

## INTEGRATION AND DEFINITE INTEGRAL: INTRODUCTION

MATH 152, SECTION 55 (VIPUL NAIK)

**Corresponding material in the book:** Section 5.1, 5.2.

**Difficulty level:** Hard if you have not seen this before. Medium if you have.

**What students should definitely get:** The definitions of partition, upper sum, and lower sum. A rough idea of what it means to take finer partitions and how this limiting process can be used to define integrals.

**What students should hopefully get:** The intuition behind an integral as an infinite summation; how it measures cumulative quantities. The intuitive relation with the area of a curve.

### EXECUTIVE SUMMARY

Words ...

- (1) The definite integral of a continuous (though somewhat weaker conditions also work) function  $f$  on an interval  $[a, b]$  is a measure of the signed area between the graph of  $f$  and the  $x$ -axis. It measures the *total value* of the function.
- (2) For a partition  $P$  of  $[a, b]$ , the lower sum  $L_f(P)$  adds up, for each subinterval of the partition, the length of that interval times the minimum value of  $f$  over that interval. The upper sum adds up, for each subinterval of the partition, the length of that interval times the maximum value of  $f$  on that subinterval.
- (3) Every lower sum of  $f$  is less than or equal to every upper sum of  $f$ .
- (4) The *norm* or *size* of a partition  $P$ , denoted  $\|P\|$ , is defined as the maximum of the lengths of its subintervals.
- (5) If  $P_1$  is a finer partition than  $P_2$ , i.e., every interval of  $P_1$  is contained in an interval of  $P_2$ , then the following three things are true: (a)  $L_f(P_2) \leq L_f(P_1)$ , (b)  $U_f(P_1) \leq U_f(P_2)$ , and (c)  $\|P_1\| \leq \|P_2\|$ .
- (6) If  $\lim_{\|P\| \rightarrow 0} L_f(P) = \lim_{\|P\| \rightarrow 0} U_f(P)$ , then this common limit is termed the *integral* of  $f$  on the interval  $[a, b]$ .
- (7) We can define  $\int_a^b f(x) dx$  as above if  $a < b$ . If  $a = b$  the integral is defined to be 0. If  $a > b$ , the integral is defined as  $-\int_b^a f(x) dx$ .
- (8) A continuous function on  $[a, b]$  has an integral on  $[a, b]$ . A piecewise continuous function where one-sided limits exist and are finite at every point is also integrable.

Actions ...

- (1) For constant functions, the integral is just the product of the value of the function and the length of the interval.
- (2) Points don't matter. So, if we change the value of a function at one point while leaving the other values unaffected, the integral does not change.
- (3) A first-cut lower and upper bound on the integral can be obtained using the *trivial* partition, where we do not subdivide the interval at all. The upper bound is thus the maximum value times the length of the interval, and the lower bound is the minimum value times the length of the interval.
- (4) The finer the partition, the closer the lower and upper bounds, and the better the approximation we obtain for the integral.
- (5) A very useful kind of partition is a *regular partition*, which is a partition where all the parts have the same length. If the integral exists, we can calculate the actual integral as  $\lim_{n \rightarrow \infty}$  of the upper sums or the lower sums for a regular partition into  $n$  parts.
- (6) When a function is increasing on some parts of the interval and decreasing on other parts, it is useful to choose the partition in such a way that on each piece of the partition, the function is either

increasing throughout or decreasing throughout. This way, the maximum and minimum occur at the endpoints in each piece. In particular, try to choose all points of local extrema as points of partition.

## 1. MOTIVATION AND BASICS

In this lecture, we will introduce some of the ideas behind integration. Integration is a continuous analogue of summation (or adding things up) with a few additional complications because of the infinitely divisible nature of the real line.

**1.1. Summation: numerically.** Suppose we have a real-valued function  $f$  defined on all integers. Given any two integers  $a < b$ , we can legitimately ask for the sum of the values of  $f(n)$  for all  $n$  in the interval  $[a, b)$  (including  $a$ , excluding  $b$ ). This can be interpreted as the total value of  $f$  on this interval. This summation poses no problems because we are adding finitely many real numbers. We could alternatively be interested in the sum of the values of  $f(n)$  for all  $n$  in the interval  $(a, b]$  (excluding  $a$ , including  $b$ ).

