FOUNDATION OF ANALYSIS, LECTURE NOTES, FALL 2011
PART I: RIEMANN AND RIEMANN–STIELTJES INTEGRATION

It may be a comfort to students just beginning to work in the calculus to know that Newton and Leibniz, two of the greatest mathematicians, did not fully understand what they themselves had produced

M. Kline, Calculus

1. The Riemann integral

1.1. **Definition.** We consider a fixed bounded interval $[a,b]$ and a bounded (real-valued) function f on $[a,b]$. A finite set $\mathscr{C}$ of points of $[a,b]$ containing a and b is called a *partition* of $[a,b]$.

Let $\mathscr{C}=\{c_0,c_1,\dots,c_n\}$ be a partition of $[a,b]$, where $a=c_0<c_1<\dots<c_n=b$, then

$$\|\mathscr{C}\|=\max\{c_j-c_{j-1},\, j=1,\dots,n\}$$

is called the *mesh size* of $\mathscr{C}$. We say that $\eta=\{\eta_1,\dots,\eta_n\}$ is a *choice sequence* for $\mathscr{C}$ if $c_{j-1}\le\eta_j\le c_j$ for $j=1,\dots,n$. Given a partition $\mathscr{C}$ of $[a,b]$ and a choice sequence η, we define the **Riemann sum of** f **over** $\mathscr{C}$ **relative to** η by

$$\mathscr{R}(f;\mathscr{C},\eta)=\sum_{j=1}^{n}f(\eta_j)(c_j-c_{j-1}).$$

Definition. We say that f is **Riemann integrable on** $[a,b]$ if there is a number $A\in\mathbb{R}$ such that for any $\varepsilon>0$ there exists a $\delta>0$ for which $|\mathscr{R}(f;\mathscr{C},\eta)-A|<\varepsilon$ whenever $\|\mathscr{C}\|<\delta$.

Clearly if such number A exists then it is unique. Q1: Explain that!

When f is Riemann integrable on $[a,b]$, we call the corresponding number A the **Riemann integral of** f **on** $[a,b]$ and write

$$A=\int_a^b f(x)\,dx=\int_{[a,b]}f(x)\,dx.$$

In other words, we say that f is Riemann integrable if $\lim_{\|\mathscr{C}\|\to 0}\mathscr{R}(f;\mathscr{C},\eta)$ exists. Here the limit is not standard, the precise definition is given above (there is A such that for any $\varepsilon>0$ there exists δ etc). We will later use limits like $\lim_{f(\alpha)\to a}F(\alpha)$ without additional explanations.

The question when a function is Riemann integrable is not trivial, but it is not difficult to see (and we will do it in Section 3.1) that each continuous function is Riemann integrable. The complete description of the Riemann integrable functions will be given later in the course.

1.2. **Riemann–Darboux sums.** Assume that $m\le f(x)\le M$ on $[a,b]$, and let $\mathscr{C}=\{c_0,c_1,\dots,c_n\}$ be a partition of $[a,b]$. We define

$$m_j=\inf\{f(x):c_{j-1}\le x\le c_j\}\quad\text{and}\quad M_j=\sup\{f(x):c_{j-1}\le x\le c_j\},$$

and introduce the **lower Riemann–Darboux sum of** f **over** $\mathscr{C}$:

$$\mathscr{L}(f;\mathscr{C})=\sum_{j=1}^{n}m_j(c_j-c_{j-1}),$$

and the **upper Riemann–Darboux sum of f over $\mathscr{C}$**:

$$\mathscr{U}(f;\mathscr{C}) = \sum_{j=1}^{n} M_j(c_j - c_{j-1}).$$

Clearly, we have

$$\mathscr{L}(f;\mathscr{C}) \le \mathscr{R}(f;\mathscr{C},\eta) \le \mathscr{U}(f;\mathscr{C})$$

for any partition $\mathscr{C}$ and any choice sequence η. Moreover, for any partition $\mathscr{C}$,

$$\mathscr{L}(f;\mathscr{C}) = \inf_{\eta} \mathscr{R}(f;\mathscr{C},\eta) \quad \text{and} \quad \mathscr{U}(f;\mathscr{C}) = \sup_{\eta} \mathscr{R}(f;\mathscr{C},\eta).$$

1.3. **Criterion for Riemann integrability.** We shall use the Riemann–Darboux sums to give a criterion for integrability that is easier to verify than the initial definition.

Lemma 1.1. *The function f is Riemann integrable if and only if*

$$\lim_{\|\mathscr{C}\|\to 0} \mathscr{L}(f;\mathscr{C}) = \lim_{\|\mathscr{C}\|\to 0} \mathscr{U}(f;\mathscr{C}).$$

By this, we mean that limits above exist and are equal.

Proof. Clearly, if f is Riemann integrable and $A = \int_a^b f(x)\,dx$, then

$$\lim_{\|\mathscr{C}\|\to 0} \mathscr{L}(f;\mathscr{C}) = \lim_{\|\mathscr{C}\|\to 0} \mathscr{U}(f;\mathscr{C}) = A.$$

Q2: Write it down!

Now suppose that $\lim_{\|\mathscr{C}\|\to 0} \mathscr{L}(f;\mathscr{C}) = \lim_{\|\mathscr{C}\|\to 0} \mathscr{U}(f;\mathscr{C}) = A$. It means that for any $\varepsilon > 0$ there exist δ_1 and δ_2 such that $|\mathscr{L}(f;\mathscr{C}) - A| < \varepsilon$ whenever $\|\mathscr{C}\| < \delta_1$ and $|\mathscr{U}(f;\mathscr{C}) - A| < \varepsilon$ whenever $\|\mathscr{C}\| < \delta_2$. We let $\delta = \min(\delta_1, \delta_2)$, then

$$|\mathscr{R}(f;\mathscr{C},\eta) - A| \le \max(|\mathscr{L}(f;\mathscr{C}) - A|, |\mathscr{U}(f;\mathscr{C}) - A|) \le \varepsilon$$

whenever $\|\mathscr{C}\| < \delta$. Thus f is Riemann integrable on $[a,b]$. □

Given two partitions $\mathscr{C}_1$ and $\mathscr{C}_2$ of $[a,b]$, we say that $\mathscr{C}_2$ is *more refined* than $\mathscr{C}_1$ (or $\mathscr{C}_2$ is a refinement of $\mathscr{C}_1$) if $\mathscr{C}_1 \subseteq \mathscr{C}_2$, i.e., each point of $\mathscr{C}_1$ is a point of $\mathscr{C}_2$.

Lemma 1.2. *Suppose that $\mathscr{C}_1 \subseteq \mathscr{C}_2$ are two partitions of $[a,b]$.*
Then $\mathscr{L}(f;\mathscr{C}_1) \le \mathscr{L}(f;\mathscr{C}_2)$ and $\mathscr{U}(f;\mathscr{C}_1) \ge \mathscr{U}(f;\mathscr{C}_2)$ for any bounded function f.

Proof. We may add points to $\mathscr{C}_1$ one by one to get $\mathscr{C}_2$. So it suffices to show that $\mathscr{L}(f;\mathscr{C}_1) \le \mathscr{L}(f;\mathscr{C}_2)$ where $\mathscr{C}_2 = \mathscr{C}_1 \cup \{c\}$. The inequality for the upper sums can be proved similarly, or one can use the identity $\mathscr{U}(f,\mathscr{C}) = -\mathscr{L}(-f,\mathscr{C})$ together with the result for lower sums. Suppose that $\mathscr{C}_1 = \{c_0, c_1, \dots, c_n\}$ and $c_{j-1} < c < c_j$. Then we have

$$\begin{aligned}
&\mathscr{L}(f;\mathscr{C}_2) - \mathscr{L}(f;\mathscr{C}_1) \\
&\qquad = \inf_{c_{j-1}\le x\le c} f(x)(c - c_{j-1}) + \inf_{c\le x\le c_j} f(x)(c_j - c) - \inf_{c_{j-1}\le x\le c_j} f(x)(c_j - c_{j-1}) \\
&\qquad = \Big(\inf_{c_{j-1}\le x\le c} f(x) - \inf_{c_{j-1}\le x\le c_j} f(x)\Big)(c - c_{j-1}) + \Big(\inf_{c\le x\le c_j} f(x) - \inf_{c_{j-1}\le x\le c_j} f(x)\Big)(c_j - c) \ge 0.
\end{aligned}$$

□

It follows from the proof that

$$\mathscr{L}(f;\mathscr{C}_2) - \mathscr{L}(f;\mathscr{C}_1) \le (M - m)\|\mathscr{C}_1\|$$

if $\mathscr{C}_2$ is obtained from $\mathscr{C}_1$ by adding a single point.

