

The Lebesgue Integral

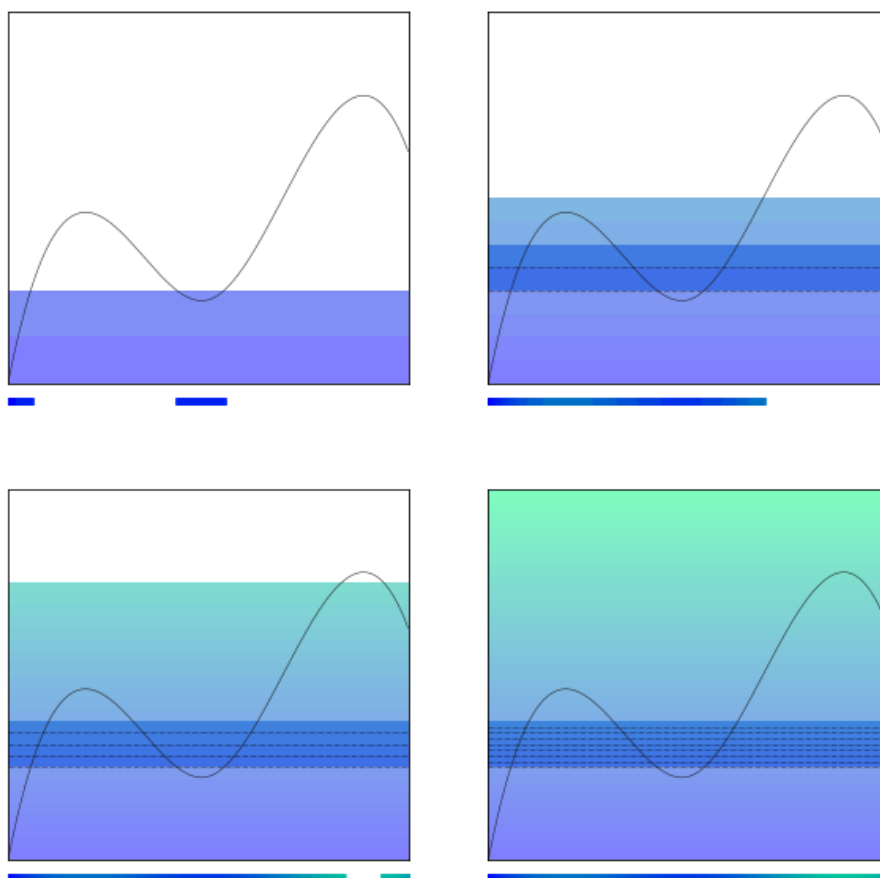
*From σ -algebras to simple functions;
building up to defining the Lebesgue Integral*