To make things simpler, we introduce a notation for summation. This notation is something we will pick up again much later, so for now this is just as a temporary device, and not something you need to learn. The notation is:

$$\sum_{n=a}^{b-1} f(n)$$

This notation means that we add up the values of  $f(n)$  for all  $n$  starting from  $n = a$  and ending at  $n = b - 1$ . In this case,  $a$  is the lower limit of the summation,  $b - 1$  is the upper limit of the summation, and  $f(n)$  is the summand.

This is also sometimes written as:

$$\sum_{a \leq n < b} f(n)$$

This means that the sum is over all the integers  $n$  satisfying  $a \leq n < b$ . The expression  $a \leq n < b$  can be replaced by any condition that restricts the  $n$  to certain integers.

**1.2. Summations: graphically.** We can think of these summations graphically as *areas*. For the first area (the summation on  $[a, b)$ ), consider the following: for each integer  $n$ , draw a rectangle with base on the  $x$ -axis from  $n$  to  $n + 1$  and height  $f(n)$ . The total area above  $[a, b)$  is the summation of the values of  $f(n)$  on  $[a, b)$ . Note that there's a little caveat: rectangles with negative height are given a negative area.

For the other case ( $(a, b]$ ), we make the rectangle from  $n - 1$  to  $n$  with height  $f(n)$ , i.e., the rectangle height is given by the value of the function at the *right end* of the rectangle.

This suggests some relationship between summations and areas. Here's one way to think about it. The area is the sum of the lengths of all the vertical slices of the figure, with each vertical slice length weighed by how much horizontal length it continues for. Thus, if the vertical length is 3 for a horizontal length of 2 and then 4 for a horizontal length of 1, the total area is  $(3 * 2) + (4 * 1) = 10$ .

**1.3. Piecewise constant functions: integration.** We now try to define a notion of integration for piecewise constant functions. What this notion of integration should do is measure the total value of the function, based on the ideas that we discussed above. Geometrically, it measures the *signed area* between the graph of the function and the  $x$ -axis, with a negative sign when the graph of the function is below the  $x$ -axis.

- (1) For each interval  $[a, b]$  where the function takes a constant value  $L$ , the integral on that interval is  $L(b - a)$ .
- (2) The overall integral is the sum of the integrals on each of the pieces where it is constant.

This makes sense geometrically – we are breaking the area to be measured into rectangles and then finding the area of each rectangle as the product of its height  $L$  and base length  $b - a$ .

For instance, consider the signum function, which is  $-1$  for  $x$  negative,  $0$  at  $0$ , and  $+1$  for  $x$  positive. The integral of this function on the interval  $[-3, 7]$  is  $(-1) * (0 - (-3)) + (1) * (7 - 0) = -3 + 7 = 4$ .

**1.4. Extending the idea to other functions.** We want to define a notion of integration for a function over an interval when the function is not piecewise constant, such that:

- (1) This notion measures what we intuitively think of as the area between the curve and the  $x$ -axis, with suitable signs: a positive contribution for the regions where the curve is above the  $x$ -axis and a negative contribution for the regions where the curve is below the  $x$ -axis.
- (2) This notion measures some kind of *total value* of the function.
- (3) If we subdivide the interval into smaller intervals, the integral over the whole interval is the sum of the integrals over the smaller intervals.

Our goal is to find something that roughly satisfies all these properties. We do, however, need to qualify the kinds of functions that we are willing to consider, because it is not possible to define a notion of integral for every function in a consistent and intuitive manner. One thing that seems to be desirable when trying to integrate is *continuity* – for a well-defined region to take the area of, the graph of the function should not randomly jump about. A slightly weaker formulation, *piecewise continuity*, will also do. Piecewise continuous means that there are only finitely many points of discontinuity. A piecewise continuous function can be integrated if it has the property that one-sided limits exist and are finite at all points of discontinuity. (Some other piecewise continuous functions can be integrated)

The way we integrate them is to break the interval into subintervals where the function is continuous, integrate the function on those subintervals, and then add up the values.