Lemma 1.3. *For any two partitions $\mathscr{C}_1$ and $\mathscr{C}_2$ of $[a,b]$*

$$\mathscr{L}(f;\mathscr{C}_1) \le \mathscr{U}(f;\mathscr{C}_2).$$

Proof. Let $\mathscr{C} = \mathscr{C}_1 \cup \mathscr{C}_2$, clearly, $\mathscr{C}$ is a partition of $[a,b]$ which is more refined than both $\mathscr{C}_1$ and $\mathscr{C}_2$. Applying the previous lemma, we obtain

$$\mathscr{L}(f;\mathscr{C}_1) \le \mathscr{L}(f;\mathscr{C}) \le \mathscr{U}(f;\mathscr{C}) \le \mathscr{U}(f;\mathscr{C}_2).$$

□

Lemma 1.4. *The limits $\lim_{\|\mathscr{C}\|\to 0}\mathscr{L}(f;\mathscr{C})$ and $\lim_{\|\mathscr{C}\|\to 0}\mathscr{U}(f;\mathscr{C})$ exist and*

$$\lim_{\|\mathscr{C}\|\to 0}\mathscr{L}(f;\mathscr{C}) = \sup_{\mathscr{C}}\mathscr{L}(f;\mathscr{C}), \qquad \lim_{\|\mathscr{C}\|\to 0}\mathscr{U}(f;\mathscr{C}) = \inf_{\mathscr{C}}\mathscr{U}(f;\mathscr{C}).$$

Proof. We shall prove that $\lim_{\|\mathscr{C}\|\to 0}\mathscr{L}(f;\mathscr{C}) = s$, where $s = \sup_{\mathscr{C}}\mathscr{L}(f;\mathscr{C})$. The second identity can be proved in the same way. It is sufficient to show that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$s - \mathscr{L}(f;\mathscr{C}) < \varepsilon$$

whenever $\|\mathscr{C}\| < \delta$. (The expression above is clearly positive, so we do not take the absolute value.)

By the definition of the supremum, there exists a partition $\mathscr{C}^* = \{c_0^*,\dots,c_n^*\}$ such that $\mathscr{L}(f;\mathscr{C}^*) > s - \varepsilon/2$. Now let $\mathscr{C}$ be any partition of $[a,b]$, we write

$$\mathscr{L}(f;\mathscr{C}) = \mathscr{L}(f,\mathscr{C}\cup\mathscr{C}^*) - \big(\mathscr{L}(f;\mathscr{C}\cup\mathscr{C}^*) - \mathscr{L}(f;\mathscr{C})\big). \tag{1}$$

Applying lemma 1.3, we have

$$\mathscr{L}(f,\mathscr{C}\cup\mathscr{C}^*) \ge \mathscr{L}(f,\mathscr{C}^*) > s - \frac{\varepsilon}{2}. \tag{2}$$

Now, we want estimate $\mathscr{L}(f;\mathscr{C}\cup\mathscr{C}^*) - \mathscr{L}(f;\mathscr{C})$, note that the partition $\mathscr{C}\cup\mathscr{C}^*$ is a refinement of $\mathscr{C}$ and it is obtained by adding less than n new points (n is the number of points in $\mathscr{C}^*$). By adding each new point to $\mathscr{C}$ we increase the lower sum by at most $(M-m)\|\mathscr{C}\|$, where $M = \sup_{[a,b]} f(x)$ and $m = \inf_{[a,b]} f(x)$. Thus, we get

$$\mathscr{L}(f;\mathscr{C}\cup\mathscr{C}^*) - \mathscr{L}(f;\mathscr{C}) \le (M-m)n\|\mathscr{C}\|.$$

Finally, we combine the last inequality with (1) and (2),

$$\mathscr{L}(f;\mathscr{C}) \ge s - \frac{\varepsilon}{2} - (M-m)n\|\mathscr{C}\| > s - \varepsilon,$$

when $\|\mathscr{C}\|$ is small enough.

We begin with ε, find a partition $\mathscr{C}^*$ as above and let $n = n(\varepsilon)$ be the number of intervals in that partition, finally we take

$$\delta = \frac{\varepsilon}{2n(M-m)}.$$

□

Theorem 1.5. *Let f be a bounded function on $[a,b]$. Then f is Riemann integrable if and only if $\sup_{\mathscr{C}}\mathscr{L}(f;\mathscr{C}) = \inf_{\mathscr{C}}\mathscr{U}(f;\mathscr{C})$.*

Proof. The theorem follows from Lemma 1.1 and Lemma 1.4. □

Corollary. *Let f be a bounded function on $[a,b]$. Then f is Riemann integrable if and only if for any $\varepsilon > 0$ there is a partition $\mathscr{C}$ of $[a,b]$ such that $\mathscr{U}(f;\mathscr{C}) - \mathscr{L}(f,\mathscr{C}) < \varepsilon$.*

Q3: Explain how the statement above follows from the theorem.

1.4. **Riemann's criterion for integrability.** Assume that $m \le f(x) \le M$ is a bounded function on $[a,b]$. Let $\mathscr{C} = \{c_0, c_1, \dots, c_n\}$ be a partition of $[a,b]$ and let d be a positive number. We introduce

$$s(f;\mathscr{C},d) = \sum_{j:\, M_j - m_j > d} (c_j - c_{j-1}),$$

the total length of the intervals of the partition in which the variability of f is greater then d.

Theorem 1.6. *The function f is Riemann integrable if and only if for any $d > 0$*

$$\lim_{\|\mathscr{C}\| \to 0} s(f;\mathscr{C},d) = 0.$$

Proof. Assume first that f is integrable. It follows from Lemma 1.1 that for any $\varepsilon > 0$ there exists δ such that $\mathscr{U}(f;\mathscr{C}) - \mathscr{L}(f;\mathscr{C}) < \varepsilon$ whenever $\|\mathscr{C}\| < \delta$. Suppose that $\|\mathscr{C}\| < \delta$, and $d > 0$ is fixed. Then we have

$$\mathscr{U}(f;\mathscr{C}) - \mathscr{L}(f;\mathscr{C}) = \sum_j (M_j - m_j)(c_j - c_{j-1}) \ge d \sum_{j:\, M_j - m_j > d} (c_j - c_{j-1}) = d s(f,\mathscr{C},d).$$

From the other hand $\mathscr{U}(f;\mathscr{C}) - \mathscr{L}(f;\mathscr{C}) < \varepsilon$ and we have

$$s(f,\mathscr{C},d) < \frac{\varepsilon}{d}$$

whenever $\|\mathscr{C}\| < \delta$. Clearly, $s(f;\mathscr{C},d)$ tends to zero when $\|\mathscr{C}\|$ does.

Now, suppose that $\lim_{\|\mathscr{C}\| \to 0} s(f;\mathscr{C},d) = 0$ for any $d > 0$. We have

$$\begin{aligned}
\mathscr{U}(f;\mathscr{C}) - \mathscr{L}(f;\mathscr{C}) &= \sum_j (M_j - m_j)(c_j - c_{j-1}) \\
&= \sum_{j:\, M_j - m_j > d} (M_j - m_j)(c_j - c_{j-1}) + \sum_{j:\, M_j - m_j \le d} (M_j - m_j)(c_j - c_{j-1}) \\
&\le (M - m) s(f,\mathscr{C},d) + d(b-a).
\end{aligned}$$

Given $\varepsilon > 0$ we take $d = \varepsilon/(2(b-a)$ and chose δ such that $s(f,\mathscr{C},d) < \varepsilon/\big(2(M-m)\big)$ if $\|\mathscr{C}\| < \delta$. It implies $\mathscr{U}(f;\mathscr{C}) - \mathscr{L}(f;\mathscr{C}) < \varepsilon$. □

1.5. **Examples.**

(1) $f(x) = 1$ on $[0,1]$, clearly $\mathscr{R}(f;\mathscr{C},\eta) = 1$ and the function is Riemann integrable, $\int_0^1 1\, dx = 1$.

