Lecture 6 : Integration of step functions

Let $f : [a,b] \longrightarrow \mathbb{R}_{\geq 0}$ be a function with domain the closed interval $[a,b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$.

Definition 1: The ordinate set of f is

$$Q_f := \{ (x,y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ 0 \leq y \leq f(x) \}.$$

Question: For what functions f is this "well-approximable", hence measurable — that is, $Q_f \in \mathcal{M}$ and $a(Q_f)$ is defined ?

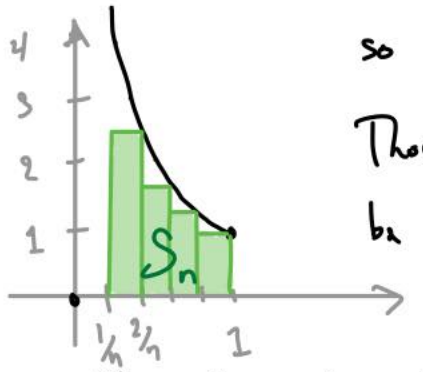
We'll see that the class of such functions is quite large in the next lecture; for now, we will consider a small class of functions. I'd also like to point out that it isn't true for just any function:

Ex/ On $[0,1]$, neither $f(x) := \begin{cases} 1 & \text{if } x \in \mathbb{Q} \quad (x \text{ is rational}) \\ 0 & \text{otherwise} \quad (x \text{ is irrational}) \end{cases}$

nor $g(x) := \begin{cases} 1/x, & x > 0 \\ 0, & x = 0 \end{cases}$ is well-approximable.

For f, the smallest step-region enclosing Q_f is $T = [0,1] \times [0,1]$, while the largest step-region contained in Q_f is $S = [0,1] \times \{0\}$. So there is certainly not a unique real number c satisfying $a(S) \leq c \leq a(T)$ for all step-regions S, T with $S \subseteq Q_f \subseteq T$: the largest $a(S)$ can be is 0 and the smallest $a(T)$ can be is 1.

For g, we have a different issue: we can choose the "lower step region" S to be the S_n in the picture below,



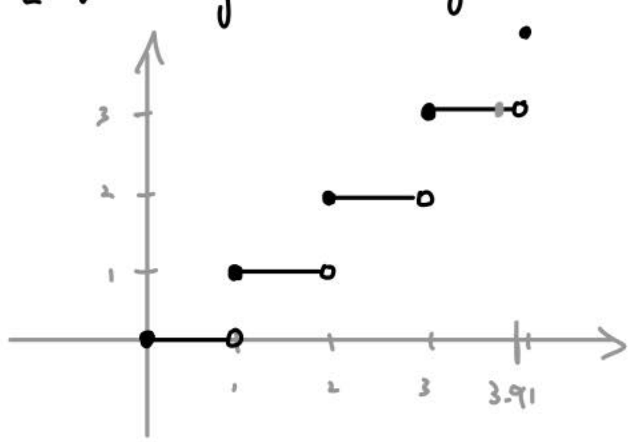
so that $a(S_n) = \sum_{j=2}^{n} \frac{1}{n} \cdot \frac{1}{\left(\frac{j}{n}\right)} = \sum_{j=2}^{n} \frac{1}{j}$.

Though we won't prove this now, this can be made arbitrarily large by taking n arbitrarily large. So in this case the problem is not uniqueness of c but existence of c: there is no number at all which is $\geq a(S)$ for every step region S with $S \subseteq Q_g$. //

<u>Definition 2</u>: A partition of $[a,b]$ is a finite set
$$P = \{x_0, x_1, \ldots, x_n\} \subset \mathbb{R}$$
of real numbers with $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$.
It determines a subdivision $[a,b] = \bigcup_{i=1}^{n} [x_{i-1}, x_i]$.

<u>Definition 3</u>: A step function on $[a,b]$ is a function $s : [a,b] \to \mathbb{R}$ which is constant on each open subinterval (x_{i-1}, x_i) of some partition P of $[a,b]$. (Write s_i for this constant value. Note that s may take completely unrelated values at the points x_i.)

Ex/ $s(x) = [x] :=$ greatest integer $\leq x$. Its graph on $[0,4]$:



e.g. $[3.91] = 3$

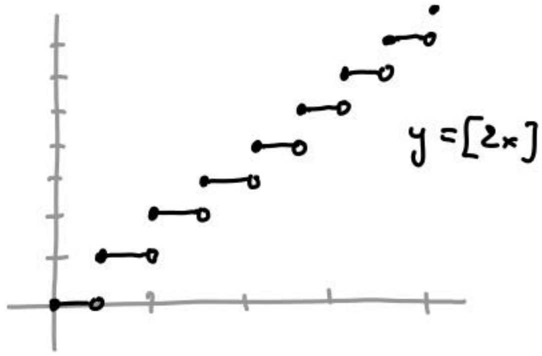
If we change s so that $s(3) = 2{,}718{,}281$, it is still a step function (according to Def. 3). //

Note that if s takes nonnegative values on $[a,b]$, and (for each i) $s(x_i)$ is the larger of s_i (the value on (x_{i-1}, x_i)) and s_{i+1} (the value on (x_i, x_{i+1})), then the ordinate set Q_s is a step region, i.e. a union of closed rectangles. But even if we change the values at the points $\{x_i\}$, that only adds or subtracts line segments from Q_s, which doesn't affect its area.

