Adding y into P' , we get a new partition $P'' = P' \cup y = \{x_0, x_1, \dots, x_{n+1}\}$, and now $c = x_{i-1} = x'_{i-1}$, $y = x_i$. For P'' , we have

$$U(P'', f, \alpha) - L(P'', f, \alpha) \geq [M_i(f) - m_i(f)][\alpha(y) - \alpha(c)].$$

Here $M_i(f) = \sup f([c, y])$ and $m_i(f) = \inf f([c, y])$. Since f is right-discontinuous at c , we can find $x \in (c, y)$ such that

$$|f(x) - f(c)| \geq \varepsilon.$$

This implies $M_i(f) - m_i(f) \geq \varepsilon$. Because $P \supseteq P''$,

$$U(P, f, \alpha) - L(P, f, \alpha) \geq U(P'', f, \alpha) - L(P'', f, \alpha) \geq \varepsilon^2.$$

So f does not satisfy Riemann's condition, hence $\int_a^b f d\alpha$ does not exist. ■

9 Mean value theorems for integrals

Theorem 9.1 (First MVT). *Assume $\alpha \uparrow [a, b]$ and $f \in \mathcal{R}(\alpha, [a, b])$. Then there exists a real number $c \in [\inf f, \sup f]$ such that*

$$\int_a^b f d\alpha = c \int_a^b d\alpha = c[\alpha(b) - \alpha(a)].$$

Proof. If $\alpha(b) = \alpha(a)$ then the proof is done. Assume $\alpha(b) > \alpha(a)$. For every partition P , we have

$$\inf f \cdot [\alpha(b) - \alpha(a)] \leq L(P, f, \alpha) \leq \underline{I} \leq \bar{I} \leq U(P, f, \alpha) \leq \sup f \cdot [\alpha(b) - \alpha(a)].$$

Since $f \in \mathcal{R}$, then $\int_a^b f d\alpha = \underline{I} = \bar{I}$. Therefore

$$\inf f \leq \int_a^b f d\alpha / [\alpha(b) - \alpha(a)] \leq \sup f.$$

So we have found $c = \int_a^b f d\alpha / \int_a^b d\alpha \in [\inf f, \sup f]$. ■

Using integration by parts, we can obtain the extended second MVT.

Theorem 9.2 (Second MVT). *Assume $\alpha \in C[a, b]$ and $f \uparrow [a, b]$. Then there exists $\xi \in [a, b]$ such that*

$$\int_a^b f d\alpha = f(a) \int_a^\xi d\alpha + f(b) \int_\xi^b d\alpha.$$

Proof. Note that α is continuous, therefore the number $c \in [\inf f, \sup f]$ is attained by f , so c can be written as $f(\xi)$.

$$\begin{aligned} \int_a^b f d\alpha &= [\alpha f]_a^b - \int_a^b \alpha df \\ &= [\alpha f]_a^b - \alpha(\xi) \int_a^b df \\ &= f(a) [\alpha(\xi) - \alpha(a)] + f(b) [\alpha(b) - \alpha(\xi)] \\ &= f(a) \int_a^\xi d\alpha + f(b) \int_\xi^b d\alpha. \end{aligned}$$

■

Note the difference between two MVT theorems:

- First MVT: $\alpha \uparrow$, mean value taken between $[\inf f, \sup f]$
- Second MVT: $f \uparrow$ and $\alpha \in C$, mean value taken between interval $[a, b]$

10 Integral with variable upper limit

Suppose $f \in \mathcal{R}(\alpha, [a, b])$ and $\alpha \in \text{BV}$, then theorem 7.2 told us f is integrable on arbitrary subinterval $[c, d] \subseteq [a, b]$. Therefore for every $x \in [a, b]$, the integral $\int_a^x f d\alpha$ exists and can be seen as a function of x . Define

$$F(x) = \int_a^x f d\alpha.$$

F is closely related to f and α .