(2) $f(x) = x$ on $[0,1]$. We take $\mathscr{C}_n = \{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$, then

$$\mathscr{L}(f;\mathscr{C}_n) = \sum_{j=1}^{n} \frac{j-1}{n}\frac{1}{n} = \frac{1}{n^2}\sum_{j=1}^{n} j - 1 = \frac{(n-1)n}{2n^2} = \frac{n-1}{2n},$$

$$\mathscr{U}(f;\mathscr{C}_n) = \sum_{j=1}^{n} \frac{j}{n}\frac{1}{n} = \frac{1}{n^2}\sum_{j=1}^{n} j = \frac{n(n+1)}{2n^2} = \frac{n+1}{2n},$$

and we get $\mathscr{U}(f;\mathscr{C}_n) - \mathscr{L}(f;\mathscr{C}_n) = 1/n$. Using the corollary we conclude that f is Riemann integrable. Moreover, we have

$$\sup_{\mathscr{C}} \mathscr{L}(f;\mathscr{C}) \ge \sup_n \mathscr{L}(f;\mathscr{C}_n) = \frac{1}{2} \quad \text{and}$$

$$\inf_{\mathscr{C}} \mathscr{U}(f;\mathscr{C}) \le \inf_n \mathscr{U}(f;\mathscr{C}_n) = \frac{1}{2}$$

on the other hand, $\sup_{\mathscr{C}} \mathscr{L}(f;\mathscr{C}) \le \inf_{\mathscr{C}} \mathscr{U}(f;\mathscr{C})$. Thus we get $\sup_{\mathscr{C}} \mathscr{L}(f;\mathscr{C}) = \inf_{\mathscr{C}} \mathscr{U}(f;\mathscr{C})$ and $\int_0^1 x\, dx = \frac{1}{2}$.

(3) $f(x) = 0$ if $x \in [0,1]$ is irrational, $f(x) = 1$ if $x \in [0,1]$ is rational (this function is called the *Dirichlet function*). Clearly, for any partition $\mathscr{C}$ we have $\mathscr{L}(f;\mathscr{C}) = 0$ and $\mathscr{U}(f;\mathscr{U}) = 1$. This function is not Riemann integrable.

(4) $f(x) = 0,\ x \in [0, \frac{1}{2}],\ f(x) = 1,\ x \in (\frac{1}{2}, 1]$. We consider partitions $\mathscr{C}_n = \{0, \frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n}, 1\}$. Then, readily, $\mathscr{L}(f;\mathscr{C}_n) = \frac{1}{2} - \frac{1}{n}$ and $\mathscr{U}(f;\mathscr{C}_n) = \frac{1}{2} + \frac{1}{n}$, thus we get that f is Riemann integrable and $\int_0^1 f(x)\,dx = \frac{1}{2}$.

(5) $f(x) = 0$ if $x \in [0,1]$ is irrational and $f(p/q) = 1/q$ if $p, q \in \mathbb{N}$ are mutually prime. This function is continuous at irrational points and discontinuous at rational. Q4: Explain that. We want to use Riemann's criterion to find out if it is integrable. Let d be a positive number; intervals of the partition for which $M_j - m_j > d$ should contain rational points p/q with $q < 1/d$. Let N_d be the number of rational points of this kind in $[0,1]$. Q5 Estimate it from above. We have

$$s(f;\mathscr{C},d) \le N_d \|\mathscr{C}\| \to 0 \quad \text{as} \quad \|\mathscr{C}\| \to 0.$$

Exercises.

1.1 Find $\int_0^1 x^2\,dx$ using the Riemann–Darboux sums.
(Hint: $\sum_1^n k^2 = \frac{1}{6}n(n+1)(2n+1)$)

1.2 Use the Riemann–Darboux sums to show that $f(x) = 1/x$ is Riemann integrable on $[1,2]$.

1.3 Use the Riemann integrability criterion to show that any monotone bounded function is integrable.

1.4 Suppose that f and g are Riemann integrable functions on $[a,b]$. Show that

a) $f + g$ is integrable and

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

b) cf (where c is a constant) is integrable and

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx.$$

c) fg is integrable. Hint: you may assume first that $f, g \ge 0$.

1.5 Prove that if f and g are Riemann integrable on $[a,b]$ and $f \le g$, then

$$\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx.$$

2. Topology of the real line

In this section the list of necessary facts about the topology of the real line is given. Students are assumed to know most of the material from previous courses or otherwise are recommended to work through the section very thoroughly. We end the section by proving that any function continuous on a compact subset of the real line is uniformly continuous.

2.1. **Preliminaries.** First, we remind briefly basic facts about countable sets. We always use $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ to denote the sets of positive integers, integers, rational, and real numbers.

Definition. A set X is called **countable** if there exists an injective map $s\colon X \to \mathbb{N}$, i.e. $s(x) \ne s(y)$ whenever x and y are two distinct elements of X.

Fact 2.1. *If X is a countable set, then either X is finite or there exists a one-to-one correspondence (bijection) $s\colon X \to \mathbb{N}$.*

- Any subset of $\mathbb{N}$ is countable.

- $\mathbb{Z}$ and $\mathbb{Q}$ are countable.
- Countable union of countable sets is countable.
- Product of two countable sets is countable.
- $\mathbb{R}$ is uncountable (i.e. is not countable).

Let X be an infinite set and $f : X \to [0, +\infty]$, then we define

$$\sum_{x\in X} f(x) = \sup\Big\{\sum_{x\in F} f(x) : F \subset X, F \text{ is finite}\Big\}.$$

Fact 2.2. *Let $f : X \to [0, +\infty]$, consider $Y = \{x \in X : f(x) > 0\}$.*

If Y is uncountable then $\sum_{x\in X} f(x) = +\infty$.

If Y is countably infinite (=*countable and infinite*) *then for any bijection $t : N \to Y$ one has $\sum_{x\in X} f(x) = \sum_1^\infty f\big(t(n)\big)$.*

2.2. **Completeness of the real line.**

Definition. Let E be a non-empty subset of $\mathbb{R}$. We say that a real number M is an **upper bound** for E if $x \le M$ for any $x \in E$.

We say that M_0 is the **least upper bound** for E if M_0 is an upper bound and any upper bound for E is greater than or equal to M_0.

Fact 2.3. *If a nonempty subset E of $\mathbb{R}$ has an upper bound then it has a least upper bound, which is denoted by* $\sup E$*.*

If E is non-empty and does not have an upper bound we write $\sup E = +\infty$. For completeness, it is also handy to define $\sup\emptyset = -\infty$.

Definition. A sequence of points $\langle x_k\rangle_k \subset \mathbb{R}$ is called **convergent** if $\lim_{k\to\infty} x_k = x$ for some $x \in \mathbb{R}$.

A sequence of points $\langle x_k\rangle_k \subset \mathbb{R}$ is called a **Cauchy sequence** if for any $\varepsilon > 0$ there exists n such that $|x_l - x_m| < \varepsilon$ whenever $l, m > n$.

Lemma 2.4. *Any monotone bounded sequence converges.*

Proof. Suppose that the sequence $\langle x_k\rangle_k \subset \mathbb{R}$ is non-decreasing, $x_k \le x_{k+1}$. Since the sequence is bounded, it follows from Fact 2.3 that there exists $s = \sup_k x_k$. Then $s - \varepsilon$ is not an upper bound for the sequence for any $\varepsilon > 0$. It means that there exists $k = k(\varepsilon)$ such that $s - \varepsilon < x_k \le s$. Monotonicity of the sequence implies that $s - \varepsilon < x_j \le s$ for all $j \ge k$, so $\lim_{k\to\infty} x_k = s$.

Similarly, if $\langle x_k\rangle_k \subset \mathbb{R}$ is non-increasing then $\lim_{k\to\infty} x_k = \inf_k x_k$. □

Theorem 2.5 (Bolzano–Weierstrass)**.** *Any bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

Proof. We define $b_n = \sup_{k\ge n} a_k$. Then $\langle b_n\rangle$ is a non-increasing sequence, (Q1: Explain why $b_n \ge b_{n+1}$). Then $\lim_{n\to\infty} b_n = \inf_n b_n = b$. We will show that there is a subsequence of $\langle a_n\rangle$ that converges to b.

