Chapter 5

Area

In chapter 2 we calculated the area of the set

$$\{(x, y) \in \mathbf{R}^2 \colon 0 \le x \le a \text{ and } 0 \le y \le x^2\}$$

where $a \ge 0$, and of the set

$$\{(x, y) \in \mathbf{R}^2 \colon 1 \le x \le b \text{ and } 0 \le y \le x^{-2}\}$$

where $b > 1$.

The technique that was used for making these calculations can be used to find the areas of many other subsets of $\mathbf{R}^2$. The general procedure we will use for finding the area of a set S will be to find two sequences of polygons $\{I_n\}$ and $\{O_n\}$ such that

$$I_n \subset S \subset O_n \text{ for all } n \in \mathbf{Z}^+.$$

We will then have

$$\text{area}(I_n) \le \text{area}(S) \le \text{area}(O_n) \text{ for all } n \in \mathbf{Z}^+. \tag{5.1}$$

We will construct the polygons I_n and O_n so that $\text{area}(O_n) - \text{area}(I_n)$ is arbitrarily small when n large enough, and we will see that then there is a unique number A such that

$$\text{area}(I_n) \le A \le \text{area}(O_n) \text{ for all } n \in \mathbf{Z}^+. \tag{5.2}$$

We will take A to be the area of S.

5.1 Basic Assumptions about Area

5.3 Definition (Bounded Sets.) A subset S of $\mathbf{R}^2$ is *bounded* if $S \subset B(a, b\colon c, d)$ for some box $B(a, b\colon c, d)$. A subset of $\mathbf{R}^2$ that is not bounded is said to be *unbounded*.

It is clear that every subset of a bounded set is bounded. It is not difficult to show that if B is a bounded set then $\mathbf{a} + B$ is bounded for every $\mathbf{a} \in \mathbf{R}^2$, and $S(B)$ is bounded for every symmetry of the square, S.

5.4 Example. The set

$$\{(n, \frac{1}{n}) : n \in \mathbf{Z}^+\}$$

is an unbounded subset of $\mathbf{R}^2$. I cannot draw a picture of an unbounded set, because the sheet of paper on which I make my drawing will represent a box containing any picture I draw.

5.5 Definition (Bounded Function.) Let S be a set. A function $f : S \to \mathbf{R}$ is called a *bounded function* if there is a number M such that $|f(x)| \leq M$ for all $x \in S$. It is clear that if f is a bounded function on an interval $[a, b]$, then graph(f) is a bounded subset of $\mathbf{R}^2$, since graph$(f) \subset B(a, b : -M, M)$. If f is bounded on S, then any number M satisfying

$$|f(x)| \leq M \text{ for all } x \in S$$

is called a *bound for f* on S.

We are now ready to state our official assumptions about area. At this point you should officially forget everything you know about area. Unofficially, however, you remember everything you know so that you can evaluate whether the theorems we prove are reasonable. Our aim is not simply to calculate areas, but to see how our calculations are justified by our assumptions.

We will assume that there is a function α from the set of bounded subsets of $\mathbf{R}^2$ to the real numbers that satisfies the conditions of positivity, additivity, normalization, translation invariance and symmetry invariance described below. Any function α that satisfies these conditions will be called an *area function*.

5.6 Assumption (Positivity of area.)

$$\alpha(S) \geq 0 \text{ for every bounded subset } S \text{ of } \mathbf{R}^2.$$

5.7 Definition (Disjoint sets.) We say two sets S, T are *disjoint* if and only if $S \cap T = \emptyset$.

5.8 Assumption (Additivity of area.) If S, T are disjoint bounded subsets of $\mathbf{R}^2$, then

$$\alpha(S \cup T) = \alpha(S) + \alpha(T).$$

5.9 Assumption (Normalization property of area.) For every box $B(a, b\colon c, d)$ we have

$$\alpha\Big(B(a, b\colon c, d)\Big) = (b - a)(d - c),$$

i.e., the area of a box is the product of the length and the width of the box.

5.10 Assumption (Translation invariance of area.) Let S be a bounded set in $\mathbf{R}^2$, and let $\mathbf{a} \in \mathbf{R}^2$, then

$$\alpha(S) = \alpha(\mathbf{a} + S).$$

5.11 Assumption (Invariance under symmetry.) Let S be a bounded subset of $\mathbf{R}^2$. Then if F is any symmetry of the square

$$\alpha(F(S)) = \alpha(S).$$

(See definition 4.12 for the definition of symmetry of the square.)

Remark: I would like to replace the assumptions 5.10 and 5.11 by the single stronger assumption:

> If A and B are bounded subsets of $\mathbf{R}^2$, and A is congruent to B, then $\alpha(A) = \alpha(B)$.

However the problem of defining what *congruent* means is rather complicated, and I do not want to consider it at this point.

5.12 Entertainment (Congruence problem.) Formulate a definition of what it means for two subsets of $\mathbf{R}^2$ to be congruent.

5.13 Example. Let

$$\begin{aligned} S &= B(0,1:0,1) \cap \{(x,y) \in \mathbf{R}^2 : x \in \mathbf{Q}\} \\ T &= B(0,1:0,1) \cap \{(x,y) \in \mathbf{R}^2 : x \notin \mathbf{Q}\}. \end{aligned}$$

I do not know how to make any reasonable drawing of S or T. Any picture I draw of S would look just like a picture of T, even though the two sets are disjoint. By additivity and the normalization property for area

$$\alpha(S) + \alpha(T) = \alpha(S \cup T) = \alpha(B(0,1:0,1)) = 1.$$

Since areas are non-negative, it follows that

$$0 \leq \alpha(S) \leq 1 \text{ and } 0 \leq \alpha(T) \leq 1.$$

The problem of calculating $\alpha(S)$ exactly cannot be answered on the basis of the assumptions we have made.

Remarks: The assumptions we have just made are supposed to be intuitively plausible. When we choose to make a particular set of assumptions, we hope that the assumptions are consistent, i.e., that no contradictions follow from them. If we were to add a new assumption:

The area of a circle with radius 1 is 3.14159,

then we would have an inconsistent set of assumptions, because it follows from the assumptions we have already made that the area of a circle of radius 1 is greater than 3.141592.

In 1923 Stefan Banach(1892–1945) [5] showed that area functions exist, i.e., that the assumptions we have made about area are consistent. Unfortunately Banach showed that there is more than one area function, and different area functions assign different values to the set S described in the previous example.

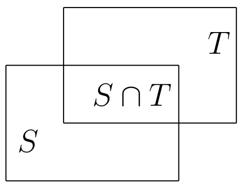
A remarkable result of Felix Hausdorff(1868–1942) [24, pp469–472] shows that the analogous assumptions for volume in three dimensional space (if we include the assumption that any two congruent sets in 3 dimensional space $\mathbf{R}^3$ have the same volume) are inconsistent. If one wants to discuss volume in $\mathbf{R}^3$ then one cannot consider volumes of arbitrary sets. One must considerably restrict the class of sets that have volumes. A discussion of Hausdorff's result can be found in [20].

5.2 Further Assumptions About Area

In this section we will introduce some more assumptions about area. The assumptions in this section can actually be proved on the basis of the basic assumptions we have already made, and in fact the proofs are easy (the proofs are given in appendix B). The reason I have made assumptions out of them is that they are as intuitively plausible as the assumptions I have already made, and I do not have time to do everything I want to do. I am omitting the proofs with regret because I agree with Aristotle that

> It is manifest that it is far better to make the principles finite in number. Nay, they should be the fewest possible provided they enable all the same results to be proved. This is what mathematicians insist upon; for they take as principles things finite either in kind or in number.[25, page 178]

5.14 Assumption (Addition rule for area.)



For any bounded sets S and T in $\mathbf{R}^2$

$$\alpha(S \cup T) = \alpha(S) + \alpha(T) - \alpha(S \cap T). \tag{5.15}$$

and consequently

$$\alpha(S \cup T) \leq \alpha(S) + \alpha(T).$$

5.16 Assumption (Subadditivity of area.) Let $n \in \mathbf{Z}_{\geq 1}$, and let A_1, A_2, $\cdots$, A_n be bounded sets in $\mathbf{R}^2$. Then

$$\alpha(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} \alpha(A_i). \tag{5.17}$$

5.18 Assumption (Monotonicity of area.) Let S, T be bounded sets such that $S \subset T$. Then $\alpha(S) \leq \alpha(T)$.