Theorem 10.1. *Assume $f \in \mathcal{R}(\alpha, [a, b])$ and $\alpha \in \text{BV}[a, b]$. Then for F defined above, we have:*

1. $F \in \text{BV}[a, b]$
2. α continuous at a point implies F continuous at this point
3. Suppose $\alpha \uparrow [a, b]$. At $x \in (a, b)$ where f is continuous and α is differentiable, F' exists and

$$F'(x) = f(x)\alpha'(x)$$

Proof. It suffice to assume $\alpha \uparrow [a, b]$. Let $M = \sup f([a, b])$, by MVT,

$$F(y) - F(x) = \int_x^y f d\alpha = c [\alpha(y) - \alpha(x)] \leq M [\alpha(y) - \alpha(x)].$$

This proves (1) and (2). Divided by $y - x$, we have

$$\frac{F(y) - F(x)}{y - x} = c \cdot \frac{\alpha(y) - \alpha(x)}{y - x}.$$

Now let $y \rightarrow x$, note $c \rightarrow f(x)$.

■

For Riemann integral, we have $\alpha(x) = x$. This theorem gives us the first fundamental theorem of calculus:

1. F is continuous on $[a, b]$
2. F is differentiable on (a, b) and $F'(x) = f(x)$

Theorem 10.2. Assume $f, g \in \mathcal{R}(\alpha, [a, b])$, and $\alpha \uparrow [a, b]$. Define F and G like above. Then $f \in \mathcal{R}(F)$ and $g \in \mathcal{R}(G)$ on $[a, b]$, also $fg \in \mathcal{R}(\alpha)$ on $[a, b]$, they satisfy

$$\int_a^b fg d\alpha = \int_a^b f dG = \int_a^b g dF.$$

Proof. We only prove $f \in \mathcal{R}(G)$ on $[a, b]$. Proof of $g \in \mathcal{R}(F)$ is analogous. Since α is monotonic, it is of bounded variation. Then f is integrable on every subinterval $[c, d] \subseteq [a, b]$. We have

$$S(P, f, G) = \sum_k f(t_k)[G(x_k) - G(x_{k-1})] = \sum_k \int_{x_{k-1}}^{x_k} f(t_k)g(t) d\alpha(t),$$

and

$$\int_a^b fg d\alpha = \sum_k \int_{x_{k-1}}^{x_k} f(t)g(t) d\alpha(t).$$

Let M be an upper bound of g , that is $|g| \leq M$ on $[a, b]$, then

$$\begin{aligned} \left| S(P, f, G) - \int_a^b fg d\alpha \right| &= \left| \sum_k \int_{x_{k-1}}^{x_k} [f(t_k) - f(t)]g(t) d\alpha(t) \right| \\ &\leq \sum_k \int_{x_{k-1}}^{x_k} |f(t_k) - f(t)| |g(t)| d\alpha(t) \\ &\leq M \cdot \sum_k \int_{x_{k-1}}^{x_k} [M_k(f) - m_k(f)] d\alpha(t). \end{aligned}$$

Note that $M_k(f) - m_k(f)$ is a constant on $[x_{k-1}, x_k]$, therefore

$$\int_{x_{k-1}}^{x_k} [M_k(f) - m_k(f)] d\alpha(t) = [M_k(f) - m_k(f)] \Delta\alpha_k.$$

Finally,

$$\left| S(P, f, G) - \int_a^b fg d\alpha \right| \leq M \sum_k [M_k(f) - m_k(f)] \Delta\alpha_k = M[U(P, f, \alpha) - L(P, f, \alpha)].$$

Since $f \in \mathcal{R}(\alpha)$, by Riemann's condition $U - L$ can be made arbitrarily small. This implies that $f \in \mathcal{R}(G)$ on $[a, b]$. ■

11 Second fundamental theorem of calculus

Theorem 11.1 (Second fundamental theorem of calculus). Assume $f \in \mathcal{R}[a, b]$. Let F be a function such that $F \in D(a, b) \cap C[a, b]$ and $F'(x) = f(x)$ then

$$\int_a^b f(x) dx = F(b) - F(a).$$

Proof. Since $f \in \mathcal{R}[a, b]$, by Riemann's condition for partition P sufficiently fine, we have

$$U(P, f) - L(P, f) < \varepsilon.$$

Therefore, using MVT, we have $F(x_k) - F(x_{k-1}) = f(\xi_k)\Delta x_k$ with $\xi_k \in (x_{k-1}, x_k)$, then

$$\begin{aligned} |S(P, f) - [F(b) - F(a)]| &= \left| \sum_k f(t_k)\Delta x_k - \sum_k [F(x_k) - F(x_{k-1})] \right| \\ &= \left| \sum_k [f(t_k) - f(\xi_k)]\Delta x_k \right| \\ &\leq U(P, f) - L(P, f) < \varepsilon. \end{aligned}$$