We take k_1 such that $a_{k_1} > b_1 - 1$, then choose $k_2 > k_1$ that satisfies $a_{k_2} > b_{k_1+1} - \frac{1}{2}$. If k_n is chosen, we take $k_{n+1} > k_n$ such that $a_{k_{n+1}} > b_{k_n+1} - 1/n$. Clearly, $a_{k_{n+1}} \le b_{k_n+1}$. We have $\lim_{n\to\infty} a_{k_n} = \lim_{n\to\infty} b_{k_n+1} = b$. □

The number $b = \inf_n \sup_{k\ge n} a_k$ is called the upper limit of the sequence $\langle a_k\rangle$ and is denoted by $\limsup_k a_k$. The lower limit is similarly defined by

$$\liminf_k a_k = \sup_n \inf_{k\ge n} a_k.$$

The definitions above have sense for unbounded sequences as well the limits are allowed to take values in $\mathbb{R}^* = \mathbb{R} \cup \{+\infty\} cup \{-\infty\}$.

Theorem 2.6. *Each Cauchy sequence in $\mathbb{R}$ converges.*

That fact above is usually formulated as: "$\mathbb{R}$ is a complete metric space".

Proof. First, any Cauchy sequence is bounded. (Q2: Explain.) Thus we can use the Bolzano–Weierstrass theorem and find a subsequence $\langle a_{k_n} \rangle$ that converges.

Suppose $a = \lim_{n\to\infty} a_{k_n}$; let us prove that $a = \lim_{k\to\infty} a_k$. For each $\varepsilon > 0$ we choose N such that $|a - a_{k_n}| < \varepsilon/2$ if $n \geq N$. Since $\langle a_k \rangle$ is a Cauchy sequence, we can also choose K such that $|a_l - a_k| < \varepsilon/2$ if $k, l \geq K$. Now if $k \geq K$ we choose $n \geq N$ and such that $k_n \geq K$, then we have

$$|a_k - a| \leq |a_k - a_{k_n}| + |a_{k_n} - a| < \varepsilon.$$

□

2.3. **Open, closed and compact subsets of the real line.**

Definition. A set $U \subset \mathbb{R}$ is called **open** if for any $x \in U$ there exists $a = a(x)$ such that $(x - a, x + a) \subset U$. We denote by $\mathfrak{G}$ the class of open subsets of the real line.

A set $F \subset \mathbb{R}$ is called **closed** if its complement $F^c = \mathbb{R} \setminus F$ is open. The class of all closed subsets of $\mathbb{R}$ is denoted by $\mathfrak{F}$.

- Any (open) interval $(a, b) = \{x : a < x < b\}$ is open, (here $a, b \in [-\infty, +\infty]$).
- The empty set is open.
- The union of any family of open sets is open.
- The intersection of any finite family of open sets is open.
- Any (closed) interval $[a, b] = \{x : a \leq x \leq b\}$ is closed.
- The empty set is closed.
- Any (closed) half-ray $[a, +\infty)$ of $(-\infty, b]$ is closed.
- The intersection of any family of closed sets is closed.
- The union of any finite family of closed sets is closed.

Open subsets of the real line can be described in a very transparent way. We will use the fact below through the whole course.

Fact 2.7. *Every open subset of $\mathbb{R}$ is a countable union of **disjoint** open intervals.*

Definition. A set $K \subset \mathbb{R}$ is called **compact** if for every cover of K by open sets, $K \subset \cup_{\alpha\in A} U_\alpha$, there is a finite set $F \subset A$ such that $K \subset \cup_{\alpha\in F} U_\alpha$. (Each open cover contains a finite subcover.)

Theorem 2.8. *Let K be a subset of $\mathbb{R}$, then the following are equivalent:*

(1) *K is compact.*
(2) *K is closed and bounded.*
(3) *Every sequence in K has a subsequence that converges to a point of K.*

Proof. (1) *implies* (2): Clearly a compact set is bounded: consider the cover $V_n = (-n, n)$, by the definition of a compact set, there is a finite subcover $K \subset \cup_j (-n_j, n_j)$ and then $K \subset (-\max_j n_j, \max_j n_j)$.

To show that K is closed, take a point $x \in \mathbb{R} \setminus K$. We have

$$K \subset \bigcup_{\varepsilon>0} \mathbb{R} \setminus [x - \varepsilon, x + \varepsilon],$$

this open cover has a finite subcover, thus $K \subset \mathbb{R} \setminus (x - \varepsilon_0, x + \varepsilon_0)$ for some $\varepsilon_0 > 0$ and $(x - \varepsilon_0, x + \varepsilon_0) \subset \mathbb{R} \setminus K$. So $\mathbb{R} \setminus K$ is open.

(2) *implies* (3): By the Bolzano – Weierstrass theorem this bounded sequence has a convergent subsequence $\langle t_k \rangle$. We want to show that the limit point t belongs to K. Otherwise $t \in \mathbb{R} \setminus K$ and $(t-a, t+a) \in \mathbb{R} \setminus K$ for some $a > 0$ since K is closed. It means that no points of the sequence $\langle t_k \rangle$ is contained in $(t-a, t+a)$, so t is not the limit point.

(3) *implies* (1): It follows easily from (3) that K is bounded. Q3: Explain!

We want to show that K is compact. Let $\cup_\alpha U_\alpha$ be an open cover of K.

We will first show that for some n each segment of length 2^{-n} that intersects K is contained in some of the sets U_α. If it is not true, there is a sequence of intervals $\langle I_n \rangle_{n=1}^\infty$ such that $I_n \cap K \neq \emptyset$, $|I_n| < 2^{-n}$ and I_n is not contained in any U_α. Let $x_n \in I_n \cap K$, by (3) there is a convergent subsequence $\langle x_{n_k} \rangle$ of this sequence and $x = \lim x_{n_k}$ is a point of K. Then $x \in U_\alpha$ for some α. Now since U_α is open, there exists m such that $(x - 2^{-m}, x + 2^{-m}) \in U_\alpha$. Finally, if $|x_{n_k} - x| < 2^{-m-1}$ and $n_k > m+1$, then $I_{n_k} \subset U_\alpha$. We got a contradiction. Thus it is true that there exists n such that each interval of length 2^{-n} that intersects K is contained in one of the elements of the cover.

Since K is bounded it can be covered by a finite number of segments of length 2^{-n}. We just divide a big interval that contains K into segments of length 2^{-n} and take those segments that intersect K. For each of them there is a set in the cover that contains this segment, those sets form a finite subcover.

□

The statement above implies that every closed bounded interval $[a, b]$ is compact.

Most of the facts above (excluding the characterization of open sets) remain true in $\mathbb{R}^n$.

2.4. **Continuous and uniformly continuous functions.** Let X be a subset of $\mathbb{R}$. We consider real-valued functions defined on X.

Definition. We say that f is **continuous at a point** $x \in X$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $y \in X$ and $|x - y| < \delta$; f is called **continuous on** X if it is continuous at each point of X.

We say that f is **uniformly continuous** on X if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in X$ and $|x - y| < \delta$.

Q4: Show that any function on $\mathbb{N}$ is uniformly continuous.

- A uniformly continuous function is continuous at each point.
- Function $f(x) = 1/x$ is continuous on $(0, +\infty)$ but is NOT uniformly continuous on this set. Q5:Construct a bounded function with this property.