Author: Daniel Ajibola Adelodun

Supervisor: Professor Gromov

Third Year Project

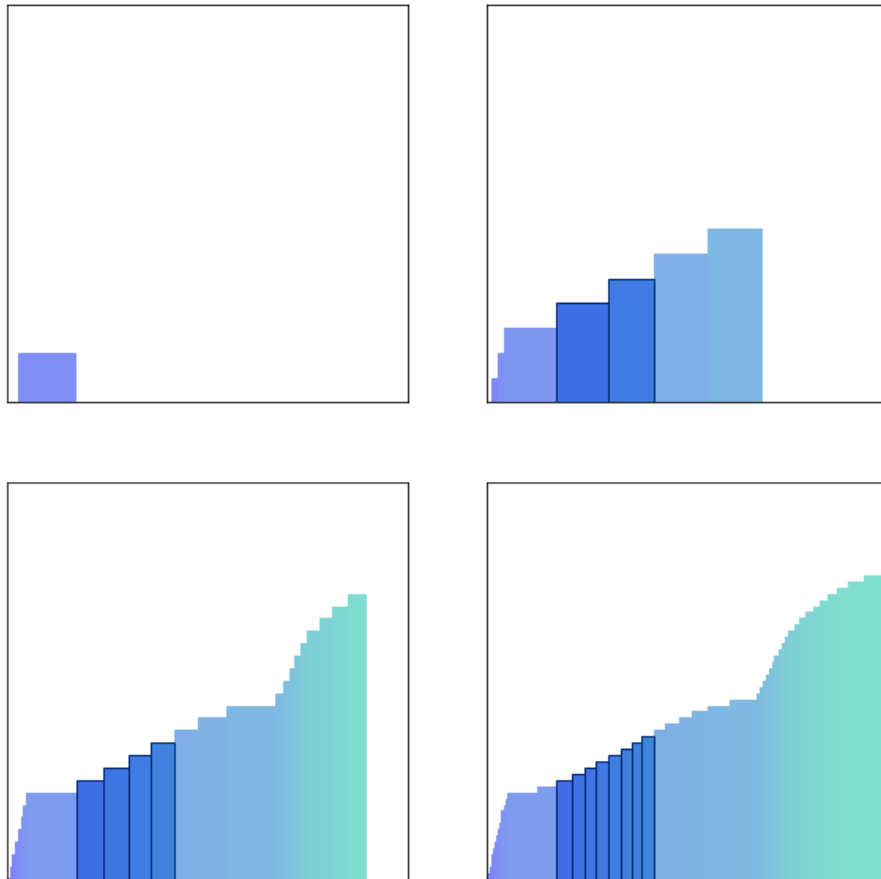
19th April 2019



Abstract

In this paper we work our way up to defining the *Lebesgue Intergral* by introducing relevent results from Set Theory and Measure Theory. We make comparisions between the Lebesgue Intergral and the Reimann Intergral, noting the relative streghths and weaknesses of each.

Over the course of the paper, we will touch on other uses of Measures Theory, including it's use in Axiomatic Set Theory in the definition of a random variable. Finally, we will look very briefly at \mathcal{L}^p spaces; we'll simply state what they are and make vague but interesting statements about what it means for the \mathcal{L}^p s to be normed vector spaces, and that \mathcal{L}^2 in particular is an inner product space.



Contents

1	Introduction	4
1.1	The Definition	4
1.2	Why Not Just Stop Here?	4
2	Riemann vs. Lebesgue	5
2.1	The Motivation	5
2.2	The Definition of the Riemann Integral	6
2.3	The Indicator function $\mathbb{1}_{\mathbb{Q}}(x)$	9

1 Introduction

1.1 The Definition

The aim of this paper is to eventually define the Lebesgue Integral - so lets try to do that...

Definition 1.1 (*Lebesgue Integration*). Given

- a Measure Space; (X, \mathcal{A}, μ) ,
- an \mathcal{A}/\mathcal{B} Measurable Function $u: X \rightarrow \bar{\mathbb{R}}$,
- the two positive functions u^+ and u^- s.t. $u = u^+ - u^-$,
- the two sets S^+ and S^- of *all* positive Simple Functions ($S^\pm \subset \mathcal{E}_\mu^+$) on (X, \mathcal{A}, μ) s.t. $s \in S^\pm \Rightarrow s \leq u^\pm$,

we define the μ -integral of u - which we write as $\int u d(\mu)$ - in terms of the integrals of the positive functions u^+ and u^- .

The integral of a general *positive* function (so including both u^- and u^+ but not necessarily u itself) is defined as

$$\int u^+ d(\mu) := \sup\{I_\mu(s) : s \in \mathcal{E}^+ \text{ and } s \leq u\}, \quad (1.1')$$

and going back to our more general u ,

$$\int u d(\mu) := \int u^+ d(\mu) - \int u^- d(\mu). \quad (1.1)$$

If the domain of our function is \mathbb{R}^n , with the 'standard' n -dimensional Borel σ -algebra and Lebesgue Measure on it - i.e. $(X, \mathcal{A}, \mu) = (\mathbb{R}^n, \mathcal{B}, \lambda^n)$ - then this becomes the λ^n -integral (read as Lebesgue integral instead of lambda integral - or we just call it 'the integral') and we write

$$\int u d(\lambda^n) = \int u(x) d(\lambda^n) := \int u(x) d(x) \quad (1.1'')$$

1.2 Why Not Just Stop Here?

That's quite a long definition, one which raises more questions than it answers. The biggest issue here is that even assuming you've taken a course in Analysis, there are many undefined terms;

- What are *Measures* and σ -algebras? And what is a *Measure Space*?
- What is the *Borel* σ -algebra
- What is a *Measurable Function*? What is a *Simple Function*?
- What is I_μ and why can it seemingly only be used on positive *Simple Functions*?