By definition of Riemann integral, $\int_a^b f(x) dx = F(b) - F(a)$. ■

In theorem 4.1 that reduces Riemann-Stieltjes integrals to Riemann integrals, α' is required to be continuous on $[a, b]$. Now we can relax this condition by only requiring α' to be integrable on $[a, b]$.

Theorem 11.2. *If $f \in \mathcal{R}(\alpha, [a, b])$, $\alpha \in C[a, b]$ and $\alpha' \in \mathcal{R}[a, b]$, then*

$$\int_a^b f d\alpha = \int_a^b f(x)\alpha'(x) dx.$$

Proof. Let $g = \alpha'$ and $G(x) = \int_a^x g(x) dx = \alpha(x) - \alpha(a)$. Then

$$\begin{aligned} \int_a^b f(x)\alpha'(x) dx &= \int_a^b f(x) dG(x) = \int_a^b f(x) d(\alpha(x) - \alpha(a)) \\ &= \int_a^b f(x) d\alpha(x) - \int_a^b f(x) d(\alpha(a)) \\ &= \int_a^b f d\alpha. \end{aligned}$$
■

Theorem 11.2 says that we can differentiate the integrator and move it to integrand. While theorem 10.2 says that we can integrate part of integrand and move it to integrator.

12 More on Riemann integral

12.1 Change of variable

If we want to apply theorem 3.1 to Riemann integrals to get change of variable theorem for this type, we assume that g' is continuous on $[c, d]$. Also theorem 3.1 requires g to be strictly monotonic. We can remove this condition, so g need not to be monotonic (injective).

Theorem 12.1 (Change of variable for Riemann integrals). *Assume g is continuously differentiable on $[c, d]$. Let f be continuous on $g([c, d])$.*

$$\int_{g(c)}^{g(d)} f(x) dx = \int_c^d f(g(t))g'(t) dt.$$

Proof. Since $f \circ g$ and g' are continuous, $(f \circ g) \cdot g'$ is also continuous, the integral $\int_c^d f(g(t))g'(t) dt$ exists. For every $x \in [c, d]$, the same integral on subinterval $[c, x]$ still exists since integrator (t) is of bounded variation.

Define F on $g([c, d])$ by

$$F(x) = \int_{g(c)}^x f(t) dt.$$

We have $F'(x) = f(x)$ for $x \in g([c, d])$, thus by first fundamental theorem of calculus

$$\int_c^x f(g(t))g'(t) dt = \int_c^x [F(g(t))]' dt = F(g(x)).$$

Take $x = d$ and done. ■

Two sets of conditions for change of variable:

1. $g \in C^1[c, d]$ and strictly monotonic
2. $g \in C^1[c, d]$ and $f \in C^0(g([c, d]))$

12.2 Second MVT

Theorem 12.2. Suppose $f \uparrow [a, b]$ and $g \in C[a, b]$. If $A, B \in \mathbb{R}$ satisfies $A \leq f(a+)$ and $B \geq f(b-)$, then

1. there is $\xi \in [a, b]$ such that

$$\int_a^b f(x)g(x) dx = A \int_a^\xi g(x) dx + B \int_\xi^b g(x) dx.$$

2. (Bonnet's theorem) If $f \geq 0$, choose $A = 0$ and we have

$$\int_a^b f(x)g(x) dx = B \int_\xi^b g(x) dx.$$

Proof. Let $G(x) = \int_a^x g(x) dx$, by first fundamental theorem of calculus, $G \in C[a, b]$. If we redefine f so that $f(a) = A$ and $f(b) = B$, then f is still increasing, and the integral $\int_a^b f(x)g(x) dx$ will not change since we only changed the value of fg at two points. So if we now use second MVT for Riemann-Stieltjes integral:

$$\begin{aligned} \int_a^b f(x)g(x) dx &= \int_a^b f(x) dG(x) = f(a) \int_a^\xi dG(x) + f(b) \int_\xi^b dG(x) \\ &= A \int_a^\xi g(x) dx + B \int_\xi^b g(x) dx. \end{aligned}$$
■