**1.5. Points and zero length idea.** In the case of finite sums, changing the value at any single point changes the final sum. However, when dealing with integration, the picture is a little different. The value of the function at a particular point  $a$  makes a very small contribution – in fact a zero contribution, to the integral. This is because the rectangle corresponding to the interval  $[a, a]$  has base length zero. Thus, changing the value of the function at just one point, without changing it elsewhere, has no effect on the integral. Another way of saying this is that our sample size, or base of aggregation, is so large, that measurement errors in one data point have no effect on the final answer.

**1.6. Brief note on terminology and notation.** If  $a < b$  and  $f$  is a function defined on  $[a, b]$ , we use the notation:

$$\int_a^b f(x) dx$$

Here,  $f$  is termed the *integrand* or the *function being integrated*,  $a$  and  $b$  are termed the *limits of integration*, with  $a$  the *lower limit* and  $b$  the *upper limit*,  $x$  is the variable of integration, and  $[a, b]$  is the *interval of integration* (also called the *domain of integration* or *region of integration*). The answer that we get is termed the *integral* of  $f$  over  $[a, b]$  or the integral of  $f$  from  $a$  to  $b$ . This integral is also sometimes called a *definite integral*, to distinguish it from indefinite integrals, that we will encounter later.

As already noted, the value of the function at any one point is irrelevant, so we often do not care much if the function is not defined at finitely many of the points on  $[a, b]$ . Similarly, we do not care whether the function is defined at the endpoints  $a$  and  $b$ . As far as integration is concerned, we shall not make very fine distinctions between the open, closed, left-open right-closed, and right-open left-closed intervals.

We now proceed to make sense of  $\int_a^b f(x) dx$ . In a later lecture, we will extend the meaning so that we can interpret  $\int_a^b f(x) dx$  for  $a = b$  and  $a > b$  as well.

## 2. PARTITIONS: TECHNICAL DETAILS BEGIN

**2.1. Partitions, upper sums, and lower sums.** Consider a closed interval  $[a, b]$ . By a *partition* of  $[a, b]$  we mean a sequence of points  $x_0 < x_1 < \dots < x_n$  with  $a = x_0$  and  $b = x_n$ . The nontrivial cases of partitions are when  $n \geq 2$ . We use the term *partition* because given the  $x_i$ , we can divide  $[a, b]$  into the *parts*  $[x_0, x_1]$ ,  $[x_1, x_2]$ , and so on, right till  $[x_{n-1}, x_n]$ . The union of these parts is  $[a, b]$ . Moreover, two adjacent parts intersect at a single point, and two non-adjacent parts do not intersect. *For our purposes, single points are too small to matter, as discussed above.* So, for our purposes, this is a partition into (almost) disjoint pieces.

The idea behind using partitions is to break up the behavior of the function into smaller intervals, wherein the variation in the value of the function within each interval is smaller than the overall variation in the

value. Thus, if we choose a partition with small enough parts, and find reasonable approximations for the integral on each part, adding those approximations up should give a reasonable approximation of the overall area.

**2.2. Upper bounds and lower bounds.** For the notion of integral to be reasonable, it should be true that if  $f(x) \leq g(x)$  for all  $x \in [a, b]$ , then the integral of  $f$  is less than or equal to the integral of  $g$ . Verbally, if the function gets bigger everywhere on the interval, its total value should also get bigger. Thus, we can try determining upper and lower bounds on the integral of  $f$  by finding functions slightly smaller and slightly larger than  $f$  that we know how to integrate. The integral of  $f$  is bounded between those two integrals.

Now, the only kinds of functions that we have already decided how to integrate are the piecewise constant functions, so we need to find good piecewise constant functions. We do this using the partition.