That last bullet point touches on the second key issue with this definition; why. Why any of this. What about this definition of integration makes it any better than the *Riemann Integral*? Do they even describe similar things? This is not *at all* immediately obvious, since the Riemann integral is defined on functions $f: Y \rightarrow \mathbb{R}$ where Y is some compact¹ subset of \mathbb{R} , and this integral is defined on...well, any *Measurable Function* on any set at all, apparently.

Even if we can decipher this definition, and then somehow convince ourselves they are compatible, it still seems like a convoluted definition - like, what about Simple Functions is so special that the integrals of all other functions are defined in terms of them (or to be more precise, why is the integral of a function defined in terms of the integrals of positive functions, and then the integrals of positive functions defined in terms of positive simple ones)?

Interestingly (to me, at least) the Riemann integral is defined in a similarish way - we find the upper Riemann integral of a function by finding all the *Step Functions* which bound the function from above, define/find the integrals of these using their 'steppyness', then take the infimum - we do the same thing from below, and then if these limits coincided then we call that the Riemann integral. In order to understand the motivations behind Lebesgue Integration better, let's take a slightly closer look at the similarities and differences between these two.

¹any closed and bounded interval - so an interval in the form $[a, b]$

2 Riemann vs. Lebesgue

2.1 The Motivation

Our goal when we integrate a function is to find the amount of ‘space’ between the graph and x-axis (or the Abscissa if you’re fancy). It’s pretty difficult to say how much space an arbitrary curve takes up, but we can work out the area of rectangles very easily — by calculating the *base* \times *height*. Riemann integration takes advantage of this, and defines the integral of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ by first bounding it from above with rectangular functions and finding the smallest of these areas (which would be an upper bound for the area of the function), then we bound the function from below and find an lower bound, then we hope that these two bounds match up.

