Let $y_0 = a < y_1 < \cdots < y_{m-1} < y_m = b$ be the points in $[a,b]$ where f is discontinuous. Define $f_k(x)$ to be the restriction of f to the interval $[y_{k-1}, y_k]$, where at the endpoints f_k is defined to be the limit of its values from the interior points:

$$f_k(x) = \begin{cases} \lim_{x \to y_{k-1}^+} f(x) & \text{when } x = y_{k-1}, \\ f(x) & \text{when } x \in (y_{k-1}, y_k), \\ \lim_{x \to y_k^-} f(x) & \text{when } x = y_k. \end{cases}$$

Then we can define

$$\begin{aligned} \int_a^b f(x)\,\mathrm{d}x &= \int_a^{y_1} f_1(x)\,\mathrm{d}x + \int_{y_1}^{y_2} f_2(x)\,\mathrm{d}x + \cdots + \int_{y_m}^{b} f_m(x)\,\mathrm{d}x \\ &= \sum_{k=1}^{m} \int_{y_{k-1}}^{y_k} f_k(x)\,\mathrm{d}x. \end{aligned}$$

and, thus, can still define the definite (Cauchy) integral despite the discontinuities.

Example 10.7: Consider the function $f(x) = x$ on $[0,1]$. We know that the area under the curve is a triangle with area $\frac{1}{2}$. Let us show that the Cauchy integral gives us the same result. We take an n-partition of $[0,1]$, giving us partition points $x_k = \frac{k}{n}$, $k = 0, 1, \ldots, n$, so that $\Delta x = \frac{1}{n}$. On each I_k, the function $f(x) = x$ will attain its minimum on the left endpoint, and the maximum on the right end point. We therefore have

$$\begin{aligned} s_n &= \sum_{k=1}^{n} x_{k-1}\Delta x = \frac{1}{n}\left(0 + \frac{1}{n} + \frac{2}{n} + \cdots + \frac{n-1}{n}\right) = \frac{1}{n^2}\left(1 + 2 + \cdots + (n-1)\right) \\ &= \frac{1}{n^2}\frac{n(n-1)}{2} = \frac{1}{2}\frac{n-1}{n} \xrightarrow{n\to\infty} \frac{1}{2} \end{aligned}$$

and

$$\begin{aligned} S_n &= \sum_{k=1}^{n} x_k \Delta x = \frac{1}{n}\left(\frac{1}{n} + \frac{2}{n} + \cdots + \frac{n}{n}\right) = \frac{1}{n^2}\left(1 + 2 + \cdots + n\right) \\ &= \frac{1}{n^2}\frac{(n+1)n}{2} = \frac{1}{2}\frac{n+1}{n} \xrightarrow{n\to\infty} \frac{1}{2} \end{aligned}$$

which confirms that we obtain the correct result:

$$\int_0^1 x\,\mathrm{d}x = \frac{1}{2}.$$

10.5 Riemann integral

The problem with the Cauchy integral

In the previous section we saw that the limits of the lower and upper sums of a continuous function f over an interval $[a,b]$ exist and are equal to one another. Thus it was natural to conclude that this limit is indeed the area, which is called the definite integral of f over $[a,b]$ and is denoted $\int_a^b f(x)\,\mathrm{d}x$. The mathematical problem with this process is the following:

> The partitions we used were very specific: we split $[a,b]$ into n subintervals of equal length. *Could it be that the outcome was influenced by this?*

The solution

The solution to this question is to allow ***un*even partitions**. Instead of choosing an even n-partition with $x_0 = a$ and $x_k = x_0 + k\frac{b-a}{n}$, we take some points

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

which might *not* be equally distributed within the interval $[a,b]$. We denote, as before, the kth subinterval

$$I_k = [x_{k-1}, x_k], \qquad k = 1, \dots, n.$$

As before, a **refinement** is a further partition of an existing partition (only that now we do not require an even partition). In what follows we denote

$$I = [a,b].$$

Step functions

Definition 10.7: A function $f : I \to \mathbb{R}$ is called **step function** if it is constant on subintervals of I. We say that it is **adapted** to a particular partition $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ if it is constant on the interior of any of the subintervals I_k. We denote

$$\mathcal{S}(I) = \text{the set of all step functions on } I.$$

Definition 10.8: Let $f \in \mathcal{S}(I)$ and let $a = x_0 < x_1 < \cdots < x_n = b$ be an adapted partition. Denote by c_k the value of f on I_k. Then we define the **definite integral** of f on I to be