Suppose  $P = \{x_0, x_1, \dots, x_n\}$  is a partition of the interval  $[a, b]$ . Define piecewise constant functions  $f_l$  and  $f_u$  as follows: on each interval  $(x_{i-1}, x_i)$ ,  $f_l$  is constant at the minimum (more precisely infimum) of  $f$  over the interval  $[x_{i-1}, x_i]$  and  $f_u$  is constant at the maximum (more precisely supremum) of  $f$  over the interval  $[x_{i-1}, x_i]$ . So, both  $f_l$  and  $f_u$  are piecewise constant functions (define them whatever way you want at the points  $x_i$  – as mentioned earlier, the values at individual points do not matter). Note that for continuous functions, the extreme-value theorem guarantees that the function attains its maximum and minimum over any closed interval, so we do not need to make fine distinctions between infimum and minimum or between supremum and maximum.

The integral of  $f_l$  is given by the summation, for  $1 \leq i \leq n$ , of the product of  $(x_i - x_{i-1})$  and the minimum value of  $f$  over  $[x_{i-1}, x_i]$ . This value is known as the *lower sum* of  $f$  for the partition  $P$ , and it is denoted  $L_f(P)$ . In symbols:

$$L_f(P) = \sum_{i=1}^n (x_i - x_{i-1}) * (\text{minimum value for } f \text{ over } [x_{i-1}, x_i])$$

The integral of  $f_u$  is given by the summation, for  $1 \leq i \leq n$ , of the product of  $(x_i - x_{i-1})$  and the maximum value of  $f$  over  $[x_{i-1}, x_i]$ . This value is known as the *upper sum* of  $f$  for the partition  $P$ , and it is denoted  $U_f(P)$ . For obvious reasons,  $L_f(P) \leq U_f(P)$ .

$$U_f(P) = \sum_{i=1}^n (x_i - x_{i-1}) * (\text{maximum value for } f \text{ over } [x_{i-1}, x_i])$$

**2.3. Finer partitions and integral as limiting value.** Given two partitions  $P_1$  and  $P_2$ , we say that  $P_2$  is *finer* than  $P_1$  if the points of  $P_1$  form a subset of the points of  $P_2$ . In other words,  $P_2$  has all the points of  $P_1$  and perhaps more. This means that each interval for the partition  $P_2$  is contained in an interval for the partition  $P_1$ . The finer the partition, the *better* in some sense, since the smaller the interval, the more legitimate the process of approximating by a constant function on that interval.

If  $P_2$  is finer than  $P_1$ , then it turns out that  $U_f(P_2) \leq U_f(P_1)$  and  $L_f(P_2) \geq L_f(P_1)$ . In other words, the upper sums get smaller (though not necessarily strictly smaller) and the lower sums get bigger (though not necessarily strictly bigger) as the partition becomes finer. This can be seen formally as well. The idea is that when one part is subdivided further, the maximum over the entire part is greater than or equal to the maximum over each subpart. Thus, after subdivision, we are multiplying potentially smaller numbers with the same interval lengths, and the overall upper sum thus either remains the same or becomes smaller.

What we hope is that, as the partition gets finer and finer, the lower sums converge upward and the upper sums converge downward to a particular value, and we can then declare that value to be the integral of the function. Formally, for a function  $f$  on  $[a, b]$  and partitions  $P$  of  $[a, b]$ :

If  $\lim_{\|P\| \rightarrow 0} U_f(P) = \lim_{\|P\| \rightarrow 0} L_f(P)$ , then this common value is termed the *integral* of  $f$  over the interval  $[a, b]$ , and is denoted  $\int_a^b f(x) dx$ .

What precisely does  $\lim_{\|P\| \rightarrow 0}$  mean? For  $P = \{x_0, x_1, x_2, \dots, x_n\}$ , we define  $\|P\| = \max_{1 \leq i \leq n} (x_i - x_{i-1})$ . In other words, it is the maximum of the lengths of the intervals in the partition  $P$ . Sending this limit to zero means that we are considering partitions that get smaller and smaller in the sense that their largest part's size approaches zero.