13 Riemann-Stieltjes integral with parameter

Theorem 13.1 (Continuity of integral). Let f be continuous on $[a, b] \times [c, d]$. Assume $\alpha \in \text{BV}[a, b]$. Define F on $[c, d]$ by

$$F(y) = \int_a^b f(x, y) d\alpha(x).$$

Then $F \in C[a, b]$. Suppose $y_0 \in [c, d]$ then we can exchange limit and integral:

$$\lim_{y \rightarrow y_0} \int_a^b f(x, y) d\alpha(x) = \int_a^b \lim_{y \rightarrow y_0} f(x, y) d\alpha(x) = \int_a^b f(x, y_0) d\alpha(x).$$

Proof. Assume $\alpha \uparrow$. Since f is continuous on $[a, b] \times [c, d]$, a closed set, f is also uniformly continuous on $[a, b] \times [c, d]$. Given arbitrary $\varepsilon > 0$, we can find $\delta > 0$ such that $\|(x, y) - (x', y')\| < \delta$ implies $|f(x, y) - f(x', y')| < \varepsilon/[\alpha(b) - \alpha(a)]$. Therefore for every y such

that $|y - y_0| < \delta$, we have $\|(x, y) - (x, y_0)\| < \delta$, and $|f(x, y) - f(x, y_0)| < \varepsilon / [\alpha(b) - \alpha(a)]$. Now we have

$$\begin{aligned} |F(y) - F(y_0)| &= \left| \int_a^b f(x, y) - f(x, y_0) d\alpha(x) \right| \leq \int_a^b |f(x, y) - f(x, y_0)| d\alpha(x) \\ &< \frac{\varepsilon}{\alpha(b) - \alpha(a)} \int_a^b d\alpha(x) = \varepsilon. \end{aligned}$$

This shows that F is continuous at arbitrary $y_0 \in [c, d]$. Hence $f \in C[c, d]$. ■

13.1 Differentiation under the integral sign

Theorem 13.2 (Differentiation of integral). *Let f be defined on $[a, b] \times [c, d]$. Suppose $\alpha \in \text{BV}[a, b]$, define F on $[c, d]$ by $F(y) = \int_a^b f(x, y) d\alpha(x)$. If D_2f is continuous on $[a, b] \times [c, d]$, then F is differentiable on (c, d) with*

$$F'(y) = \int_a^b D_2f(x, y) d\alpha(x).$$

Proof. Consider $y_0 \in (c, d)$,

$$\begin{aligned} F'(y) &= \lim_{y \rightarrow y_0} \frac{F(y) - F(y_0)}{y - y_0} = \lim_{y \rightarrow y_0} \int_a^b \frac{f(x, y) - f(x, y_0)}{y - y_0} d\alpha(x) \\ &= \lim_{y \rightarrow y_0} \int_a^b D_2f(x, \xi) d\alpha(x) \\ &= \int_a^b \lim_{y \rightarrow y_0} D_2f(x, \xi) d\alpha(x) \\ &= \int_a^b D_2f(x, y_0) d\alpha(x). \end{aligned}$$

■

A more rigorous proof is similar to the previous theorem about continuity.