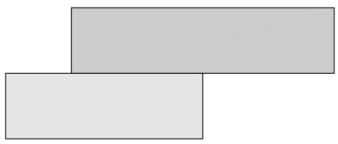
5.19 Definition (Zero-area set.) We will call a set with zero area a *zero-area set.*

From the normalization property it follows that every horizontal or vertical segment has area equal to 0. Thus every horizontal or vertical segment is a zero-area set.

5.20 Corollary (to assumption 5.18.) [1] *Every subset of a zero-area set is a zero-area set. In particular the empty set is a zero-area set.*

5.21 Corollary (to assumption 5.16.) *The union of a finite number of zero-area sets is a zero-area set.*

5.22 Definition (Almost disjoint.) We will say that two bounded subsets S, T of $\mathbf{R}^2$ are *almost disjoint* if $S \cap T$ is a zero-area set.



Almost disjoint sets

5.23 Examples. If a, b, c are real numbers with $a < b < c$, then since

$$B(a, b\colon p, q) \cap B(b, c\colon s, t) \subset B(b, b\colon p, q),$$

the boxes $B(a, b\colon p, q)$ and $B(b, c\colon s, t)$ are almost disjoint.

Any zero-area set is almost disjoint from every set – including itself.

5.24 Assumption (Additivity for almost disjoint sets.) Let $\{R_1, \cdots, R_n\}$ be a finite set of bounded sets such that R_i and R_j are almost disjoint whenever $i \neq j$. Then

$$\alpha(\bigcup_{i=1}^{n} R_i) = \sum_{i=1}^{n} \alpha(R_i). \tag{5.25}$$

[1] Usually a corollary is attached to a *theorem* and not to an *assumption.* A corollary is a statement that follows immediately from a theorem without a proof. By etymology, it is a "small gift".

5.26 Notation (Area functions α or area) Any real valued function α, whose domain is the family of bounded subsets of $\mathbf{R}^2$, and which satisfies all of the assumptions listed in sections 5.1 and 5.2 will be called an *area function*. In these notes I will use the names "α" and "area" to denote area function. Thus

$$\alpha(B(a,b:c,d)) = \text{area}(B(a,b:c,d)) = (b-a)(d-c).$$

5.3 Monotonic Functions

5.27 Definition (Partition.) Let a, b be real numbers with $a \leq b$. A *partition* P of the interval $[a,b]$ is a finite sequence of points

$$P = \{x_0, x_1, \cdots, x_n\}$$

with $a = x_0 \leq x_1 \leq x_2 \cdots \leq x_n = b$. The intervals $[x_0, x_1], [x_1, x_2], \cdots,$ $[x_{n-1}, x_n]$ are called the *subintervals of the partition* P, and $[x_{j-1}, x_j]$ is the j^{th} *subinterval of* P for $1 \leq j \leq n$. The largest of the numbers $x_j - x_{j-1}$ is called the *mesh of the partition* P, and is denoted by $\mu(P)$. The partition

$$\{a,\ a + \frac{(b-a)}{n},\ a + \frac{2(b-a)}{n}, \cdots, a + \frac{n(b-a)}{n} = b\}$$

is called the *regular partition of* $[a,b]$ *into* n *equal subintervals.*

5.28 Example. Let

$$P = \{0, \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1\}$$

Then P is a partition of $[0,1]$ into 5 subintervals and $\mu(P) = 1 - \frac{1}{2} = \frac{1}{2}$.

The regular partition of $[1,2]$ into 5 subintervals is $\{1, \frac{6}{5}, \frac{7}{5}, \frac{8}{5}, \frac{9}{5}, 2\}$.

If Q_n is the regular partition of $[a,b]$ into n equal subintervals, then $\mu(Q_n) = \frac{b-a}{n}$.

5.29 Exercise. Find a partition P of $[0,1]$ into five subintervals, such that $\mu(P) = \frac{4}{5}$, or explain why no such partition exists.

5.30 Exercise. Find a partition Q of $[0,1]$ into five subintervals, such that $\mu(Q) = \frac{1}{10}$, or explain why no such partition exists.

5.31 Definition (Monotonic function.) Let J be an interval in $\mathbf{R}$, and let $f\colon J \to \mathbf{R}$ be a function. We say that f is *increasing on* J if

$$\text{for all } x, y \text{ in } J\Big[(x \leq y) \implies (f(x) \leq f(y))\Big] \tag{5.32}$$

and we say that f is *decreasing on* J if

$$\text{for all } x, y \text{ in } J\Big[(x \leq y) \implies (f(x) \geq f(y))\Big].$$

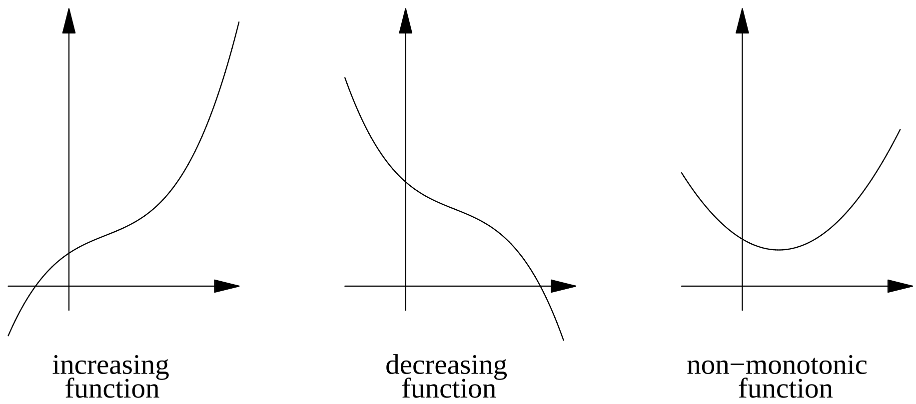
We say that f is *strictly increasing on* J if

$$\text{for all } x, y \text{ in } J\Big[(x < y) \implies (f(x) < f(y))\Big]$$

and we say that f is *strictly decreasing on* J if

$$\text{for all } x, y \text{ in } J\Big[(x < y) \implies (f(x) > f(y))\Big].$$

We say that f is *monotonic on* J if f is either increasing on J or decreasing on J, and we say that f is *strictly monotonic* on J if f is either strictly increasing or strictly decreasing on J.



A constant function on J is both increasing and decreasing on J.

5.33 Notation ($S_a^b f$) Let f be a function from the interval $[a, b]$ to the non-negative numbers. We will write

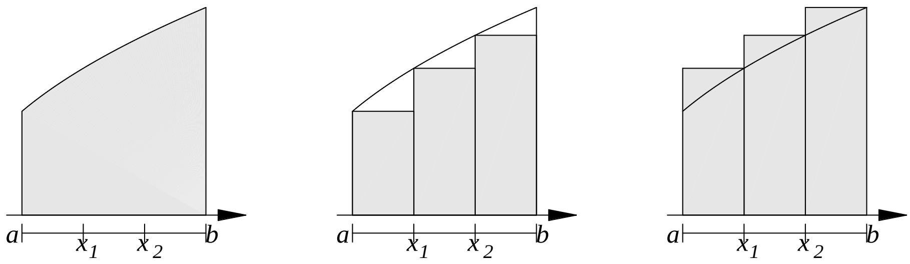
$$S_a^b f = \{(x, y) \in \mathbf{R}^2\colon a \leq x \leq b \text{ and } 0 \leq y \leq f(x)\},$$

i.e., $S_a^b f$ is the set of points under the graph of f over the interval $[a, b]$.