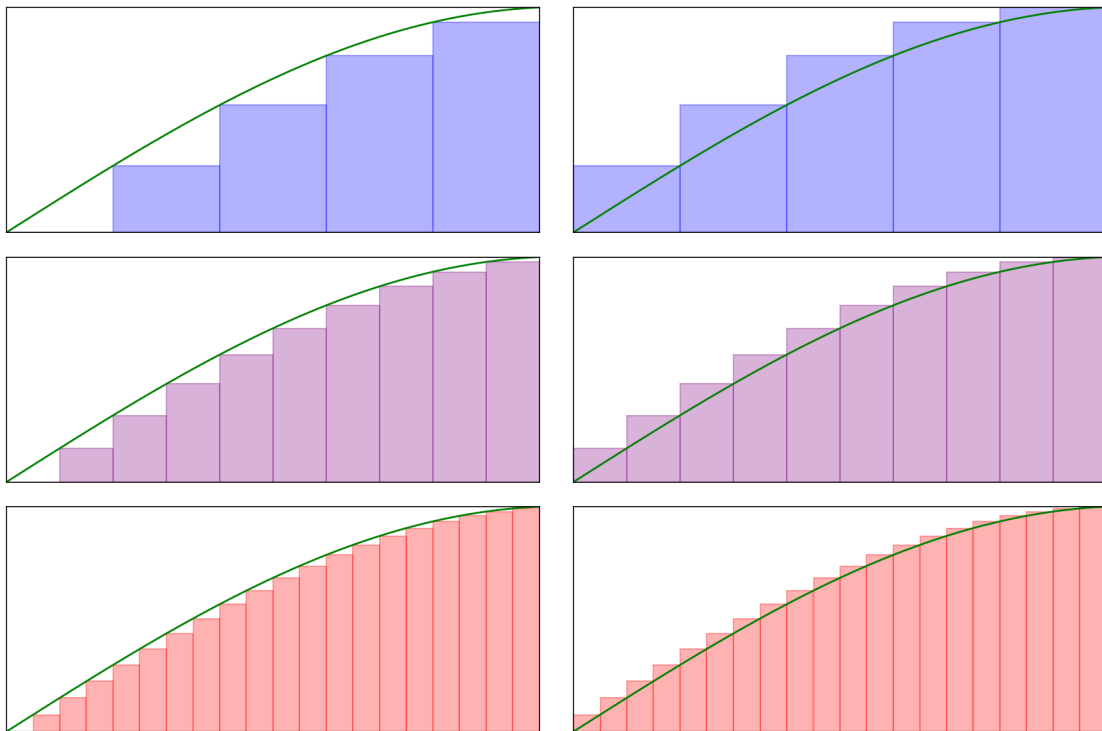


Figure 2.1: Bounding from Below

Figure 2.2: Bounding from Above

If the domain of our function (the ‘x-axis’), is \mathbb{R}^2 instead of just \mathbb{R} , we’d actually calculating a volume by finding the ‘space’ between the graph and the axis. Curves are bad enough, but it’s even harder to find the volume under a surface. Basic Riemann integration doesn’t apply to these functions — but we can still make a similar simplification and think of the ‘area of a rectangle’ (which in $\mathbb{R}^2 \rightarrow \mathbb{R}$ would actually be the volume of a cuboid) as being calculated in much the same way as in \mathbb{R} ; the size of the base (which is now an area, instead of a length) \times the height. It’d be nice if there were a way to take advantage of this so that we could integrate any function where we can make sense of ‘the size of the base.’ With this in mind, let’s clear up some of the terms we are going to use:

Lebesgue integration is the method of integration defined above. Just like Riemann Integration, it is used to find the space between the graph of a function and the domain. However, unlike Riemann, the domain can be any set at all. Thinking back to rectangles and cuboids and *base* \times *height*, if we wanted make sense of what it means for there to be space between the graph and the x-axis, we should be able to make sense of the ‘size’ of parts of the domain, so that we can figure out how big our base is and multiply that by the height...

The Lebesgue Measure seems then like it’d be a way to assign sizes to, or ‘measure’ sets any domain; but it is not! It’s a term used specifically when we give sizes to the subsets of \mathbb{R}^n , and those sizes (or ‘measures’) coincided with our usual idea of the size of a set in \mathbb{R}^n — i.e. it is used when we are considering functions $\mathbb{R}^n \supset Y \rightarrow \mathbb{R}$, and in \mathbb{R} the measure (length) of the interval $[1, 0]$ is 1, in \mathbb{R}^2 the measure (area) of the rectangle $[1, 0] \times [1, 0]$ is 1, in \mathbb{R}^3 the measure (volume) of the cube $[1, 0] \times [1, 0] \times [1, 0]$ is 1... etc.

The Lebesgue Integral is what we get when we do Lebesgue Integration on a set with the Lebesgue Measure on it. i.e. it's the integral of normal functions $\mathbb{R}^n \supset Y \rightarrow \mathbb{R}$. This is the (basically) the same domain as the Riemann integral and so we'd want to check that the Lebesgue Integral and the Riemann integral match up — and then see if the Lebesgue integral is better in some way.

Already, we can see an improvement over Riemann Integration; *Lebesgue Integration* is defined on any set at all, and even the more restricted *Lebesgue Integral* can be calculated directly for any \mathbb{R}^n — so no double/triple integration nonsense. To see another benefit of the Lebesgue integral over Riemann, let's look at the canonical example of a function for which the Riemann integral fails; $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = \mathbb{1}_{\mathbb{Q}}(x)$ (Figure 2.3). This is also an example of a *Simple Function* since it is just the *Indicator Function* of a measurable set, but let's not get ahead of ourselves... First, we should define the Riemann integral.

2.2 The Definition of the Riemann Integral

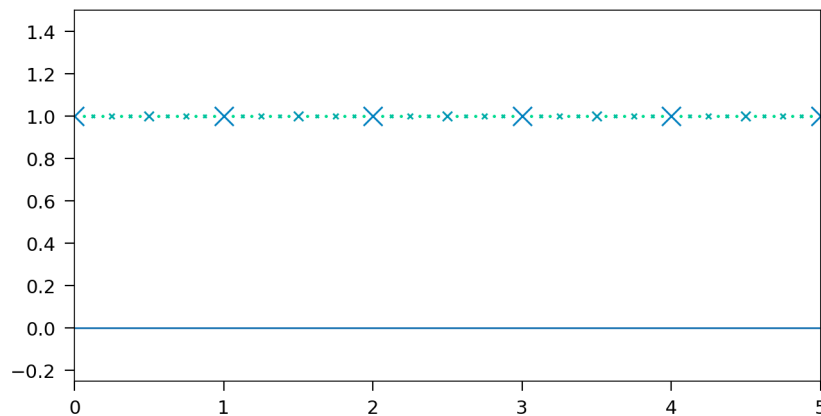


Figure 2.3: An illustration of the indicator function $\mathbb{1}_{\mathbb{Q}}(x)$, showing a few rationals that are multiples of negative powers of 2.