After using MVT we want to prove $\lim_{y \rightarrow y_0} \int_a^b D_2f(x, \xi) d\alpha(x)$ is $\int_a^b D_2f(x, y_0) d\alpha(x)$. Since D_2f is continuous on $[a, b] \times [c, d]$, it is also uniformly continuous. For every $\varepsilon > 0$, we can find $\delta > 0$, such that for every (x', y') and (x'', y'') in $[a, b] \times [c, d]$ with $\|(x', y') - (x'', y'')\| < \delta$ we have:

$$|D_2f(x', y') - D_2f(x'', y'')| < \varepsilon / [\alpha(b) - \alpha(a)].$$

Therefore if $|y - y_0| < \delta$, then we also have $\|(x, \xi) - (x, y_0)\| < \delta$, hence

$$|D_2f(x, \xi) - D_2f(x, y_0)| < \frac{\varepsilon}{\alpha(b) - \alpha(a)}.$$

And

$$\begin{aligned} \left| \int_a^b D_2f(x, \xi) d\alpha(x) - \int_a^b D_2f(x, y_0) d\alpha(x) \right| &\leq \int_a^b |D_2f(x, \xi) - D_2f(x, y_0)| d\alpha(x) \\ &< \frac{\varepsilon}{\alpha(b) - \alpha(a)} \int_a^b d\alpha(x) = \varepsilon. \end{aligned}$$

13.2 Interchanging the order of integration

Theorem 13.3. Let $f \in C([a, b] \times [c, d])$, $\alpha \in \text{BV}[a, b]$, $\beta \in \text{BV}[c, d]$. Define

$$F(y) = \int_a^b f(x, y) d\alpha(x), \quad G(x) = \int_a^b f(x, y) d\beta(y).$$

Then

$$\int_c^d F(y) d\beta(y) = \int_a^b G(x) d\alpha(x).$$

In other words, we can change the order of integration:

$$\int_a^b \int_c^d f(x, y) d\beta(y) d\alpha(x) = \int_c^d \int_a^b f(x, y) d\alpha(x) d\beta(y).$$

Let $Q = [a, b] \times [c, d]$. The idea is to subdivide $[a, b]$ and $[c, d]$ both into n subintervals, getting n^2 subrectangles Q_{ij} . The integral can be split into n^2 integrals on Q_{ij} , which can be estimated by f at some point in Q_{ij} using MVT for integrals. Since f is continuous on each compact Q_{ij} , it must be uniformly continuous. The distance between $f(x', y')$ and $f(x'', y'')$ can be made arbitrarily small, if we subdivide $[a, b] \times [c, d]$ fine enough. That's how we control the difference $\int_a^b \int_c^d - \int_c^d \int_a^b$.

Proof. It suffices to assume $\alpha, \beta \uparrow$. Let $Q = [a, b] \times [c, d]$. Given arbitrary $\varepsilon > 0$, continuity of f on Q implies uniform continuity, hence there is $\delta > 0$ such that $|f(x', y') - f(x'', y'')| < \varepsilon$ as long as $\|(x', y') - (x'', y'')\| < \delta$.

Subdivide $[a, b]$ and $[c, d]$ both into n subintervals such that each subinterval has length less than $\delta/2$. Suppose partition of $[a, b]$ is $\{x_0, \dots, x_n\}$ and partition of $[c, d]$ is $\{y_0, \dots, y_n\}$. Let the Q_{ij} be the subrectangle $[x_{i-1}, x_i] \times [y_{j-1}, y_j]$, $i = 1, \dots, n$ and $j = 1, \dots, n$.

The integral $\int_a^b G(x) d\alpha(x)$ can be split into n^2 integrals on each Q_{ij} :

$$\int_a^b \int_c^d f(x, y) d\beta(y) d\alpha(x) = \sum_{i=1}^n \sum_{j=1}^n \int_{x_{i-1}}^{x_i} \int_{y_{j-1}}^{y_j} f(x, y) d\beta(y) d\alpha(x).$$

Using first MVT for integral twice:

$$\begin{aligned} \int_a^b G(x) d\alpha(x) &= \sum_{i=1}^n \sum_{j=1}^n \int_{x_{i-1}}^{x_i} f(x, y'_j) \Delta\beta_j d\alpha(x) \\ &= \sum_{i=1}^n \sum_{j=1}^n f(x'_i, y'_j) \Delta\alpha_i \Delta\beta_j. \end{aligned}$$

Here $(x'_i, y'_j) \in Q_{ij}$. Similarly we have

$$\int_c^d F(y) d\beta(y) = \sum_{i=1}^n \sum_{j=1}^n f(x''_i, y''_j) \Delta\alpha_i \Delta\beta_j \quad (x''_i, y''_j) \in Q_{ij}.$$