Theorem 2.9. *Let f be a continuous function on a compact subset K of $\mathbb{R}$. Then f is uniformly continuous on K.*

Proof. Fix $\varepsilon > 0$, for each $x \in K$ there exist $I(x) = (x - \delta(x), x + \delta(x))$ such that $|f(x) - f(y)| < \varepsilon/2$ for any $y \in I(x) \cap K$, we use the condition that function is continuous at each point. Consider the intervals $J(x) = (x - \frac{1}{2}\delta(x), x + \frac{1}{2}\delta(x))$, clearly $\cup_{x \in K} J(x)$ is an open cover of K. By the definition of a compact set, there is a finite subset $\langle x_1, \ldots, x_m \rangle$ such that $K \subset \cup_1^m J(x_j)$. Now let

$$\delta = \frac{1}{2} \min_{1 \le j \le m} \delta(x_j),$$

and suppose that $y_1, y_2 \in K$ are such that $|y_1 - y_2| < \delta$. Then $y_1 \in J(x_k)$ for some $k, 1 \le k \le m$. It means that $|y_1 - x_k| < \frac{1}{2}\delta(x_k)$ and we have

$$|y_2 - x_k| \le |y_2 - y_1| + |y_1 - x_k| < \delta + \frac{\delta(x_k)}{2} \le \frac{\delta(x_k)}{2} + \frac{\delta(x_k)}{2} = \delta(x_k).$$

Then $y_2 \in I(x_k)$, clearly $y_1 \in I(x_k)$ and

$$|f(y_1) - f(y_2)| \le |f(y_1) - f(x_k)| + |f(x_k) - f(y_2)| < \varepsilon.$$

□

In particular, any continuous function on a bounded closed interval is uniformly continuous.

Exercises.

2.1 Let X be a subset of $\mathbb{R}$. We define $\operatorname{dist}(y, X) = \inf\{|x - y| : x \in X\}$. Prove that for any $\varepsilon > 0$ the set $\{y \in \mathbb{R} : \operatorname{dist}(y, X) < \varepsilon\}$ is open.
2.2 Show that each closed set is the intersection of a countable family of open sets.
2.3 Let f be a continuous function on $\mathbb{R}$. Proof that the set $\{x \in \mathbb{R} : f(x) > c\}$ is open for any $c \in \mathbb{R}$.
2.4 Let $f : \mathbb{R} \to \mathbb{R}$ be a non-decreasing function. Show that f is continuous at each but countably many points.
2.5 Let f be a function on X, for each positive δ we define

$$\omega_{f,X}(\delta) = \sup\{|f(x) - f(y)| : x, y \in X, |x - y| < \delta\}.$$

Show that f is uniformly continuous on X if and only if $\lim_{\delta\to 0} \omega_{f,X}(\delta) = 0$. Function $\omega_{f,X}(\delta)$ is called the modulus of continuity of f on X.

3. Riemann integration of continuous functions

3.1. **Main theorem.** Finally we are ready to prove that the class of Riemann integrable functions contains all continuous functions. As we have seen in Section 1.3, some discontinuous functions are also Riemann integrable; all piecewise continuous functions are Riemann integrable.

Theorem 3.1. *Let f be a continuous function on the bounded interval $[a, b]$, then f is Riemann integrable.*

Proof. We will use that f is **uniformly** continuous on $[a, b]$ and apply the criterion of integrability we got at the end of Section 1.3.

Uniform continuity means that for any positive ε there is a positive $\delta(\varepsilon)$ such that $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in [a, b]$ and $|x - y| < \delta(\varepsilon)$. Now if we consider a partition $\mathscr{C} = \{c_0, \dots, c_n\}$ with $\|\mathscr{C}\| < \delta(\varepsilon)$, we get

$$M_j - m_j = \sup_{c_{j-1}\le x\le c_j} f(x) - \inf_{c_{j-1}\le x\le c_j} f(x) \le \varepsilon.$$

Thus we can estimate the difference between the upper and the lower sums

$$\mathscr{U}(f;\mathscr{C}) - \mathscr{L}(f;\mathscr{C}) = \sum_j (M_j - m_j)(c_j - c_{j-1}) \le \varepsilon(b - a).$$

In particular, we can easily find partitions $\mathscr{C}_n$ such that $\mathscr{U}(f;\mathscr{C}_n) - \mathscr{L}(f;\mathscr{C}_n)$ tends to 0 when n goes to infinity. Hence f is Riemann integrable. □

3.2. **Riemann integration on $\mathbb{R}^N$.** The definition we developed in Section 1 can be extended to functions of several variables, the same argument as above shows that continuous functions are Riemann integrable.

The role of intervals is now played by rectangles whose sides are parallel to the coordinate axes, for a such rectangle $I = \prod_{k=1}^{N} [a_k, b_k]$ we define

$$\operatorname{diam} I = \Big(\sum_k (b_k - a_k)^2\Big)^{1/2} \quad \text{and} \quad \operatorname{vol} I = \prod_{k=1}^{N} (b_k - a_k).$$

By I° we denote the interior of the rectangle I, $I^\circ = \prod_{k=1}^N (a_k, b_k)$.

Let $\mathscr{C}$ be a collection of rectangles, we say that $\mathscr{C}$ is non-overlapping if elements of $\mathscr{C}$ have disjoint interiors.

Fact 3.2. *If $\mathscr{C}$ is a non-overlapping finite collection of rectangles each of which is contained in the rectangle J, then* $\mathrm{vol}(J) \geq \sum_{I\in\mathscr{C}} \mathrm{vol}(I)$.

If $\mathscr{C}$ is a finite collection of rectangles and the rectangle J is covered by $\mathscr{C}$, i.e. $J \subset \cup_{I\in\mathscr{C}} I$, then $\mathrm{vol}(J) \leq \sum_{I\in\mathscr{C}} \mathrm{vol}(I)$.

A finite collection of rectangles $\mathscr{C}$ is called a partition of the rectangle J if $\mathscr{C}$ is non-overlapping and $J = \cup_{I\in\mathscr{C}} I$. As before $\|\mathscr{C}\| = \max\{\mathrm{diam}(I) : I \in \mathscr{C}\}$ is called the mesh size of $\mathscr{C}$. Given a partition $\mathscr{C}$ of J, we say that $\eta : \mathscr{C} \to J$ is a choice map for $\mathscr{C}$ if $\eta(I) \in I$ for each $I \in \mathscr{C}$.

Let f be a bounded function defined on J, $\mathscr{C}$ be a partition of J and η be a choice map for $\mathscr{C}$, then we define the Riemann sum of f over $\mathscr{C}$ relative to η to be

$$\mathscr{R}(f;\mathscr{C},\eta) = \sum_{I\in\mathscr{C}} f\big(\eta(I)\big)\,\mathrm{vol}(I).$$

The lower and upper Riemann–Darboux sums are defined by

$$\mathscr{L}(f;\mathscr{C}) = \sum_{I\in\mathscr{C}} \inf_{x\in I} f(x)\,\mathrm{vol}(I), \quad \mathscr{U}(f;\mathscr{C}) = \sum_{I\in\mathscr{C}} \sup_{x\in I} f(x)\,\mathrm{vol}(I).$$

A bounded function f on J is called Riemann integrable if $\lim_{\|\mathscr{C}\|\to 0} \mathscr{R}(f;\mathscr{C},\eta)$ exists. The function f is Riemann integrable if and only if $\sup_{\mathscr{C}} \mathscr{L}(f;\mathscr{C}) = \inf_{\mathscr{C}} \mathscr{U}(f;\mathscr{C})$.

3.3. **Basic properties of the Riemann integral.** We briefly recall the basic facts about the integration of continuous functions. We denote by $C(X)$ the class of all functions continuous on X. Clearly $C(X)$ is a linear space with standard operations: the addition of function and the multiplication by a scalar. There is a natural norm on the this space that we call the uniform norm:

$$\|f\|_X = \sup_{x\in X} |f(x)|.$$

If X is compact subset of $\mathbb{R}$ one can write max above, since a function continuous on a compact set achieves its supremum on this set (see exercises).

Definition. We say that a sequence of functions f_n converges **uniformly** on X to the function f if $\lim_{n\to\infty} \|f_n - f\|_X \to 0$.

Another convergence of sequence of functions is so called pointwise convergence. We say that a sequence of functions f_n converges to a function f **pointwise** on X if $\lim_{n\to\infty} f_n(x) = f(x)$ for any $x \in X$.

Clearly, if f_n converges uniformly on X to f, then it converges to f pointwise. Q1: Explain.

The converse is not true, consider $f_n(x) = x^n$ on $X = [0,1]$, then f_n converges pointwise to function

$$f(x) = \begin{cases} 0 & \text{if } 0 \leq x < 1, \\ 1 & \text{if } x = 1. \end{cases}$$

But f_n does not converge uniformly to f. Q2: Calculate $\|f_n - f\|_{[0,1]}$.