The definition of the Riemann integral involves a few steps, and so it's going to take a little patience. The first step is defining a step function (no pun intended).

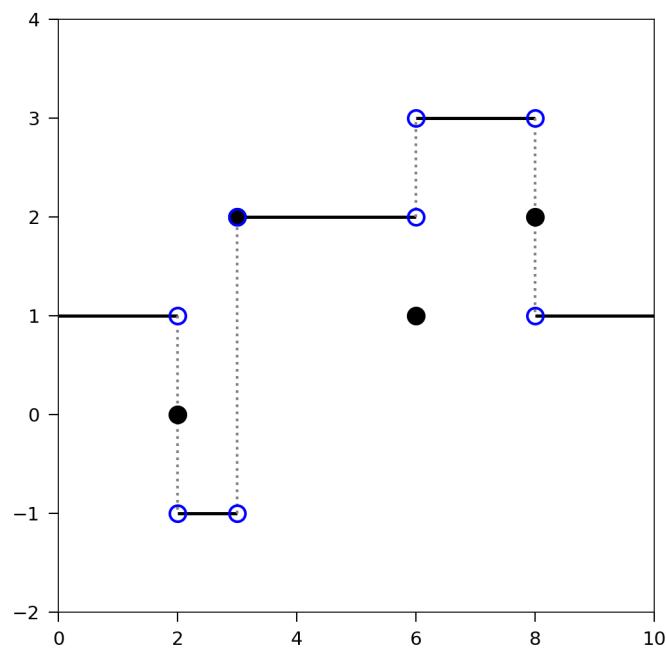


Figure 2.4: A Step Function

A function is called a step function if and only if there exists some finite sequence of points such that the function is constant between any two adjacent points. So, for the function with the graph shown below in Figure 2.4, the points $\{0, 2, 3, 6, 8, 10\}$ are the finite sequence which make this a step function. We will often index the elements of sequence using n , and label each element x_n . We can call sequence this a ‘*representation*’ of the function, and we can call the sequence of constant heights just *before* each point the corresponding ‘*height sequence*’. The *height sequence* along with the representation fully describe the function.

Note two important points;

- We could have picked $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ to be our representation, since the function also happens to be constant between any two adjacent integers.
- It doesn’t really matter what the function is *at* the points x_n — we only care whether the function is constant *between* the points. In our graph, 2, 3, 6, and 8 are discontinuities, but that doesn’t matter since they’re also in our sequence of points.

The formal definition is as follows;

Definition 2.1 (Step Function). $\phi: [a, b] \rightarrow \mathbb{R}$ is a *step function*

$$:\Leftrightarrow \phi \in \mathcal{S}([a, b])$$

$$:\Leftrightarrow \exists N \in \mathbb{N} \text{ and } \exists X = (x_n)_{n=0}^N : a = x_0 < x_1 < \dots < x_{N-1} < x_N = b, \text{ and } \forall m \in \mathbb{N}_{<N} \text{ we have } \exists! c_m \in \mathbb{R} : \forall z \in (x_{m-1}, x_m) \text{ we have } \phi(z) = c_m.$$

This also defines *the set of all step functions on $[a, b]$* as $\mathcal{S}[a, b]$. This kind of thing is done a lot in analysis — we often want to look at the set of all things which satisfy some property. We normally don’t have to *find* them all — we just need a way to represent the full set. The fact that it’d be near impossible to try to imagine the set $\mathcal{S}([a, b])$ is irrelevant — it’s just important we know it exists.

Next, we want to define what it means to integrate a step function. Remember, the point of working with step functions is because it’s easy to find the area of rectangles, and integration is all about finding areas. So where are the rectangles?

As is hinted at by Figure 2.4, we can make the bases of the rectangles the width the intervals (x_{n-1}, x_n) , and the heights the constant c which values in this interval take. This is made even clearer by Figure 2.5, where we can also see that the sometimes our heights will have negative values, and that’s okay — these areas are just taken away from the total area instead of added to it.