For every $i \in [1..n]$ and every $j \in [1..n]$, we have $\|(x'_i, y'_j) - (x''_i, y''_j)\| < \delta$, hence $|f(x'_i, y'_j) - f(x''_i, y''_j)| < \varepsilon$. Therefore

$$\begin{aligned} \left| \int_a^b G(x) d\alpha(x) - \int_c^d F(y) d\beta(y) \right| &\leq \sum_{i=1}^n \sum_{j=1}^n |f(x'_i, y'_j) - f(x''_i, y''_j)| \Delta\alpha_i \Delta\beta_j \\ &< \varepsilon [\alpha(b) - \alpha(a)] [\beta(d) - \beta(c)]. \end{aligned}$$

This completes the proof. ■

14 Lebesgue's criterion for Riemann integrability

14.1 Measure zero

Definition 14.1 (Measure zero). A set $S \subseteq \mathbb{R}$ is said to have measure zero or to be null set if for every $\varepsilon > 0$, there is a countable covering of S by open intervals, the sum of whose lengths is less than ε . In other words, we can use open intervals that have arbitrarily small total length to cover S .

Countable union of null sets is a null set.

Theorem 14.1. Let $F = \{F_1, F_2, \dots\}$ where each F_i has measure zero, then $\cup F$ has measure zero.

Proof. For every $\varepsilon > 0$, each F_k can be covered by a countable set of open intervals S_k , and the total length of these intervals is less than $\varepsilon/2^k$. Now $S = \bigcup_{k=1}^{\infty} S_k$ is a countable set of open intervals. S covers F and the total length of intervals in S is less than

$$\sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon.$$

■

Example. Some example of null set:

- Finite sets of real numbers $\{x_1, \dots, x_n\}$
- Countable sets of real numbers $\{x_1, x_2, \dots\}$: since it's the countable union of sets with single real number
- Set of rational numbers \mathbb{Q} : since it's countable subset of \mathbb{R}
- Cantor set is an uncountable null set

14.2 Oscillation

Definition 14.2. Assume f is bounded on an interval S and $T \subseteq S$. The number

$$\Omega_f(T) = \sup\{f(x) - f(y) : x, y \in T\}$$

is call the oscillation of f on T . The oscillation at a point x is defined by

$$\omega_f(x) = \lim_{r \rightarrow 0+} \Omega_f(B(x, r) \cap S).$$

If $T \subseteq S$ then $\{f(x) - f(y) : x, y \in T\} \subseteq \{f(x) - f(y) : x, y \in S\}$, therefore $\Omega_f(T) \leq \Omega_f(S)$. This means $\Omega_f(B(x, r) \cap S)$ is a decreasing bounded function and hence $\omega_f(x)$ must exists.

Also we have $\omega_f(x) = 0$ iff. f is continuous at x .

Next theorem shows that if $\omega_f(x) < \varepsilon$ on compact interval, then $\Omega_f([c, d])$ can be made less than ε if the length is sufficiently small.

Theorem 14.2. Let f be bounded on $[a, b]$. Given $\varepsilon > 0$, assume that $\omega_f(x) < \varepsilon$ for $x \in [a, b]$. Then there exists $\delta > 0$ (depending only on ε), such that for every closed subinterval $[c, d] \subseteq [a, b]$, $d - c < \delta$ implies $\Omega_f([c, d]) < \varepsilon$.

Proof. For every $x \in [a, b]$ we can find $\delta_x > 0$ (depending on both x and ε) such that $\Omega_f(B(x, \delta_x) \cap [a, b]) < \varepsilon$.