$$\int_I f = \sum_{k=1}^{n} c_k(x_k - x_{k-1}).$$

Proposition 10.9: Let $g, h \in \mathcal{S}(I)$ be two step functions on I and assume that $g(x) \le h(x)$ for all $x \in I$. Then

$$\int_I g \le \int_I h.$$

Proof. We take a partition $\{x_k\}_{k=0}^n$ that is adapted to both g and h (i.e., a partition that includes all points where either g or h have a jump). Then on any interval I_k there exist constants c_k, d_k such that

$$c_k = g(x) \le h(x) = d_k, \qquad \forall x \in I_k.$$

Then we have

$$\int_I g = \sum_{k=1}^{n} c_k(x_k - x_{k-1}) \le \sum_{k=1}^{n} d_k(x_k - x_{k-1}) = \int_I h.$$

□

Bounded functions

Let $f : I \to \mathbb{R}$ be a bounded function (not necessarily continuous or piecewise continuous). Define the sets $\mathcal{S}_f^+$ and $\mathcal{S}_f^-$ containing all step functions greater than f and less than f, respectively:

$$\mathcal{S}_f^+ = \left\{h \in \mathcal{S}(I) \;\mid\; f(x) \le h(x),\ \forall x \in I\right\}$$
$$\mathcal{S}_f^- = \left\{g \in \mathcal{S}(I) \;\mid\; g(x) \le f(x),\ \forall x \in I\right\}$$

These sets are not empty, since they contain, respectively, any constant function that is greater than the upper bound of f on I and any constant function that is smaller than the lower bound of f on I. We can therefore define:

Lower and upper integral

For $f : I \to \mathbb{R}$ that is bounded, we define the **upper integral of f on I** to be

$$\overline{\int_I} f = \inf\left\{\int_I h \;\mid\; h \in \mathcal{S}_f^+\right\},$$

and the **lower integral of f on I** to be

$$\underline{\int_I} f = \sup\left\{\int_I g \;\mid\; g \in \mathcal{S}_f^-\right\}.$$

Proposition 10.10: If f is bounded on I, then

$$\underline{\int_I} f \le \overline{\int_I} f$$

Proof. Let $g \in \mathcal{S}_f^-$ and $h \in \mathcal{S}_f^+$. Then:

$$g(x) \le f(x) \le h(x), \qquad \forall x \in I.$$

It follows that

$$\int_I g \le \int_I h.$$

Taking the infimum over all $h \in \mathcal{S}_f^+$ in this inequality gives us

$$\int_I g \le \overline{\int_I} f$$

and now taking the supremum over all $g \in \mathcal{S}_f^-$ gives us

$$\underline{\int_I} f \le \overline{\int_I} f.$$

□

The lower and upper integrals aren't necessarily equal

Consider the bounded function (called the *Dirichlet function*)

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Regardless of the partition $\{x_k\}_{k=0}^n$ of I, any subinterval I_k will include both rational and irrational points, and therefore f will attain the values 0 and 1 on any such subinterval. Hence $\mathcal{S}_f^+$ will contain step functions whose values are at least 1 and $\mathcal{S}_f^-$ will contain step functions whose values are at most 0. It is not hard to conclude that

$$\underline{\int_I} f = 0 \qquad \text{and} \qquad \overline{\int_I} f = 1.$$

This example demonstrates that it is not evident that the lower and upper integral should be equal. In fact, it motivates the following definition:

Riemann integrable functions

A bounded function $f : I \to \mathbb{R}$ is said to be (Riemann) **integrable** on I if

$$\underline{\int_I} f = \overline{\int_I} f.$$

This value, called the definite integral, is denoted $\int_a^b f(x)\,\mathrm{d}x$ or $\int_I f(x)\,\mathrm{d}x$.

Theorem 10.11: The following functions are (Riemann) integrable on I:

1. Continuous functions on I.
2. Piecewise-continuous functions on I.
3. Functions that are continuous on (a, b) and bounded on $[a, b]$.
4. Monotone functions on $[a, b]$.

Proof. We skip this proof. □

Example 10.8: The function $f(x) = x$ is Riemann integrable (it was also Cauchy integrable). We saw that the result of the Cauchy integral was $\int_0^1 x\,\mathrm{d}x = \frac{1}{2}$. The Riemann integral will give the same result, and to see that one could take the step functions

$$h_n(x) = \begin{cases} 0 & x = 0 \\ \frac{k+1}{n} & \frac{k}{n} < x \le \frac{k+1}{n},\ k = 0, \dots, n-1 \end{cases}$$