What’s not okay is that any one step function may have different sequences which we can use to prove that it’s a step function — or in other words, the same step function may have different but equally valid representations. We defined the integral of a step function in terms of its representation, but how do we know that different representations of the same function produce the same integral? Figure 2.5 demonstrates that, given the representations $\{0, 2, 3, 6, 8, 10\}$ and $\{0, 1, 2, \dots, 10\}$ of 2.4, we’re fine, and the areas are equal. It turns out that this is the case in general — the integral of a step function is independent of the chosen representation.

Definition 2.2 (Riemann Integral of a Step Function). Given

- $\phi \in \mathcal{S}([a, b])$,
- $N \in \mathbb{N}$ and the representation $(x_n)_{n=0}^N$ of ϕ ,
- the corresponding *height sequence* $(c_n)_{n=1}^N : \phi(x) = c_n \quad \forall x \in (x_{n-1}, x_n)$,

then the Riemann Integral of ϕ (as a function of x) on $[a, b]$ is

$$\begin{aligned} & \int_a^b \phi(x) dx \\ &:= \sum_{i=1}^N c_i (x_i - x_{i-1}). \end{aligned} \tag{2.1}$$

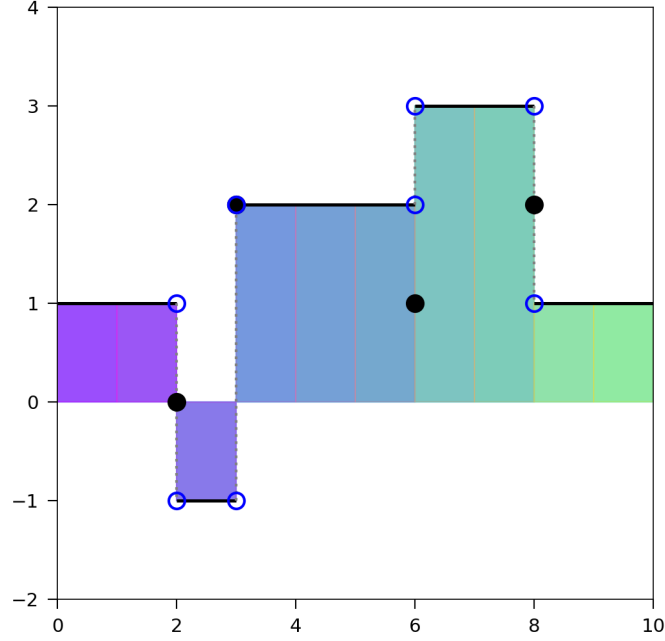


Figure 2.5: Integral of 2.4

To show that the integral is independent of the choice of representation, we'd need to show that if we had two arbitrary representations $(x_n)_{n=0}^N$ and $(x'_m)_{m=0}^M$ with height sequences $(c_n)_{n=1}^N$ and $(c'_m)_{m=1}^M$ resp. then

$$\sum_{i=1}^N c_i(x_i - x_{i-1}) = \sum_{j=1}^M c'_j(x'_j - x'_{j-1})$$

. The idea behind why this is true is the following;

- × First, we will focus on the representation that is more ‘fine’ — i.e. the one that has smaller width rectangles; ones which are divisions of the rectangles in the ‘coarser’ representation.
- × We look at it in sections; with the sections being the steps of the coarser representation. So, in 2.4, we'd look at $\{0, 1, 2 \dots, 10\}$ between 0 and 2, then 2 and 3, then 3 and 6 ... etc.
- × The two representations are for a single the step function, and we use this group adjacent rectangles with the same height. $\phi(x)$ must be constant in the section we are looking at, we know that all the rectangles here have the same height
- × Then we can use the distributive property of multiplication to view the smaller width rectangles of the same height as one big rectangle with the width of the whole section. And boom, we're done, since those big rectangles are exactly the ones which you would use when integrating using the coarser representation.
- × If neither representation is contained within the other, then we can just create a third representation which is the union of the first ones. Both original sequences will be coarser than this new one, and so we can use the same reasoning as before, plus transitivity and symmetry, to prove the integrals are the same.