Let f be an increasing function from the interval $[a, b]$ to the non-negative numbers. Let $P = \{x_0, \cdots, x_n\}$ be a partition of $[a, b]$ and let

$$\begin{aligned} I_a^b(f, P) &= \bigcup_{i=1}^{n} B(x_{i-1}, x_i \colon 0, f(x_{i-1})) \\ O_a^b(f, P) &= \bigcup_{i=1}^{n} B(x_{i-1}, x_i \colon 0, f(x_i)). \end{aligned}$$

Then



$S_a^b(f)$ $I_a^b(f, P)$ $O_a^b(f, P)$

$$I_a^b(f, P) \subset S_a^b f \subset O_a^b(f, P). \tag{5.34}$$

To see this, observe that since f is increasing

$$x_{i-1} \le x \le x_i \implies f(x_{i-1}) \le f(x) \le f(x_i),$$

so

$$\begin{aligned} (x, y) \in I_a^b(f, P) &\implies (x, y) \in B(x_{i-1}, x_i : 0, f(x_{i-1})) \text{ for some } i \\ &\implies x_{i-1} \le x \le x_i \text{ and } 0 \le y \le f(x_{i-1}) \le f(x) \text{ for some } i \\ &\implies a \le x \le b \text{ and } 0 \le y \le f(x) \implies (x, y) \in S_a^b f. \end{aligned}$$

and also

$$\begin{aligned} (x, y) \in S_a^b f &\implies x_{i-1} \le x \le x_i \text{ and } 0 \le y \le f(x) \le f(x_i) \text{ for some } i \\ &\implies (x, y) \in B\Big(x_{i-1}, x_i \colon 0, f(x_i)\Big) \text{ for some } i \\ &\implies (x, y) \in O_a^b(f, P). \end{aligned}$$

By equation (5.34) and monotonicity of area, we have

$$\alpha\Big(I_a^b(f,P)\Big) \leq \alpha(S_a^b f) \leq \alpha\Big(O_a^b(f,P)\Big). \tag{5.35}$$

Now

$$\begin{aligned} &\alpha\Big(O_a^b(f,P)\Big) - \alpha\Big(I_a^b(f,P)\Big) \\ &= \sum_{i=1}^{n}(x_i - x_{i-1})f(x_i) - \sum_{i=1}^{n}(x_i - x_{i-1})f(x_{i-1}) \\ &= \sum_{i=1}^{n}(x_i - x_{i-1})\Big(f(x_i) - f(x_{i-1})\Big). \end{aligned} \tag{5.36}$$

Now let $\mu(P)$ be the mesh of P (cf. definition 5.27) so that

$$0 \leq x_i - x_{i-1} \leq \mu(P) \text{ for } 1 \leq i \leq n.$$

Since $f(x_i) - f(x_{i-1}) \geq 0$ for all i, we have

$$(x_i - x_{i-1})\Big(f(x_i) - f(x_{i-1})\Big) \leq \mu(P)\Big(f(x_i) - f(x_{i-1})\Big)$$

for all i, and hence

$$\begin{aligned} \sum_{i=1}^{n}(x_i - x_{i-1})\Big(f(x_i) - f(x_{i-1})\Big) &\leq \sum_{i=1}^{n}\mu(P)\Big(f(x_i) - f(x_{i-1})\Big) \\ &= \mu(P)\sum_{i=1}^{n}\Big(f(x_i) - f(x_{i-1})\Big). \end{aligned} \tag{5.37}$$

Now

$$\begin{aligned} \sum_{i=1}^{n}\Big(f(x_i) - f(x_{i-1})\Big) &= \Big(f(x_n) - f(x_{n-1})\Big) + \Big(f(x_{n-1}) - f(x_{n-2})\Big) \\ &\quad + \cdots + \Big(f(x_1) - f(x_0)\Big) \\ &= f(x_n) - f(x_0) = f(b) - f(a). \end{aligned}$$

so by equations (5.37) and (5.36), we have

$$\alpha\Big(O_a^b(f,P)\Big) - \alpha\Big(I_a^b(f,P)\Big) \leq \mu(P)\Big(f(b) - f(a)\Big).$$

Now suppose that A is any real number that satisfies

$$\alpha(I_a^b(f,P)) \leq A \leq \alpha(O_a^b(f,P)) \text{ for every partition } P \text{ of } [a,b].$$

We will show that $A = \alpha(S_a^b f)$. We have

$$-\alpha(O_a^b(f,P)) \leq -A \leq -\alpha(I_a^b(f,P)).$$

It follows from (5.35) that

$$\alpha(I_a^b(f,P)) - \alpha(O_a^b(f,P)) \leq \alpha(S_a^b f) - A \leq \alpha(O_a^b(f,P)) - \alpha(I_a^b(f,P)).$$

Thus

$$-\mu(P)(f(b) - f(a)) \leq \alpha(S_a^b f) - A \leq \mu(P)(f(b) - f(a)) \tag{5.38}$$

for every partition P of $[a,b]$. Since we can find partitions P with $\mu(P)$ smaller than any preassigned number, it follows that

$$A = \alpha(S_a^b f). \tag{5.39}$$

(After we have discussed the notion of *limit*, we will come back and reconsider how (5.39) follows from (5.38). For the present, I will just say that the implication is intuitively clear.) We have now proved the following theorem:

5.40 Theorem. *Let f be an increasing function from the interval $[a,b]$ to $\mathbf{R}_{\geq 0}$, and let $P = \{x_0, x_1, \cdots, x_n\}$ be a partition of $[a,b]$. Let*

$$S_a^b f = \{(x,y) \in \mathbf{R}^2 \colon a \leq x \leq b \text{ and } 0 \leq y \leq f(x)\},$$

$$\begin{aligned}
I_a^b(f,P) &= \bigcup_{i=1}^{n} B\Big(x_{i-1}, x_i \colon 0, f(x_{i-1})\Big), && (5.41)\\
O_a^b(f,P) &= \bigcup_{i=1}^{n} B\Big(x_{i-1}, x_i \colon 0, f(x_i)\Big), && (5.42)\\
A_a^b f &= \alpha(S_a^b f).
\end{aligned}$$

Then

$$\alpha\Big(I_a^b(f,P)\Big) \leq A_a^b f \leq \alpha\Big(O_a^b(f,P)\Big) \tag{5.43}$$

and

$$\alpha\Big(O_a^b(f,P)\Big) - \alpha\Big(I_a^b(f,P)\Big) \leq \mu(P)\Big(f(b) - f(a)\Big). \tag{5.44}$$

Also

$$\alpha\Big(I_a^b(f,P)\Big) = \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1}) \tag{5.45}$$

$$\alpha\Big(O_a^b(f,P)\Big) = \sum_{i=1}^{n} f(x_i)(x_i - x_{i-1}). \tag{5.46}$$