Lemma 3.3. *A uniform limit of a sequence of continuous functions is continuous.*

Proof. We want to show that f is continuous at each point $x \in X$. Let ε be a positive number, there exists n such that $\|f - f_n\|_X < \varepsilon/3$. Further, f_n is continuous at x, thus there exists $\delta > 0$ such that $|f_n(y) - f_n(x)| < \varepsilon/3$ when $x, y \in X$ and $|x - y| < \delta$. Thus we have

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < \varepsilon$$

when $x, y \in X$ and $|x-y| < \delta$. □

The main properties of the Riemann integral are listed below:

(1) The Newton-Leibnitz formula: Let f and F be continuous functions on $[a,b]$, if $f(x) = F'(x)$ on $[a,b]$ then
$$\int_a^b f(x)\,dx = F(b) - F(a)$$
(2) If f is a continuous function on $[a,b]$ and $F(x) = \int_a^x f(t)\,dt$, then $F'(x) = f(x)$. (see exercises below)
(3) Integration by parts formula: Let $f, g \in C^1([a,b])$. Here and below we denote by $C^k[a,b]$ the space of all functions on $[a,b]$ that are continuous on this interval with all their derivatives up to the order k. Then
$$\int_a^b f(x)g'(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)\,dx.$$
(4) Let f_n be a sequence of Riemann integrable functions that converges uniformly on $[a,b]$ to the function f, then f is Riemann integrable and
$$\lim_n \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$
(see exercises below)
(5) Let $f(x,y)$ be a continuous function on $[a,b]\times[c,d]$, then
$$\int_a^b\int_c^d f(x,y)\,dx\,dy = \int_a^b\left(\int_c^d f(x,y)\,dy\right)dx.$$

3.4. **The idea of the Lebesgue integral.** We describe very roughly the idea behind the definition of the Lebesgue integral. Careful definition will be given later.

Given a positive function f on the interval $[a,b]$, we want to calculate the area of the subgraph $\{(x,y): a \le x \le b, 0 \le y \le f(x)\}$. The Riemann integration procedure consists of dividing the interval $[a,b]$ into small intervals and approximating the "area" by the sum of the areas of rectangles. One of the ways to look at the Lebesgue integration is to divide the other axis and consider parts of the subgraph that correspond to the points where the function f has close values:
$$\left\{(x,y): 0 \le y \le f(x), \frac{k-1}{n} < f(x) \le \frac{k}{n}\right\}.$$
We want to say that the area of the set above is approximately k/n times the length of the set $\{x \in [a,b]: (k-1)/n < f(x) \le k/n\}$. If we denote by $|E|$ the length of E (whatever it means), we would like to say that
$$\int_a^b f(x)\,dx = \lim_{n\to+\infty}\sum_k \frac{k}{n}\left|\left\{x \in [a,b]: \frac{k-1}{n} < f(x) \le \frac{k}{n}\right\}\right|.$$
Our first task is to understand how the notion of the length can be extended from the class of intervals to all (actually all not very bad) subsets of the line. And then we would look closer at the limit above. We will describe a class of integrable functions, show that the Lebesgue integral is really an extension of the Riemann integral and discuss the analogs of the basic properties of the Riemann integral of continuous functions that are listed in Section 3.3.

Exercises.

3.1 Let f be a continuous function on $[a,b]$ and $F(x)=\int_a^x f(t)\,dt$. Show that
 a) $F(x)\in C[a,b]$,
 b) $F(x)$ is differentiable on (a,b) and $F'(x)=f(x)$.

3.2 Consider the sequence of functions

$$f_n = \begin{cases} 0 & \text{if } |x-2/n|>1/n, \\ n-n^2|x-2/n| & \text{if } |x-2/n|\le 1/n. \end{cases}$$

 a) Draw the graph of f_n, calculate $\|f_n\|_{[0,1]}$.
 b) Show that f_n converges pointwise to 0 on $[0,1]$.
 c) Calculate $\int_0^1 f_n(x)\,dx$.

3.3 Show that the uniform limit of a sequence of uniformly continuous functions is uniformly continuous. (In other words, if each function f_n is uniformly continuous on X and $f_n\to f$ uniformly on X, then f is uniformly continuous.)

3.4 Let f_n be a sequence of Riemann integrable functions that converges uniformly on $[a,b]$ to the function f. Show that f is Riemann integrable and

$$\lim_n \int_a^b f_n(x)dx = \int_a^b f(x)\,dx.$$

3.5 Let f be a continuous function on $[a,b]$, $\mathscr{C}$ be a partition of $[a,b]$. Show that there exists a choice sequence η for $\mathscr{C}$ such that

$$\int_a^b f(x)\,dx = \mathscr{R}(f;\mathscr{C},\eta).$$

3.5 Let $f(x,y)$ be a continuous function on $[a,b]\times[c,d]$. For any $x\in[a,b]$ we define

$$F(x)=\int_c^d f(x,y)\,dy.$$

 a) Show that F is continuous function on $[a,b]$.
 b) Prove the identity

$$\int_a^b\int_c^d f(x,y)\,dx\,dy = \int_a^b F(x)\,dx.$$

 Hint: you may use the result of the previous exercise and consider the partition of the rectangle $[a,b]\times[c,d]$ into n^2 congruent rectangles.

4. RIEMANN–STIELTJES INTEGRATION AND FUNCTIONS OF BOUNDED VARIATION

4.1. **Definition and examples of the Riemann–Stieltjes integral.** Before we start the discussion of the Lebesgue integral, we will define one extension of the Riemann integral, so called Stieltjes or Riemann–Stieltjes integral.

Let $[a,b]$ be an interval and let f,g be two bounded functions defined on $[a,b]$. Given a partition $\mathscr{C}=\{c_0,\dots,c_n\}$ of $[a,b]$ and a choice sequence $\eta=\{\eta_1,\dots\eta_n\}$, we define the **Riemann sum of** f **over** $\mathscr{C}$ **with respect to** g **relative to** η to be

$$\mathscr{R}(f|g;\mathscr{C},\eta)=\sum_j f(\eta_j)\big(g(c_j)-g(c_{j-1})\big).$$

We say that f **is Riemann integrable on** $[a,b]$ **with respect to** g if there exists

$$\lim_{\|\mathscr{C}\|\to 0}\mathscr{R}(f|g;\mathscr{C},\eta).$$

The limit is called the **Riemann-Stiltjes integral of f with respect to** g and is denoted by

$$\int_a^b f(x)\,dg(x) = \int_a^b f\,dg.$$

Examples.

(1) If $g(x) = x$ for $x \in [a,b]$, then the Riemann–Stieltjes integral is just the classical Riemann integral.
(2) If $g(x) = c$ is a constant function, then any function f is Riemann integrable with respect to g and the Riemann–Stieltjes integral is equal to 0.
(3) if $f \in C([a,b])$ and $g \in C^1([a,b])$ then f is Riemann integrable with respect to g and

$$\int_a^b f(x)\,dg(x) = \int_a^b f(x)g'(x)\,dx.$$

(see exercises below)
(4) If f_1 and f_2 are Riemann integrable with respect to g then so is $a_1 f_1 + a_2 f_2$, moreover

$$\int_a^b (a_1 f_1 + a_2 f_2)(x)\,dg(x) = a_1 \int_a^b f_1(x)\,dg(x) + a_2 \int_a^b f_2(x)\,dg(x).$$

To see it, note that

$$\mathscr{R}((a_1 f_1 + a_2 f_2)|g;\mathscr{C},\eta) = a_1\mathscr{R}(f_1|g;\mathscr{C},\eta) + a_2\mathscr{R}(f_2|g;\mathscr{C},\eta).$$

4.2. **Integration by parts formula.** We will prove a generalization of the classical integration by parts formula.