This is a kind of limiting process that you have not seen in the past. So far, you have only seen limits as one real-valued variable approaches one constant value. But a partition  $P$  is not a real number; it is a more complex collection of information. In order to make sense of limiting to zero, we invent a way of measuring the size of the partition (by looking at the maximum of the sizes of the parts) and then apply the constraint that this size needs to go to zero. The limit is being taken over the space of all partitions, which is not a line.

To make matter simpler, we can restrict attention to what are called *regular partitions*. A regular partition is a partition where all the parts have equal size. For an interval  $[a, b]$ , there is a unique regular partition with  $n$  parts, and in that, each part has size  $(b - a)/n$ . Restricted to regular partitions, the above just means that we are sending  $n$  to  $\infty$ .

**2.4. Integrating the identity function.** We illustrate the technique of using partitions to integrate the function  $f(x) = x$  over the interval  $[0, 1]$ .

We begin by looking at the trivial partition  $P_1 = \{0, 1\}$ . This basically means that we do not subdivide the interval into smaller pieces. For the function  $f(x) = x$ , the maximum value over the interval  $[0, 1]$  is 1 and the minimum value is 0. Thus,  $U_f(P_1) = 1(1 - 0) = 1$  and  $L_f(P_1) = 0(1 - 0) = 0$ . Thus, even without breaking the interval up further, we already know that the integral is somewhere between 0 and 1.

Next, consider  $P_2 = \{0, 1/2, 1\}$ . In this case, we have the two intervals  $[0, 1/2]$  and  $[1/2, 1]$ . On the first interval, the minimum value is 0 and the maximum value is  $1/2$ . And on the second interval, the minimum value is  $1/2$  and the maximum value is 1.

We thus get  $L_f(P_2) = (0)(1/2 - 0) + (1/2)(1 - 1/2) = 1/4$  and  $U_f(P_2) = (1/2)(1/2 - 0) + (1)(1 - 1/2) = 3/4$ . Thus, the integral is somewhere between  $1/4$  and  $3/4$ . We have thus narrowed the value of the integral to within a smaller interval.

Let us now consider a regular partition into  $n$  pieces, i.e., the partition  $P_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$ . In each interval  $[(i - 1)/n, i/n]$ , the maximum is  $i/n$  and the minimum is  $(i - 1)/n$ . Thus, we get:

$$L_f(P_n) = \sum_{i=1}^n \frac{i-1}{n} \left( \frac{i}{n} - \frac{i-1}{n} \right)$$

That summation is given by:

$$L_f(P_n) = \frac{1}{n^2} \sum_{i=1}^n (i-1)$$

The summation inside is the sum of the numbers  $0, 1, \dots, n - 1$ . The summation (which we proved by induction in the first quarter) is  $n(n - 1)/2$ , and we thus get:

$$L_f(P_n) = \frac{n-1}{2n} = \frac{1}{2} - \frac{1}{2n}$$

Similarly, we can calculate that:

$$U_f(P_n) = \frac{n+1}{2n} = \frac{1}{2} + \frac{1}{2n}$$

As  $n \rightarrow \infty$ , the fraction  $1/2n$  tends to zero, and we obtain that both  $L_f(P_n)$  and  $U_f(P_n)$  tend to  $1/2$  (with  $L_f(P_n)$  approaching from the left and  $U_f(P_n)$  approaching from the right). Thus, the integral of the identity function on  $[0, 1]$  equals  $1/2$ .

More generally, it turns out that the integral  $\int_a^b f(x)dx = (b^2 - a^2)/2$ . In a later lecture, we will look at general ways of finding the integral.

**2.5. Brief note: integral of piecewise constant functions.** As mentioned earlier, the integral of a piecewise constant function is given by the sum of the signed areas of the rectangles corresponding to each interval where it is constant. For instance, consider the function  $f$  on  $[0, 3]$  such that  $f(x) = 5$  on  $[0, 1)$  and  $f(x) = -7$  on  $[1, 3]$ . Then, the integral of  $f$  is given by:

$$5 * (1 - 0) + (-7) * (3 - 1) = -9$$

For a piecewise constant function, it turns out that we can (almost) choose a partition such that both the upper and lower sum for the partition equal the value of the integral.