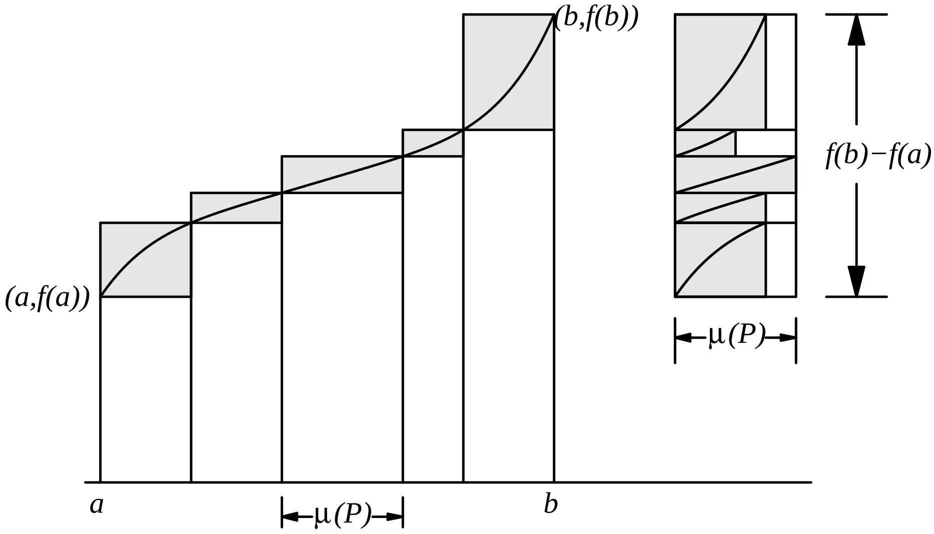
If A is any real number such that

$$\alpha(I_a^b(f,P)) \leq A \leq \alpha(O_a^b(f,P)) \text{ for every partition } P \text{ of } [a,b],$$

then

$$A = A_a^b(f).$$

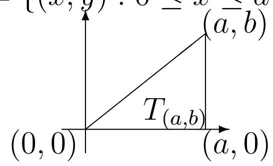
The following picture illustrates the previous theorem.



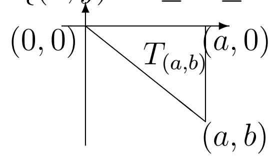
5.47 Exercise. A version of theorem 5.40 for decreasing functions is also valid. To get this version you should replace the word "increasing" by "decreasing" and change lines (5.41), (5.42), (5.44), (5.45) and (5.46). Write down the proper versions of the altered lines. As usual, use I to denote areas inside $S_a^b f$ and O to denote sets containing $S_a^b f$. Draw a picture corresponding to the above figure for a decreasing function.

5.48 Definition (Right triangle $T_{\mathbf{c}}$) Let a and b be non-zero real numbers, and let $\mathbf{c} = (a, b)$. We define the triangle $T_{\mathbf{c}} = T_{(a,b)}$ to be the set of points between the line segment $[\mathbf{0c}]$ and the x-axis. By example 4.8, we know that the equation of the line through $\mathbf{0}$ and $\mathbf{c}$ is $y = \frac{b}{a}x$. Hence we have:

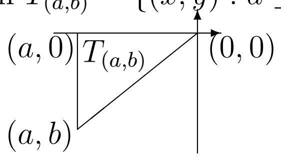
If $a > 0$ and $b > 0$, then $T_{(a,b)} = \{(x, y) : 0 \leq x \leq a \text{ and } 0 \leq y \leq \frac{b}{a}x\}$



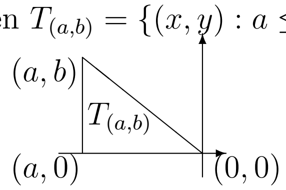
If $a > 0$ and $b < 0$, then $T_{(a,b)} = \{(x, y) : 0 \leq x \leq a \text{ and } \frac{b}{a}x \leq y \leq 0\}$



If $a < 0$ and $b < 0$, then $T_{(a,b)} = \{(x, y) : a \leq x \leq 0 \text{ and } \frac{b}{a}x \leq y \leq 0\}$



If $a < 0$ and $b > 0$, then $T_{(a,b)} = \{(x, y) : a \leq x \leq 0 \text{ and } 0 \leq y \leq \frac{b}{a}x\}$



5.49 Remark. We know from Euclidean geometry that

$$\alpha(T_{(a,b)}) = \frac{1}{2}|a||b|. \tag{5.50}$$

I would like to show that this relation follows from our assumptions about area. If H, V and R_π are the reflections and rotation defined in definition 4.9, then we can show without difficulty that for $a > 0$ and $b > 0$

$$\begin{aligned} T_{(-a,b)} &= H(T(a,b)), \\ T_{(a,-b)} &= V(T_{(a,b)}), \text{ and} \\ T_{(-a,-b)} &= R_\pi(T_{(a,b)}) \end{aligned}$$

so by invariance of area under symmetry,

$$\alpha(T_{(a,b)}) = \alpha(T_{(-a,b)}) = \alpha(T_{(a,-b)}) = \alpha(T_{(-a,-b)})$$

when a and b are positive. It follows that if we prove formula (5.50) when a and b are positive, then the formula holds in all cases. For example if a and b are positive, and we know that (5.50) holds when a and b are positive, we get

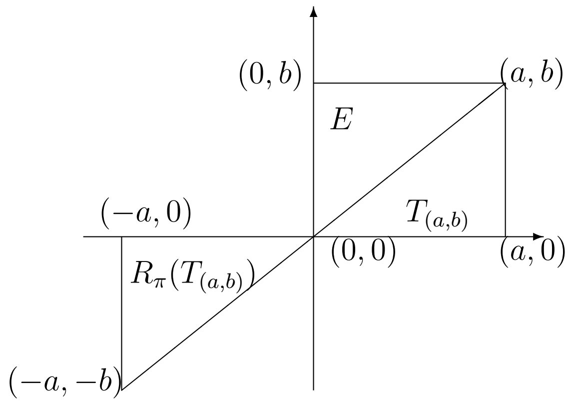
$$\alpha(T_{(-a,b)}) = \alpha(T_{(a,b)}) = \frac{1}{2}|a||b| = \frac{1}{2}|-a||b|,$$

and thus our formula holds when a is negative and b is positive.

5.51 Theorem. *Let a and b be non-zero real numbers, and let $T_{(a,b)}$ be the set defined in definition 5.48. Then*

$$\alpha(T_{(a,b)}) = \frac{1}{2}|a||b|.$$

Proof: By the previous remark, if is sufficient to prove the theorem for the case when a and b are positive. So suppose that a and b are positive.



Let $E = (a,b) + R_\pi(T_{(a,b)})$. It appears from the figure, and is straightforward to show, that

$$E = \{(x,y) : 0 \le x \le a \text{ and } \frac{b}{a}x \le y \le b\}.$$

By translation invariance of area,

$$\alpha(E) = \alpha(R_\pi(T_{(a,b)})) = \alpha(T_{(a,b)}).$$

We have

$$E \cup T_{(a,b)} = B(0, a : 0, b),$$

and

$$E \cap T_{(a,b)} = [\mathbf{0c}] \text{ where } \mathbf{c} = (a, b).$$

By the addition rule for area (assumption 5.14) we have

$$\begin{aligned} ab &= \alpha(B(0, a : 0, b)) \\ &= \alpha(E \cup T_{(a,b)}) \\ &= \alpha(E) + \alpha(T_{(a,b)}) - \alpha(E \cap T_{(a,b)}) \\ &= 2\alpha(T_{(a,b)}) - \alpha([\mathbf{0c}]), \end{aligned}$$

i.e.,

$$\alpha(T_{(a,b)}) = \frac{1}{2}ab + \frac{1}{2}\alpha([\mathbf{0c}]).$$

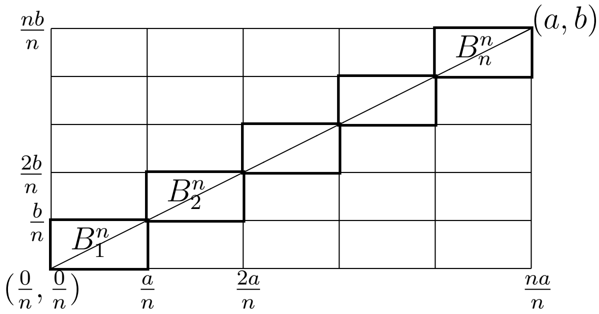
Thus our theorem will follow if we can show that the segment $[\mathbf{0c}]$ is a zero-area set. We will prove this as the next theorem.