<u>Definition 4</u>: The definite integral of a step function s on $[a,b]$ is $\displaystyle\int_a^b s(x)\,dx := \sum_{i=1}^n (x_i - x_{i-1})\cdot s_i$ $\left(= a(Q_s) \text{ if } s \geq 0 \text{ on } [a,b]\right)$.

Ex/ $$\int_0^4 [x]\,dx = (1-0)\cdot 0 + (2-1)\cdot 1 + (3-2)\cdot 2 + (4-3)\cdot 3 = 1+2+3 = 6.$$ //

Ex/ $$\int_0^4 [2x]\,dx = ?$$



In this case, the value jumps each time we add $\frac{1}{2}$ to x. So we have to use the finer partition $P = \{0, \frac{1}{2}, 1, \frac{3}{2}, \ldots, 4\}$ to cast $s(x) = [2x]$ as a step function. The integral is then $\frac{1}{2}\cdot(0+1+2+3+\cdots+7) = 14$. //

Ex/ $$\int_0^n [t]^2\,dt = 1\cdot\left(0^2+1^2+2^2+\cdots+(n-1)^2\right) = P_2(n-1) = \frac{n(n-1)(2n-1)}{6}.$$ //

("variable of integration" is just a dummy variable — changing its name does nothing.)

But does the definition make sense? Given s, there are tons of partitions P on whose subintervals s is constant: just add more points to P — this is called refining the partition. This had better not affect the value of $\int_a^b s(x)\,dx$!

Fortunately, it doesn't: inserting "y" between x_{i-1} & x_i merely subdivides the rectangle $[x_{i-1}, x_i] \times [0, s_i]$ and "changes" $(x_i - x_{i-1}) \cdot s_i$ to $(x_i - y) \cdot s_i + (y - x_{i-1}) \cdot s_i$ (i.e. changes nothing!). We say that $\int_a^b s(x)\,dx$ is "well-defined", i.e. depends only on a, b, & $s(x)$ itself.

Properties of the definite integral of a step function:

① $\int_a^b c \cdot s(x)\,dx = c \cdot \int_a^b s(x)\,dx$ ("homogeneous property")

② $\int_a^b (s(x) + t(x))\,dx = \int_a^b s(x)\,dx + \int_a^b t(x)\,dx$ ("additive property")

③ If $s(x) < t(x)$ for each $x \in [a,b]$, then $\int_a^b s(x)\,dx < \int_a^b t(x)\,dx$ ("comparison property")

④ If $a \leq b \leq c$, then $\int_a^b s(x)\,dx + \int_b^c s(x)\,dx = \int_a^c s(x)\,dx$ ("additivity with respect to the interval")

⑤ $\int_a^b s(x)\,dx = \int_{a+c}^{b+c} s(x-c)\,dx$ ("translation invariance")

⑥ $\int_{ka}^{kb} s\left(\frac{x}{k}\right) dx = k \int_a^b s(x)\,dx$ ("expansion/contraction property")

(We also set $\int_b^a s(x)\,dx := -\int_a^b s(x)\,dx$ if $a < b$.)

<u>Proofs</u>: You can prove all of these just using the definition, but a geometric proof is more natural for some of them:

- in ⑤, the LHS (left-hand side) is $a(Q_s)$ and the RHS is $a(\tau(Q_s))$, where τ is translation by c units to the right. By translation-invariance of area, these areas are equal.

- proving ④ is part of HW #2. You can also do this one geometrically.
- ⑥: if $s(x)$ has underlying partition $\{x_0, x_1, \dots, x_n\}$ then $s\left(\frac{x}{k}\right)$ has partition $\{kx_0, kx_1, \dots, kx_n\}$ (of $[ka, kb]$). The LHS is $\sum_{i=1}^{n} (kx_i - kx_{i-1}) \cdot s_i = k \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot s_i$ = the RHS.
- ② & ③ use the same (important) idea: the common refinement of two partitions P and P' means combining all the points: $P \cup P'$. If s is constant on subintervals of P, and t is constant on subintervals of P', then s & t are both constant on subintervals of $P \cup P' = \{y_0, y_1, y_2, \dots, y_N\}$. So therefore is $s+t$, and ② simply reduces to $\sum_{i=1}^{N} (y_i - y_{i-1})(s_i + t_i) = \sum_{i=1}^{N} (y_i - y_{i-1}) s_i + \sum_{i=1}^{N} (y_i - y_{i-1}) t_i$; while ③ is $\sum_{i=1}^{N} (y_i - y_{i-1}) s_i < \sum_{i=1}^{N} (y_i - y_{i-1}) t_i$. $\square$