Here's the rough idea: we can choose a partition such that the function is constant on each part. Thus, on each of those parts, the maximum and minimum of the function are equal to the constant value, hence the contributions to both the upper sum and the lower sum are equal.

The problem is that because the partitions use closed intervals, we run into issues at places where the function changes value. If we used open intervals instead of closed intervals, this problem would not arise and we would be fine.

### 3. ADDITIONAL NOTES ON PARTITIONS, NORMS, AND REGULAR PARTITIONS

**3.1. Norm of a partition and its significance.** Recall that we defined the norm  $\|P\|$  of a partition  $P = \{x_0, x_1, x_2, \dots, x_n\}$  as the largest of the part sizes, i.e.,  $\max_{1 \leq i \leq n} |x_i - x_{i-1}|$ . What is the significance of this norm?

The norm is not important per se, but its main significance is as follows: we want a norm  $P$  with the property that  $\|P\| \rightarrow 0$  forces *all* the parts to become small. Thus, if instead of the largest of the part sizes, we took the *average* part size or the *smallest* part size, then that norm could be made arbitrarily small while keeping some of the pieces in the partition very large.

This shifts the question: why do we want a partition where *all* the parts become small? The intuition is that the smaller the part, the less the variation (hopefully) in the value of the function within each part. If we do not shrink everywhere, it may so happen that the portion of the interval where the size is large is precisely the portion where there is huge variation in the function value, so that the upper and lower sum estimates are grossly off.

**3.2. For wild functions: what can happen with lower and upper sums?** When a function is continuous on a closed interval, the integral always exists and is finite. The same holds for piecewise continuous functions. What happens for a function that is not continuous or even piecewise continuous? What if the function is discontinuous on a dense subset of the reals? In these cases, the integral does not exist.

If the function is bounded, the  $\lim_{\|P\| \rightarrow 0} L_f(P)$  and  $\lim_{\|P\| \rightarrow 0} U_f(P)$  both exist but they are not equal, i.e., the first limit is strictly smaller than the second limit. An example is for  $f$  the Dirichlet function that takes the value 1 at rationals and the value 0 at irrationals. Here, on any interval, the maximum value is 1 and the minimum value is 0, hence over any interval  $[a, b]$  and any partition  $P$  of the interval,  $U_f(P) = b - a$  and  $L_f(P) = 0$ .

**3.3. Finer partitions, norm, and incomparability.** We know that if  $P_2$  is a finer partition than  $P_1$ , then (i)  $\|P_2\| \leq \|P_1\|$ , (ii)  $L_f(P_2) \geq L_f(P_1)$ , and (iii)  $U_f(P_2) \leq U_f(P_1)$ . In other words, the norm becomes smaller, the upper sum becomes smaller, and the lower sum becomes larger. However, none of (i), (ii), or (iii) individually imply that  $P_2$  is finer than  $P_1$ . It is very much possible that two partitions are incomparable (i.e., neither is finer than the other). The mathematical jargon for this is that the relation of being finer is a partial order and not a total order on the collection of all partitions – many pairs of partitions are incomparable. In contrast, any numerical value associated with a partition has a real value and these numerical values can be totally ordered.

Despite this, any two partitions always have a common refinement that is finer than both of them.

**3.4. Regular partitions.** As mentioned, a regular partition of  $[a, b]$  into  $n$  parts is a partition with  $n$  intervals each of size  $(b - a)/n$ . The norm of a regular partition is  $(b - a)/n$ . Taking the limit as  $n \rightarrow \infty$ , we get that the norm goes to 0. The sequence of regular partitions with  $n$  parts, with  $n$  varying over the natural numbers, thus is a natural sequence to use when computing integrals using upper and lower sums.

The larger the  $n$ , the smaller the norm of a regular partition. However, it is *not* true that the regular partition for any larger  $n$  is *finer*. For instance, the partitions of  $[0, 1]$  for  $n = 2$  and  $n = 3$  are incomparable. A partition into  $n$  parts is finer than a partition into  $m$  parts if  $m$  divides  $n$ , i.e.,  $n$  is a multiple of  $m$ .