5.52 Theorem. *Let* $\mathbf{c} = (a, b)$ *be a point in* $\mathbf{R}^2$*. Then*

$$\alpha([\mathbf{0c}]) = \mathbf{0}.$$

Proof: If $a = 0$ or $b = 0$, then $[\mathbf{0c}]$ is a box with width equal to zero, or height equal to zero, so the theorem holds in this case. Hence we only need to consider the case where a and b are non-zero. Since any segment $[\mathbf{0c}]$ can be rotated or reflected to a segment $[\mathbf{0q}]$ where $\mathbf{q}$ is in the first quadrant, we may further assume that a and b are both positive. Let n be a positive integer, and for $1 \leq j \leq n$ let

$$B_j^n = B\left(\frac{a(j-1)}{n}, \frac{aj}{n} : \frac{b(j-1)}{n}, \frac{bj}{n}\right).$$



Then

$$[\mathbf{0c}] \subset \bigcup_{j=1}^{n} B_j^n, \tag{5.53}$$

since

$$\begin{aligned}
\mathbf{x} \in [\mathbf{0c}] &\implies \mathbf{x} = (ta, tb) \text{ for some } t \in [0,1] \\
&\implies \mathbf{x} = (ta, tb) \text{ where } \frac{j-1}{n} \leq t \leq, \frac{j}{n} \text{ for some } j \text{ with } 1 \leq j \leq n \\
&\implies \mathbf{x} = (ta, tb) \text{ where } \frac{a(j-1)}{n} \leq ta \leq \frac{aj}{n} \\
&\qquad\qquad \text{and } \frac{b(j-1)}{n} \leq bt \leq \frac{bj}{n} \text{ for some } j \text{ with } 1 \leq j \leq n \\
&\implies \mathbf{x} \in B_j^n \text{ for some } j \text{ with } 1 \leq j \leq n.
\end{aligned}$$

For each j we have

$$\alpha(B_j^n) = \frac{a}{n} \cdot \frac{b}{n} = \frac{ab}{n^2}.$$

Also the sets B_j^n and B_k^n are almost disjoint whenever $1 \leq j, k \leq n$ and $j \neq k$. (If j and k differ by more than 1, then B_j^n and B_k^n are disjoint, and if j and k differ by 1, then $B_j^n \cap B_k^n$ consists of a single point.) By additivity for almost-disjoint sets (assumption 5.25), it follows that

$$\alpha\Big(\bigcup_{j=1}^{n} B_j^n\Big) = \sum_{j=1}^{n} \alpha(B_j^n) = \sum_{j=1}^{n} \frac{ab}{n^2} = \frac{nab}{n^2} = \frac{ab}{n}.$$

By (5.53) and monotonicity of area we have

$$\alpha([0\mathbf{c}]) \leq \alpha(\bigcup_{j=1}^{n} B_j^n) = \frac{ab}{n} \text{ for every positive integer } n. \tag{5.54}$$

In order to conclude from this that $\alpha([\mathbf{0c}]) = 0$ We now make use of the *Archimedean property* of real numbers (see (C.79) in Appendix C) which says that for any real number x there is a positive integer n with $n > x$. We know $\alpha([\mathbf{0c}]) \geq 0$, since all areas are non-negative. Suppose (in order to get a contradiction) that $\alpha([\mathbf{0c}])$ is positive. Then by the Archimedean property, there is a positive integer N such that $N > \frac{ab}{\alpha([\mathbf{0c}])}$. This implies that $\alpha([\mathbf{0c}]) > \frac{ab}{N}$, and this contradicts (5.54). Hence $\alpha([\mathbf{0c}])$ is not positive, and we conclude that $\alpha([\mathbf{0c}]) = 0$. ‖

Archimedes' statement of the Archimedean property differs from our statement. He assumes that

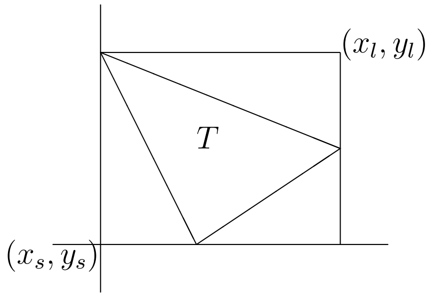
> Further, of unequal lines, unequal surfaces, and unequal solids, the greater exceeds the less by such a magnitude as, when added to itself, can be made to exceed any assigned magnitude among those which are comparable with [it and with] one another.[2, page 4]

5.55 Exercise. Let $\mathbf{a}$ and $\mathbf{b}$ be points in $\mathbf{R}^2$. Show that segment $[\mathbf{ab}]$ is a zero are set. (Use theorem 5.52. Do not reprove theorem 5.52).

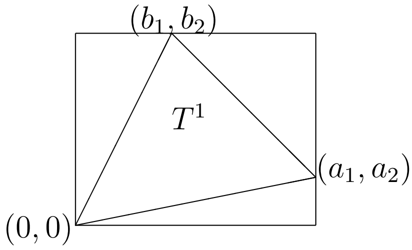
5.56 Entertainment (Area of a triangle) Let $\mathbf{x}_1 = (x_1, y_1)$, $\mathbf{x}_2 = (x_2, y_2)$ and $\mathbf{x}_3 = (x_3, y_3)$ be three points in $\mathbf{R}^2$, and let T be the triangle with vertices $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$. Let

$$\begin{aligned} x_s &= \text{smallest of } x_1,\ x_2 \text{ and } x_3 \\ x_l &= \text{largest of } x_1,\ x_2 \text{ and } x_3 \\ y_s &= \text{smallest of } y_1,\ y_2 \text{ and } y_3 \\ y_l &= \text{largest of } y_1,\ y_2 \text{ and } y_3. \end{aligned}$$

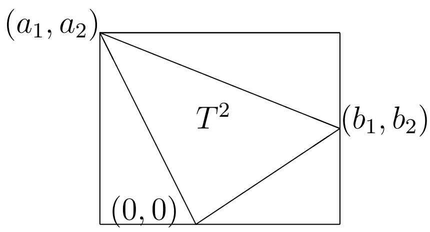
Then the box $B(x_s, x_l : y_s, y_l)$ is an almost-disjoint union of T and three triangles which are translates of triangles of the form $T_{\mathbf{c}}$. Since you know how to find the area of a box and of a triangle $T_{\mathbf{c}}$, you can find the area of T.



Using this remark show that for the triangles pictured below, $\alpha(T^1) = \frac{1}{2}(a_1b_2 - a_2b_1)$, and $\alpha(T^2) = \frac{1}{2}(a_2b_1 - a_1b_2)$.



$0 < b_1 < a_1$ and $0 < b_2 < a_2$



$a_1 < 0 < b_1$ and $0 < b_2 < a_2$

Then choose another triangle T^3 with vertices $\mathbf{0}, \mathbf{a}$ and $\mathbf{b}$, where the coordinates of the points are related in a way different from the ways shown for T^1 and T^2, and calculate the area of T^3. You should find that

$$\alpha(T^3) = \frac{1}{2}|a_1b_2 - a_2b_1|$$

in all cases. Notice that if some coordinate is zero, the formula agrees with theorem 5.51.

5.4 Logarithms.

5.57 Notation ($A_a^b f$, $A_a^b[f(t)]$.) Let f be a bounded function from the interval $[a, b]$ to $\mathbf{R}_{\geq 0}$. We will denote the area of $S_a^b f$ by $A_a^b f$. Thus

$$A_a^b f = \alpha\Big(\{(x, y) \in \mathbf{R}^2 \colon a \leq x \leq b \text{ and } 0 \leq y \leq f(x)\}\Big)$$

We will sometimes write $A_a^b[f(t)]$ instead of $A_a^b f$. Thus, for example

$$A_a^b[t^2] = \alpha\Big(\{(x,y)\in\mathbf{R}^2\colon a\leq x\leq b \text{ and } 0\leq y\leq x^2\}\Big)$$

We will also write $I_a^b([f(t)],P)$ and $O_a^b([f(t)],P)$ for $I_a^b(f,P)$ and $O_a^b(f,P)$ respectively.