Proposition 2.1. *Given,*

- $\phi \in \mathcal{S}([a, b])$,
- $N, M \in \mathbb{N}$ and the representations $(x_n)_{n=0}^N$ and $(x'_m)_{m=0}^M$ of ϕ ,
- The respective height sequences $(c_n)_{n=1}^N$ and $(c'_m)_{m=1}^M$ of these representations,

then

$$\sum_{i=1}^N c_i(x_i - x_{i-1}) = \sum_{j=1}^M c'_j(x'_j - x'_{j-1}). \quad (2.2)$$

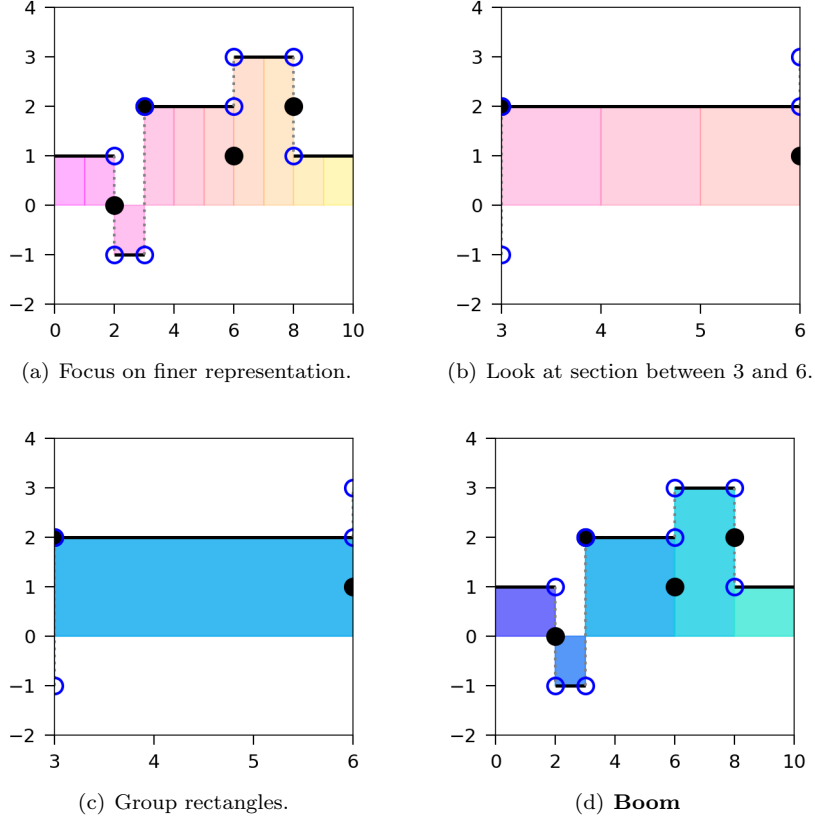


Figure 2.6: Different representations of the same function.

Proof.

Case 1: $\{x_n; n \in \mathbb{N}_{<N}\} \subset \{x'_m; m \in \mathbb{N}_{m < M}\}$:

$$\begin{aligned}
 & \{x_n; n \in \mathbb{N}_{<N}\} \subset \{x'_m; m \in \mathbb{N}_{m < M}\} \\
 \Leftrightarrow & \exists J = (j_n)_{n=0}^N \\
 & : 0 = j_0 < j_1 < \dots < j_{n-1} < j_n = m \text{ and } k \in \mathbb{N}_{<N} \text{ we have } x_k = x'_{j_k}
 \end{aligned}$$

Hence,

$$\begin{aligned}
 \sum_{j=1}^M c'_j(x'_j - x'_{j-1}) &= \sum_{k=0}^N \sum_{j=j_{k-1}}^{j_k} c'_j(x'_j - x'_{j-1}) \\
 &= \sum_{k=0}^N c'_{j_k} \sum_{j=j_{k-1}}^{j_k} (x'_j - x'_{j-1}) \\
 &= \sum_{k=1}^N c_k(x_k - x_{k-1})
 \end{aligned}$$

Case 2: Arbitrary representations $X = \{x_n; n \in \mathbb{N}_{<N}\}$ and $X' = \{x'_{n'}; n' \in \mathbb{N}_{<N'}\}$:

Let $X'' = \{x''_{n''}; x''_{n''} \in X \cup X'\}$.

$$\begin{aligned}
 & \{x_n; n \in \mathbb{N}_{<N}\} \subset \{x''_{n''}; n'' \in \mathbb{N}_{<N''}\} \supset \{x'_{n'}; n' \in \mathbb{N}_{m < M}\} \\
 \Rightarrow & \sum_{X'} = \sum_{X''} = \sum_{X'}
 \end{aligned}$$

□

2.3 The Indicator function $\mathbb{1}_{\mathbb{Q}}(x)$