Theorem 4.1. *If f is Riemann integrable on $[a,b]$ with respect to g, then g is Riemann integrable on $[a,b]$ with respect to f and*

$$\int_a^b g(x)\,df(x) = f(b)g(b) - f(a)g(a) - \int_a^b f(x)\,dg(x). \tag{3}$$

Proof. Let $\mathscr{C} = \{c_0,\dots,c_n\}$ be a partition of $[a,b]$ and let $\eta = \{\eta_1,\dots,\eta_n\}$ be a choice sequence for $\mathscr{C}$. We add the endpoints of the interval to η to get the new partition $\mathscr{C}' = \{a = \eta_0, \eta_1, \dots, \eta_n, \eta_{n+1} = b\}$, then $\eta' = \{c_0,\dots,c_n\}$ is a choice sequence for $\mathscr{C}'$. Clearly, $\|\mathscr{C}'\| \le 2\|\mathscr{C}\|$ (Q1:Explain) and we have

$$\begin{aligned}
\mathscr{R}(g|f;\mathscr{C},\eta) &= \sum_{j=1}^n g(\eta_j)(f(c_j) - f(c_{j-1})) = \sum_{j=1}^n g(\eta_j)f(c_j) - \sum_{j=1}^n g(\eta_j)f(c_{j-1}) \\
&= \sum_{j=1}^n g(\eta_j)f(c_j) - \sum_{j=0}^{n-1} g(\eta_{j+1})f(c_j) = f(c_n)g(\eta_{n+1}) - f(c_0)g(\eta_0) - \sum_{j=0}^n f(c_j)\big(g(\eta_{j+1}) - g(\eta_j)\big) \\
&= f(b)g(b) - f(a)g(a) - \mathscr{R}(f|g;\mathscr{C}',\eta').
\end{aligned}$$

Therefore, if the limit $\lim_{\|\mathscr{C}\|\to 0}\mathscr{R}(f|g;\mathscr{C},\eta)$ exists then the limit $\lim_{\|\mathscr{C}\|\to 0}\mathscr{R}(g|f;\mathscr{C},\eta)$ exists as well and the integration by parts formula (3) holds. □

Corollary (Fundamental theorem of calculus)**.** *If $g \in C^1([a,b])$ then*

$$\int_a^b g'(x)\,dx = g(b) - g(a).$$

Proof. As we have seen in example (3), the function $f(x) = 1$ is Riemann integrable with respect to g and $\int_a^b g'(x)\,dx = \int_a^b 1\,dg(x)$. Now we apply the integration by parts formula and the remark of example (2):

$$\int_a^b 1\,dg(x) = g(b) - g(a).$$

□

Corollary. *If $f, g \in C^1([a,b])$ then*

$$\int_a^b f(x)g'(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)\,dx.$$

Q2: Explain how it follows from the theorem.

4.3. **Functions of bounded variation.**

Definition. We say that a function $F : \mathbb{R} \to \mathbb{R}$ has bounded variation and write $F \in BV(\mathbb{R})$ if

$$\mathrm{TV}_F = \sup\Big\{\sum_1^n |F(x_j) - F(x_{j-1})| : -\infty < x_0 < \cdots < x_n < +\infty\Big\} < +\infty.$$

In the same way we may define functions of bounded variation on $[a,b]$, we take then $x_0 = a, x_n = b$. Suppose that F is a function of bounded variation. We define the variation function of F by

$$V_F(x) = \sup\Big\{\sum_1^n |F(x_j) - F(x_{j-1})| : -\infty < x_0 < \cdots < x_n = x\Big\}$$

Clearly, V_F is a non-decreasing function and $\mathrm{TV}_F = \lim_{x\to+\infty} V_F(x)$.

Examples.

(1) A constant function has bounded variation.
(2) Each monotone bounded function has bounded variation.
(3) If $F, G \in BV(\mathbb{R})$ then $aF + bG \in BV(\mathbb{R})$ for any $a, b \in \mathbb{R}$.
(4) $F(x) = \sin x$ has unbounded variation on $(-\infty, +\infty)$ but bounded variation on any finite interval.
(5) $F(x) = \sin(1/x)$ has unbounded variation on $(0,1)$.
(6) $F(x) = x\sin(1/x) \in C([0,1])$ and has unbounded variation on $[0,1]$.
(7) If $F \in BV(\mathbb{R})$ then $|F| \in BV(\mathbb{R})$.
(8) If F satisfies the Lipschitz condition $|F(x) - F(y)| \le k|x - y|$, then F has bounded variation.

It follows from (2) that functions of bounded variation form a vector space, variation $V(F) = Var_f$ is a seminorm on this space,

$$V(aF) \le |a|V(F) \quad \text{and} \quad V(F+G) \le V(F) + V(G).$$

Lemma 4.2. *Let $F \in BV(\mathbb{R})$, then $V_F + F$ and $V_F - F$ are non-decreasing.*

Proof. It is enough to prove that for example $V_F - F$ is non-decreasing, to get the second statement we consider function $-F$.

Suppose that $x < y$ we want to show that $V_F(x) - F(x) \le V_F(y) - F(y)$. For any $\varepsilon > 0$ we can find $-\infty < x_0 < \cdots < x_n = x$ such that $V_F(x) < \sum |F(x_{j-1}) - F(x_j)| + \varepsilon$. Then we have

$$V_F(y) - F(y) \ge \sum |F(x_{j-1}) - F(x_j)| + |F(y) - F(x)| - F(y) >$$
$$V_F(x) - \varepsilon + F(y) - F(x) - F(y) = V_F(x) - F(x) - \varepsilon.$$

Since it is true for any $\varepsilon > 0$ we get the required inequality. □

Now we may give a different characterization of functions of bounded variation.

Theorem 4.3 (Jordan's decomposition)**.** *The function $F : \mathbb{R} \to \mathbb{R}$ has bounded variation if and only if F is the difference of two bounded non-decreasing functions.*

Proof. Suppose that $F \in BV(\mathbb{R})$; then F is bounded. Q3:Check it! We can write $2F(x) = V_F(x) + F(x) - \big(V_F(x) - F(x)\big)$. Both functions $V_F + F$ and $V_F - F$ are non-decreasing; V_F is bounded by the definition of $BV(\mathbb{R})$. Thus $V_F \pm F$ is also bounded since F is bounded.

Opposite implication is obvious. Each bounded non-decreasing function has bounded variation and then the difference of two functions with bounded variations has bounded variation again. □

4.4. **Riemann–Stieltjes integral of functions of bounded variation.** It turns out that all continuous functions are Riemann integrable with respect to functions of bounded variation.

Theorem 4.4. *If $f \in BV([a,b])$ and $g \in C([a,b])$ then g is Riemann integrable over $[a,b]$ with respect to f.*

Proof. We know that $f = f_1 - f_2$, where f_1 and f_2 are non-decreasing functions. It suffices to show that g is integrable with respect to f_j, $j = 1,2$. (Then f_j is integrable with respect to g and $f = f_1 - f_2$ is integrable with respect to g.) Q4: Stop here and make certain that you have understood the argument.

So we may assume that f is non-decreasing. Then one can define the upper and lower sums as in Section 1.2 and

$$\mathscr{U}(g|f;\mathscr{C}) - \mathscr{L}(g|f;\mathscr{C}) = \sum_j \Big(\sup_{[c_{j-1},c_j]} g - \inf_{[c_{j-1},c_j]} g \Big)\big(f(c_j) - f(c_{j-1})\big) \le \omega_g(\delta)(f(b) - f(a)),$$

where $\|\mathscr{C}\| < \delta$ and $\omega_g(\delta) = \sup\{|g(x) - g(y)|, x, y \in [a,b], |x-y| < \delta\}$. Since g is uniformly continuous on $[a,b]$ we have $\omega_g(\delta) \to 0$ as δ goes to 0. Then $\lim_{\|\mathscr{C}\|\to 0} \mathscr{U}(g|f;\mathscr{C}) - \mathscr{L}(g|f;\mathscr{C}) = 0$ and in the same way as it was done for the Riemann integral we may show that $\mathscr{R}(g|f;\mathscr{C},\eta)$ has a limit. □

It is not difficult to see that the following estimate holds:

$$\left| \int_a^b g(x)\, df(x) \right| \le \mathrm{TV}_{f,[a,b]} \sup_{x\in[a,b]} |g(x)| \tag{4}$$

(see exercises below).

It means that if we consider the set of continuous functions on $[a,b]$ as a linear space with the uniform norm, i.e., space $C[a,b]$, then each function of bounded variation defines a bounded linear functional on this space. The norm of this functional does not exceed the total variation of the corresponding function.

Exercises.