Since $[a, b]$ is compact, by Heine-Borel theorem, $\{B(x, \delta_x/2) : x \in [a, b]\}$ as an open covering of $[a, b]$ can be reduced to finite subcover $\{B(x_1, \delta_{x_1}/2), \dots, B(x_n, \delta_{x_n}/2)\}$. Let $\delta = \min\{\delta_{x_1}/2, \dots, \delta_{x_n}/2\}$. For an arbitrary closed subinterval $[c, d] \subseteq [a, b]$ with $d - c < \delta$, some $B(x_i, \delta_{x_i}/2)$ must have intersection with it, so $[c, d]$ is contained in the bigger ball with same center $B(x_i, \delta_{x_i})$. And therefore $\Omega_f([c, d]) \leq \Omega_f(B(x_i, \delta_{x_i}) \cap [a, b]) < \varepsilon$. ■

Theorem 14.3. Let f be a bounded function on $[a, b]$. For $\varepsilon > 0$, define

$$J_\varepsilon = \{x \in [a, b] : \omega_f(x) \geq \varepsilon\}.$$

Then J_ε is a closed set for every ε .

For arbitrary limit point p , there is a sequence in J_ε converges to p , each element of the sequence has oscillation $\omega_f \geq \varepsilon$. It natural to think that the oscillation of p will be affected because points in the sequence can get arbitrarily close to p .

Proof. Given $\varepsilon > 0$, let p be a limit point of J_ε . For every $r > 0$, $B(p, r)$ contains some point $x \in J_\varepsilon$ different from p . For this x , since $\omega_f(x) \geq \varepsilon$, we can find $\delta > 0$ small enough so that $B(x, \delta) \subseteq B(p, r)$ and $\Omega_f(B(x, \delta) \cap [a, b]) \geq \varepsilon$. Therefore for every $r > 0$:

$$\Omega_f(B(p, r) \cap [a, b]) \geq \Omega_f(B(x, \delta) \cap [a, b]) \geq \varepsilon.$$

This implies that $\lim_{r \rightarrow 0+} \Omega_f(B(p, r) \cap [a, b]) \geq \varepsilon$, so $p \in J_\varepsilon$. Hence J_ε contains all its limit points and is closed. ■

14.3 Lebesgue's criterion

Theorem 14.4. Let f be a bounded function on $[a, b]$ and let D denote the set of discontinuities of f in $[a, b]$. Then $f \in \mathcal{R}[a, b]$ iff. D has measure zero.

We say a property holds *almost everywhere* on S if the set of points where this property doesn't hold has measure zero. "Almost everywhere" can also be abbreviated as "a.e." for short. Now Lebesgue's criterion can be rephrased: $f \in \mathcal{R}[a, b]$ if and only if f is continuous a.e. on $[a, b]$.

Proof sketch (\Rightarrow):

1. Suppose D is not null set, then some D_r is not null set
2. Show that Riemann's condition cannot be satisfied
 - (a) **Split** $U - L$ into $S_1 + S_2$ where subintervals in S_1 contains points of D_r
 - (b) Show that $S_1 \geq$ some positive number

Proof. (\Rightarrow) Let $D_r = \{x \in [a, b] : \omega_f(x) \geq 1/r\}$. Then

$$D = \bigcup_{r=1}^{\infty} D_r.$$

Suppose D does not have measure zero, then some D_r must not have measure zero by theorem 14.1. Thus there is some $\varepsilon > 0$, such that every collection of open intervals that covers D_r has total length $\geq \varepsilon$.

We'll show that for arbitrary partition $P \in \mathcal{P}[a, b]$, the difference $U(P, f) - L(P, f)$ can not be arbitrarily small. Split this difference into two parts:

$$U(P, f) - L(P, f) = S_1 + S_2,$$

where the subintervals $[x_{k-1}, x_k]$ in the sum S_1 contains points in D_r and S_2 contains the remaining terms. The corresponding (x_{k-1}, x_k) in S_1 covers all except possibly finite points of D_r (since they happened to be partition points). But the total length of these open intervals is still $\geq \varepsilon$, since finite many points can always be covered by intervals whose total length arbitrarily small. Now for S_1 we have

$$S_1 = \sum_k [M_k(f) - m_k(f)] \Delta x_k.$$

Inside, $M_k(f) - m_k(f) \geq 1/r$ since every subinterval contains points in D_r . Therefore

$$S_1 \geq \frac{1}{r} \sum_k \Delta x_k \geq \frac{\varepsilon}{r},$$

meaning Riemann's condition cannot be satisfied. ■

Proof sketch (\Leftarrow):