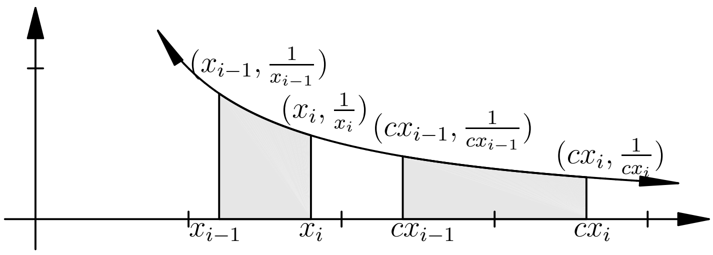
5.58 Lemma. [2] Let a, b, and c be real numbers such that $0 < a < b$ and $c > 0$. Then

$$A_{ac}^{bc}\Big[\frac{1}{t}\Big] = A_a^b\Big[\frac{1}{t}\Big].$$

Proof: Let $P = \{x_0, x_1, \cdots, x_n\}$ be a partition of $[a,b]$, and let

$$cP = \{cx_0, cx_1, \cdots, cx_n\}$$

be the partition of $[ca, cb]$ obtained by multiplying the points of P by c.



Then

$$\begin{aligned}
\alpha(I_{ac}^{bc}(\Big[\frac{1}{t}\Big], cP)) &= \sum_{i=1}^{n}\frac{1}{cx_i}(cx_i - cx_{i-1}) = \sum_{i=1}^{n}\frac{1}{cx_i}\cdot c(x_i - x_{i-1}) \\
&= \sum_{i=1}^{n}\frac{1}{x_i}(x_i - x_{i-1}) = \alpha(I_a^b(\Big[\frac{1}{t}\Big], P)) \qquad (5.59)
\end{aligned}$$

and

$$\begin{aligned}
\alpha(O_{ac}^{bc}(\Big[\frac{1}{t}\Big], cP)) &= \sum_{i=1}^{n}\frac{1}{cx_{i-1}}(cx_i - cx_{i-1}) = \sum_{i=1}^{n}\frac{1}{cx_{i-1}}\cdot c(x_i - x_{i-1}) \\
&= \sum_{i=1}^{n}\frac{1}{x_{i-1}}(x_i - x_{i-1}) = \alpha(O_a^b(\Big[\frac{1}{t}\Big], P)) \qquad (5.60)
\end{aligned}$$

[2] A lemma is a theorem which is proved in order to help prove some other theorem.

We know that

$$\alpha(I_{ac}^{bc}(\Big[\frac{1}{t}\Big], cP)) \leq A_{ac}^{bc}\Big[\frac{1}{t}\Big] \leq \alpha(O_{ac}^{bc}(\Big[\frac{1}{t}\Big], cP)).$$

Hence by (5.59) and (5.60) we have

$$\alpha(I_a^b(\Big[\frac{1}{t}\Big], P)) \leq A_{ac}^{bc}\Big[\frac{1}{t}\Big] \leq \alpha(O_a^b(\Big[\frac{1}{t}\Big], P))$$

for every partition P of $[a, b]$. It follows from this and the last statement of theorem 5.40 that

$$A_{ac}^{bc}\Big[\frac{1}{t}\Big] = A_a^b\Big[\frac{1}{t}\Big]. \parallel$$

5.61 Exercise. From lemma 5.58 we see that

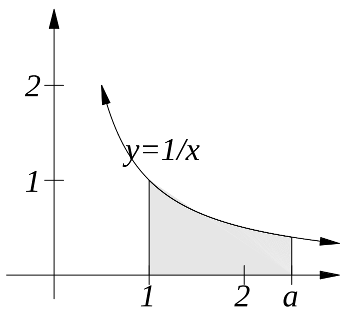
$$A_a^b\Big[\frac{1}{t}\Big] = A_{ac}^{bc}\Big[\frac{1}{t}\Big]$$

whenever $0 < a < b$, and $c > 0$. Use this result to show that for $a \geq 1$ and $b \geq 1$

$$A_1^{ab}\Big[\frac{1}{t}\Big] = A_1^a\Big[\frac{1}{t}\Big] + A_1^b\Big[\frac{1}{t}\Big]. \tag{5.62}$$

5.63 Definition ($L(x)$.) We will define a function $L\colon [1, \infty) \to \mathbf{R}$ by

$$L(a) = A_1^a\Big[\frac{1}{t}\Big] \text{ for all } a \in [1, \infty).$$



shaded area = $L(a)$

By exercise 5.61 we have

$$L(ab) = L(a) + L(b) \text{ for all } a \geq 1,\ b \geq 1. \tag{5.64}$$

In this section we will extend the domain of L to all of $\mathbf{R}^+$ in such a way that (5.64) holds for all $a, b \in \mathbf{R}^+$.

5.65 Theorem. *Let a, b, c be real numbers such that $a \leq b \leq c$, and let f be a bounded function from $[a, b]$ to $\mathbf{R}_{\geq 0}$. Then*

$$A_a^c f = A_a^b f + A_b^c f. \tag{5.66}$$

Proof: We want to show

$$\alpha(S_a^c f) = \alpha(S_a^b f) + \alpha(S_b^c f).$$

Since $S_a^c f = S_a^b f \cup S_b^c f$ and the sets $S_a^b f$ and $S_b^c f$ are almost disjoint, this conclusion follows from our assumption about additivity of area for almost disjoint sets.

I now want to extend the definition of $A_a^b f$ to cases where b may be less than a. I want equation (5.66) to continue to hold in all cases. If $c = a$ in (5.66), we get

$$0 = A_a^a f = A_a^b f + A_b^a f$$

i.e.,

$$A_b^a f = -A_a^b f.$$

Thus we make the following definition:

5.67 Definition. Let a, b be real numbers with $a \leq b$ and let f be a bounded function from $[a, b]$ to $\mathbf{R}_{\geq 0}$. Then we define

$$A_b^a f = -A_a^b f \text{ or } A_b^a[f(t)] = -A_a^b[f(t)].$$

5.68 Theorem. *Let a, b, c be real numbers and let f be a bounded non-negative real valued function whose domain contains an interval containing a, b, and c. Then*

$$A_a^c f = A_a^b f + A_b^c f.$$

Proof: We need to consider the six possible orderings for a, b and c. If $a \le b \le c$ we already know the result. Suppose $b \le c \le a$. Then $A_b^a f = A_b^c f + A_c^a f$ and hence $-A_a^b f = A_b^c f - A_a^c f$, i.e., $A_a^c f = A_a^b f + A_b^c f$. The remaining four cases are left as an exercise.