4.1 Show that $f(x) = x^2 \sin 1/x$ has bounded variation on $[0,1]$.

4.2 Let $f \in C([a,b])$ and $g \in C^1([a,b])$, apply the Mean Value theorem to show that f is Riemann integrable with respect to g and

$$\int_a^b f(x)\, dg(x) = \int_a^b f(x) g'(x)\, dx.$$

4.3 Proof inequality (4).

4.4 Give an example of a function $\psi \in BV([0,1])$ for which

$$\sup\Big\{ \int_0^1 \phi(x)\, d\psi(x) \colon \phi \in C([a,b]), \|\phi\| = 1 \Big\} < \mathrm{TV}(\psi;[0,1]).$$

4.5 Suppose that there exist $a = a_0 < a_1 < \cdots < a_n = b$ such that ψ is constant on each of the intervals (a_{m-1}, a_m), $\psi(x) = \psi_m$ for $a_{m-1} < x < a_m$. Prove that every $\phi \in C([a,b])$ is ψ-Riemann integrable on $[a,b]$ and

$$\int_a^b \phi(x)\, d\psi(x) = \sum_{m=0}^{n} \phi(a_m)(\psi_{m+1} - \psi_m),$$

where $\psi_0 = \psi(a)$ and $\psi_{n+1} = \psi(b)$.

4.6 Suppose that ϕ is ψ-Riemann integrable on $[a,b]$ and $a < c < b$. Then ϕ is ψ-Riemann integrable on $[a,c]$ and $[c,b]$ and

$$\int_a^b \phi(x)\,d\psi(x) = \int_a^c \phi(x)\,d\psi(x) + \int_c^b \phi(x)\,d\psi(x)$$

5. Historical notes and references

5.1. **Integral calculus before Riemann.** Some elements of the integral calculus can be traced back to ancient mathematics. We can mention amazing techniques developed by Greeks for computation of areas and volumes. However the discovery of the modern calculus is traditionally attributed to Isaac Newton and Gottfried Wilhelm Leibniz. Those Great Minds were not only aware of the fundamental theorem of calculus, but also realized the power of the tool they got in their hands.

Newton's applications of the antidifferentiation for calculating areas are dated back to as early as 1666. He computed integrals of the rational powers of variable and, when dealing with more complicated functions, practiced term by term integration of series. Newton also developed integration by substitution and made perhaps the first tables of integrals [1]. Leibniz published a number of articles devoted to differential and integral calculus in 1680s and 90s. It is him we should thank for the invention of the integral sign. The fundamental theorem of calculus got a proof in his paper of 1693 [2]. One of Leibniz's results, the transmutation theorem [4], gives a remarkable example of how the integral calculus was developing. Interesting comparison of ideas and methods used by Newton and Leibniz as well as further references can be found in [1].

Works of Newton and Leibniz were of an ultimate influence. Integral calculus merged into mathematical textbooks and became a powerful tool for the next generations of analysts. The most spectacular applications of integral appeared somewhat a half-century later in the works of Leonhard Euler. His tricks with integrals and infinite series showed his remarkable mathematical intuition and at the same time attracted attention to the foundations of analysis. Notions of function, limit, convergence, continuity refereed at that time to geometric intuition and were far from being precise.

The beginning of the nineteenth century starts a new era in foundation of analysis. It was Augustin-Louis Cauchy who undertake the revision of calculus. His works in 1820s, where the limit, convergence, derivative and integral got a new life, were very influential. Cauchy was much more careful with the definitions than his forerunners, with his prime interest lying in the development of the theory not in its computational applications. Cauchy introduced integral sums

$$S = (x_1 - a)f(a) + (x_2 - x_1)f(x_1) + \cdots + (b - x_{n-1})f(x_{n-1})$$

associated with a partition of the interval $[a,b]$, and defined the integral of a continuous function f "taken between the limits a and b" as the limit of this sums [3]. The theorem on integrability of continuous functions and the fundamental theorem of calculus were proved [4]. Lectures furnishing Cauchy's approach to calculus were published by a student of his, Abbe Moigno, in 1844 [3].

The role played by Cauchy in the modern development of the calculus can be hardly overestimated. Though the definitions did not get there final clarity in his works and he didn't recognize uniform continuity and uniform convergence as the main players in the integral calculus, Cauchy's works on foundations have motivated and influenced his successors.

5.2. **"What is to be understood by $\int_a^b f(x)dx$?".** The study of vibrating string led to the development of mathematical theory of trigonometric series in the middle of the 18th century. It was revisted by Fourier a half-century later. Questions arising from his work appealed to discontinuous functions and there integrals. First result on convergence of Fourier series (for piece-wise continuous functions) was proved by Dirichlet in 1829, in the same paper a famous example of nowhere continuous function was given [1].

The question taken as the title of this subsection appeared in the investigation done by Riemann in 1854, where he followed Dirichlet in the study of the convergence of Fourier series [3]. His work starts with detailed historical introduction to the theory of Fourier series, and then definitions of what we now know as Riemann sums and Riemann integrability comes. Riemann gave also a criterion for a bounded function to be integrable (but his proof implicitly refereed to the completeness of the real line). This criterion was used to construct an integrable function that is discontinuous at each rational point of the form $p/2q$.

In 1870s upper and lower sums, that are often called Darboux or Riemann–Darboux sums, were introduced. The approach we followed in the first lecture was developed by several authors, including Gaston Darboux, Vito Volterra, and Giuseppe Peano [1]. It allowed us to prove criteria of integrability using analysis developed by the time of Riemann (to be honest, we also used notions of the supremum and infimum a lot).

5.3. **Rigorous constructions of the reals.** Bernard Bolzano published an article in 1817, where he demonstrated that each Cauchy sequence has exactly one limit in [3], but his argument appealed to geometrical intuition as all previous works on limits and continuity. The reason is very simple, real numbers were not properly defined at that time. First accurate treatments of real numbers appeared in the second part of 19th century in the works by Karl Weierstrass, Richard Dedekind, Georg Cantor, Eduard Heine. (We didn't have time to go through construction of real numbers from rational ones, and thus postulated the least upper bound theorem.) This development was used to give a rigorous proof of the Bolzano–Weierstrass theorem and completeness of the real line.

From that time "$\varepsilon - \delta$" language, which was introduced by Weierstrass, replaced the geometric considerations. The calculus became less visual but much more precise and structured. Heine, whose work was influenced by Weierstrass, pointed out clearly the difference between continuity and uniform continuity in 1874 and proved that each continuous function on a closed interval is uniformly continuous (the last result was also published a year earlier by Lüroth) [3].

Interchanging of integration with taking the limit was usual in earlier works (starting from Newton as we have seen, this operation was intensively used by several generations of mathematicians, including Euler). No question about the legacy of this interchanging was asked. The delicacy of the question was fully understood by Weierstrass, who also introduced the notion of uniform convergency. The statement of Exercises 3.3 is due to him [4].

5.4. **Stieltjes's extension of Riemann integration.** The integral, that we called the Riemann–Stieltjes integral, was defined by Thomas-Jean Stieltjes in 1894 [3]. He was working on the moment problem: given a sequence of positive numbers $\{\mu_n\}$, find a "mass distribution" ϕ on $[0, +\infty)$ such that

$$\int_0^\infty x^n \, d\phi(x) = \mu_n.$$

Function ϕ describes the distribution of masses on the half-line and is assumed to be increasing. Clearly, the Riemann integration was not sufficient when addressing this problem. Stieltjes defined integral sums and considered the integration of a continuous function with respect to an increasing one [3].

Functions of bounded variations were introduced by Camille Jordan in 1881, when he generalized the Dirichlet criterion for convergence of the Fourier series [5, Variation of a function]. He also proved the Jordan decomposition theorem. This notion was later employed by Jordan to define rectifiable curves.

REFERENCES

[1] C. H. Edwards, Jr., The Historical Development of the Calculus, Springer-Verlag, 1979
[2] A Source Book in Mathematics 1200-1800, edited by D. J. Struik, Cambridge, MA, Harvard University Press, 1969.
[3] A Source book in classical analysis, edited by G. Birkhoff, Cambridge, MA, Harvard University Press, 1973.
[4] W. Dunham, The calculus gallery, Princeton University Press, 2005.
[5] Encyclopedia of mathematics, Springer-Verlag, online http://eom.springer.de