1. D is null set, then every D_r is null set
2. A is union of open intervals that covers D_r and the total length can be made smaller than arbitrarily ε' , since D_r is null set. The part of sum $U - L$ on A can be controlled by ε' we choosed
3. $B = [a, b] \setminus A$ is union of closed interval. B doesn't cover D_r so every point in B_r has oscillation $< 1/r$. If we subdivide B fine enough, every subinterval will have $M_k - m_k < 1/r$. Thus the part of sum $U - L$ on B can be made arbitrarily small for large r
4. Together $U - L$ can be made arbitrarily small

Proof. (\Leftarrow) Again we have D , D_r , and $D = \bigcup_{r=1}^{\infty} D_r$. Since D is null set, every D_r is also a null set.

Since D_r is compact (theorem 14.3), there are finite number of open intervals that covers D_r and the total length can be less than arbitrary ε' we choose. Let A_r be the open set that is the union of these finite many open intervals.

Let $B_r = [a, b] \setminus A_r$, then B_r is an union of finite number of closed sets. Every point in B_r has oscillation $< 1/r$, then by theorem 14.2, we can subdivide B_r to B'_r so that oscillation on every subinterval of B'_r is $< 1/r$.

Construct a partition from endpoints of B'_r . For every finer partition P ,

$$U(P, f) - L(P, f) = \sum_{A_r} [M_k(f) - m_k(f)] \Delta x_k + \sum_{B'_r} [M_k(f) - m_k(f)] \Delta x_k.$$

For sum on A_r , we have $\sum_k \Delta x_k$ can be less than $\varepsilon' = 1/r$ we choosed, so

$$\sum_{A_r} [M_k(f) - m_k(f)] \Delta x_k \leq (\sup f - \inf f) \sum_{A_r} \Delta x_k < \frac{\sup f - \inf f}{r}.$$

For sum on B'_r , we have $M_k(f) - m_k(f) < 1/r$, so

$$\sum_{B'_r} [M_k(f) - m_k(f)] \Delta x_k \leq \frac{1}{r} \sum_{B'_r} \Delta x_k < \frac{b - a}{r}.$$

Therefore for every $r > 0$,

$$U(P, f) - L(P, f) < \frac{\sup f - \inf f + b - a}{r},$$

this can be made arbitrarily small if we choose r large enough. Hence f satisfies Riemann's condition and is integrable. ■

Using Lebesgue's criterion, many properties of Riemann integral are now obvious:

1. $f \in \text{BV} \Rightarrow f \in \mathcal{R}$
2. $f \in \mathcal{R}[a, b]$ and $[c, d] \subseteq [a, b]$, then $f \in \mathcal{R}[c, d]$
3. $f \in \mathcal{R} \Rightarrow |f| \in \mathcal{R}$

4. $f, g \in \mathcal{R} \Rightarrow fg \in \mathcal{R}$. And $f/g \in \mathcal{R}$ if g is bounded away from 0
5. f, g bounded on $[a, b]$ and have same discontinuities, then $f \in \mathcal{R}$ iff. $g \in \mathcal{R}$
6. $g \in \mathcal{R}[a, b]$ and is bounded in a compact interval ($m \leq g \leq M$), f continuous on $[m, M]$, then $f \circ g \in \mathcal{R}[a, b]$

Example. (Using integrability of composite function) Suppose $f^3 \in \mathcal{R}$ on $[a, b]$, then $f \in \mathcal{R}$. Let $\phi(x) = x^{1/3}$.

This is because $m \leq f^3 \leq M$ and $\phi \in C^0([m, M])$. Therefore $f = \phi \circ f^3 \in \mathcal{R}$.

However $f^2 \in \mathcal{R}$ doesn't implies $f \in \mathcal{R}$. Consider the counter-example where

$$f(x) = \begin{cases} 1 & , x \in \mathbb{Q} \\ -1 & , x \notin \mathbb{Q} \end{cases}.$$

$f^2 = 1 \in \mathcal{R}$ but $f \notin \mathcal{R}$. Let $\psi(x) = \sqrt{x}$. Using the rule can only give us $|f| = \psi \circ f^2 \in \mathcal{R}$.