5.69 Exercise. Prove the remaining four cases of theorem 5.68.

5.70 Definition (Logarithm.) If a is any positive number, we define the *logarithm of* a by

$$\ln(a) = L(a) = A_1^a\Big[\frac{1}{t}\Big].$$

5.71 Theorem (Properties of Logarithms.) *For all $a, b \in \mathbf{R}^+$ and all $r \in \mathbf{Q}$ we have*

$$\begin{aligned} L(ab) &= L(a) + L(b) \\ L\Big(\frac{a}{b}\Big) &= L(a) - L(b) \\ L(a^{-1}) &= -L(a) \\ L(a^r) &= rL(a) && (5.72) \\ L(1) &= 0. && (5.73) \end{aligned}$$

Proof: Let $a, b, c \in \mathbf{R}^+$. From lemma 5.58 we know that if $a \le c$ then

$$A_a^c\Big[\frac{1}{t}\Big] = A_{ba}^{bc}\Big[\frac{1}{t}\Big] \tag{5.74}$$

If $c < a$ we get

$$A_a^c\Big[\frac{1}{t}\Big] = -A_c^a\Big[\frac{1}{t}\Big] = -A_{bc}^{ba}\Big[\frac{1}{t}\Big] = A_{ba}^{bc}\Big[\frac{1}{t}\Big]$$

so equation (5.74) holds in all cases. Let a, b be arbitrary elements in $\mathbf{R}^+$. Then

$$\begin{aligned} L(ab) &= A_1^{ab}\Big[\frac{1}{t}\Big] = A_1^a\Big[\frac{1}{t}\Big] + A_a^{ab}\Big[\frac{1}{t}\Big] \\ &= A_1^a\Big[\frac{1}{t}\Big] + A_1^b\Big[\frac{1}{t}\Big] = L(a) + L(b). \end{aligned}$$

Also

$$L(1) = A_1^1\Big[\frac{1}{t}\Big] = 0,$$

so

$$0 = L(1) = L(a \cdot a^{-1}) = L(a) + L(a^{-1})$$

and it follows from this that

$$L(a^{-1}) = -L(a).$$

Hence

$$L\Big(\frac{a}{b}\Big) = L(a \cdot b^{-1}) = L(a) + L(b^{-1}) = L(a) - L(b).$$

5.75 Lemma. *For all $n \in \mathbf{Z}_{\geq 0}$, $L(a^n) = nL(a)$.*

Proof: The proof is by induction on n. For $n = 0$ the lemma is clear. Suppose now that the lemma holds for some $n \in \mathbf{Z}_{\geq 0}$, i.e., suppose that $L(a^n) = nL(a)$. Then

$$L(a^{n+1}) = L(a^n \cdot a) = L(a^n) + L(a) = nL(a) + L(a) = (n+1)L(a).$$

The lemma now follows by induction.

If $n \in \mathbf{Z}^-$ then $-n \in \mathbf{Z}^+$ and

$$L(a^n) = L\Big((a^{-n})^{-1}\Big) = -L(a^{-n}) = -(-n)L(a) = nL(a).$$

Thus equation (5.72) holds whenever $r \in \mathbf{Z}$. If $p \in \mathbf{Z}$ and $n \in \mathbf{Z} \setminus \{0\}$, then

$$pL(a) = L(a^p) = L\Big((a^{\frac{p}{n}})^n\Big) = nL\Big(a^{\frac{p}{n}}\Big)$$

so

$$L\Big(a^{\frac{p}{n}}\Big) = \frac{p}{n}L(a).$$

Thus (5.72) holds for all $r \in Q$. |||

5.76 Theorem. *Let a and b be numbers such that $0 < a < b$. Then*

$$A_a^b\Big[\frac{1}{t}\Big] = \ln(\frac{b}{a}) = \ln(b) - \ln(a).$$

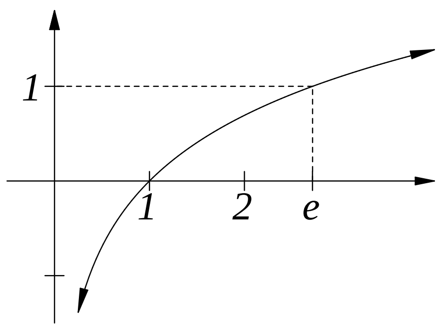
Proof: By lemma 5.58

$$A_a^b\Big[\frac{1}{t}\Big] = A_{aa^{-1}}^{ba^{-1}}\Big[\frac{1}{t}\Big] = A_1^{ba^{-1}}\Big[\frac{1}{t}\Big] = \ln(\frac{b}{a}) = \ln(b) - \ln(a). \parallel$$

Logarithms were first introduced by John Napier (1550-1632) in 1614. Napier made up the word *logarithm* from Greek roots meaning *ratio number*, and he spent about twenty years making tables of them. As far as I have been able to find out, the earliest use of ln for logarithms was by Irving Stringham in 1893[15, vol 2, page 107]. The notation $\log(x)$ is probably more common among mathematicians than $\ln(x)$, but since calculators almost always calculate our function with a key called "ln", and calculate something else with a key called "log", I have adopted the "ln" notation. (Napier did not use any abbreviation for logarithm.) Logarithms were seen as an important computational device for reducing multiplications to additions. The first explicit notice of the fact that logarithms are the same as areas of hyperbolic segments was made in 1649 by Alfons Anton de Sarasa (1618-1667), and this observation increased interest in the problem of calculating areas of hyperbolic segments.

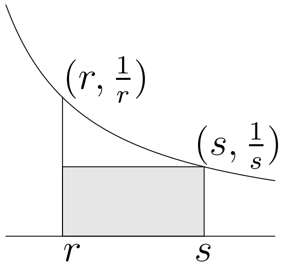
5.77 Entertainment (Calculate $\ln(2)$**.)** Using any computer or calculator, compute $\ln(2)$ accurate to 10 decimal places. You should not make use of any special functions, e.g., it is not fair to use the "ln" key on your calculator. There are better polygonal approximations to $A_1^2\Big[\frac{1}{t}\Big]$ than the ones we have discussed.

The graph of the logarithm function is shown below.



$y = \ln(x)$

We know that $\ln(1) = 0$ and it is clear that ln is strictly increasing.



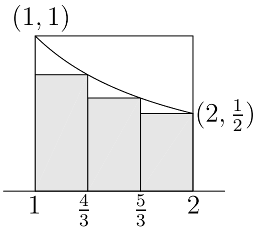
If $0 < r < s$, then

$$\ln(s) - \ln(r) = A_r^s\Big[\frac{1}{t}\Big] > (s-r)\frac{1}{s} > 0.$$

From the fact that $\ln(a^n) = n\ln(a)$ for all $n \in \mathbf{Z}$, it is clear that ln takes on arbitrarily large positive and negative values, but the function increases very slowly. Let

$$P = \{1, \frac{4}{3}, \frac{5}{3}, \frac{6}{3}\}$$

be the regular partition of $[1, 2]$ into three subintervals.



Then

$$\begin{aligned} \ln(2) &= A_1^2\Big[\frac{1}{t}\Big] \geq \alpha(I_1^2(\Big[\frac{1}{t}\Big], P)) \\ &= \frac{1}{3}\cdot\frac{3}{4} + \frac{1}{3}\cdot\frac{3}{5} + \frac{1}{3}\cdot\frac{3}{6} = \frac{1}{4} + \frac{1}{5} + \frac{1}{6} = \frac{37}{60}. \end{aligned}$$

Now

$$\ln(2) = A_1^2\Big[\frac{1}{t}\Big] \leq \alpha\Big(B(1, 2\colon 0, 1)\Big) = 1,$$

and

$$\ln(4) = \ln(2^2) = 2\ln(2) \geq 2 \cdot \frac{37}{60} > 1,$$

i.e.,

$$\ln(2) \leq 1 \leq \ln(4). \tag{5.78}$$

There is a unique number $e \in [2, 4]$ such that $\ln(e) = 1$. The uniqueness is clear because ln is strictly increasing.

The existence of such a number was taken as obvious before the nineteenth century. Later we will introduce the *intermediate value property* which will allow us to prove that such a number e exists. For the time being, we will behave like eighteenth century mathematicians, and just assert that such a number e exists.

5.79 Definition (e.) We denote the unique number in $\mathbf{R}^+$ whose logarithm is 1 by e.

5.80 Exercise. Prove that $2 \leq e \leq 3$. (We already know $2 \leq e$.)

5.81 Entertainment (Calculate e.) Using any computing power you have, calculate e as accurately as you can, e.g., as a start, find the first digit after the decimal point.

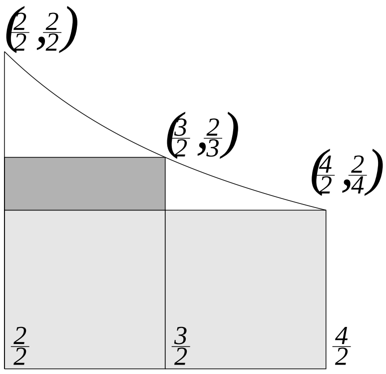
5.5 *Brouncker's Formula For $\ln(2)$

The following calculation of $\ln(2)$ is due to William Brouncker (1620-1684)[22, page 54].

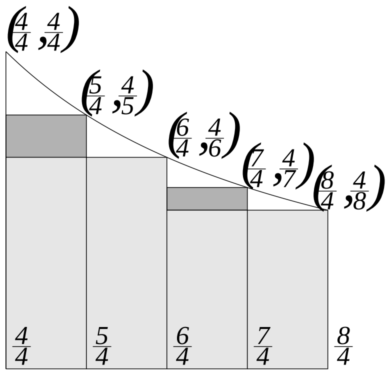
Let $P_{2^n} = \{x_0, x_1, \cdots, x_{2^n}\}$ denote the regular partition of the interval $[1, 2]$ into 2^n equal subintervals. Let

$$K(2^n) = I_1^2(\left[\frac{1}{t}\right], P_{2^n}) = \bigcup_{i=1}^{2^n} B(x_{i-1}, x_i; 0, \frac{1}{x_i}).$$

We can construct $K(2^{n+1})$ from $K(2^n)$ by adjoining a box of width $\dfrac{1}{2^{n+1}}$ to the top of each box $B(x_{i-1}, x_i; 0, \frac{1}{x_i})$ that occurs in the definition of $K(2^n)$ (see figures a) and b)).



$K(1)$ = lightly shaded region
$K(2)$ = total shaded region
figure a



$K(2)$ = lightly shaded region
$K(4)$ = total shaded region
figure b

We have

$$\alpha(K(1)) = \alpha(B(1, 2; 0, \frac{1}{2})) = 1 \cdot \frac{1}{2} = \frac{1}{2}.$$

From figure a) we see that

$$\begin{aligned}
\alpha(K(2)) &= \alpha(K(1)) + \alpha(B(\frac{2}{2}, \frac{3}{2}; \frac{2}{4}, \frac{2}{3})) \\
&= \frac{1}{2} + \frac{1}{2}\Big(\frac{2}{3} - \frac{2}{4}\Big) \\
&= \frac{1}{2} + \Big(\frac{1}{3} - \frac{1}{4}\Big) \\
&= \frac{1}{2} + \frac{1}{3 \cdot 4}.
\end{aligned}$$

From figure b) we see that

$$\begin{aligned}
\alpha(K(4)) &= \alpha(K(2)) + \alpha(B(\frac{4}{4}, \frac{5}{4}; \frac{4}{6}, \frac{4}{5})) + \alpha(B(\frac{6}{4}, \frac{7}{4}; \frac{4}{8}, \frac{4}{7})) \\
&= \alpha(K(2)) + \frac{1}{4}\Big(\frac{4}{5} - \frac{4}{6}\Big) + \frac{1}{4}\Big(\frac{4}{7} - \frac{4}{8}\Big) \\
&= \alpha(K(2)) + \Big(\frac{1}{5} - \frac{1}{6}\Big) + \Big(\frac{1}{7} - \frac{1}{8}\Big) \\
&= \frac{1}{1 \cdot 2} + \frac{1}{3 \cdot 4} + \frac{1}{5 \cdot 6} + \frac{1}{7 \cdot 8}.
\end{aligned}$$

In general we will find that

$$\alpha(K(2^n)) = \sum_{j=1}^{2^n} \frac{1}{(2j-1)(2j)}.$$

Now

$$0 \le \alpha(S_1^2(\left[\frac{1}{t}\right])) - \alpha(K(2^n)) \le (1 - \frac{1}{2})\mu(P_{2^n}),$$

i.e.

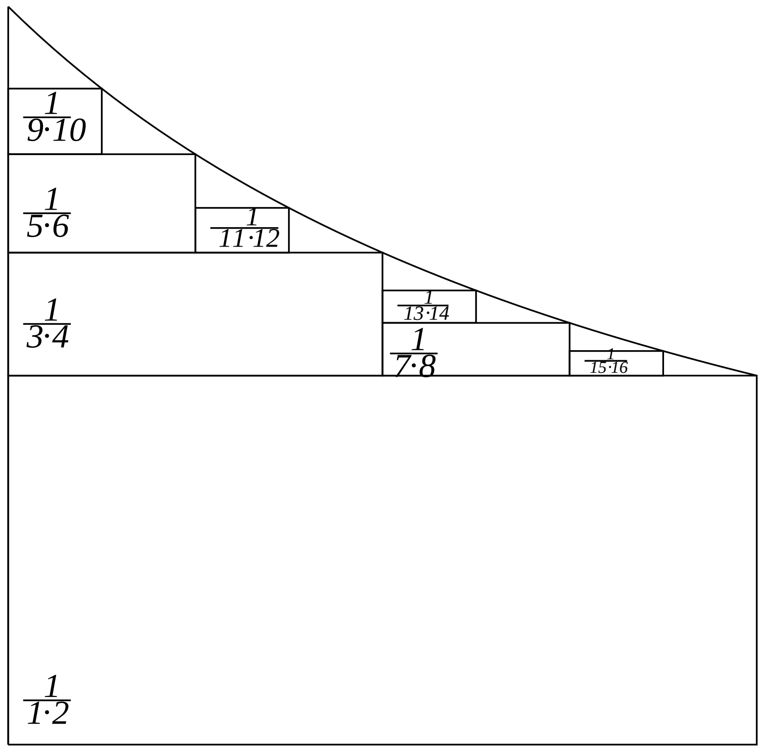
$$0 \le \ln(2) - \sum_{j=1}^{2^n} \frac{1}{(2j-1)(2j)} \le \frac{1}{2^{n+1}}.$$

Thus

$$\ln(2) = \sum_{j=1}^{2^n} \frac{1}{(2j-1)(2j)} \text{ with an error smaller than } \frac{1}{2^{n+1}}.$$

We can think of ln(2) as being given by the "infinite sum"

$$\ln(2) = \frac{1}{1 \cdot 2} + \frac{1}{3 \cdot 4} + \frac{1}{5 \cdot 6} + \frac{1}{7 \cdot 8} + \cdots. \tag{5.82}$$



$$\ln(2) = \tfrac{1}{1\cdot 2} + \tfrac{1}{3\cdot 4} + \tfrac{1}{5\cdot 6} + \cdots +$$

Equation (5.82) is sometimes called Mercator's expansion for ln(2), after Nicolaus Mercator, who found the result sometime near 1667 by an entirely different method.

Brouncker's calculation was published in 1668, but was done about ten years earlier [22, pages 56-56].

Brouncker's formula above is an elegant result, but it is not very useful for calculating: it takes too many terms in the sum to get much accuracy. Today, when a logarithm can be found by pressing a button on a calculator, we tend to think of "ln(2)" as being a known number, and of Brouncker's formula as giving a "closed form" for the sum of the infinite series $\dfrac{1}{1\cdot 2} + \dfrac{1}{3\cdot 4} + \dfrac{1}{5\cdot 6} + \cdots$.

5.6 Computer Calculation of Area

In this section we will discuss a Maple program for calculating approximate values of $A_a^b f$ for monotonic functions f on the interval $[a, b]$. The programs will be based on formulas discussed in theorem 5.40.

Let f be a decreasing function from the interval $[a, b]$ to $\mathbf{R}_{\geq 0}$, and let $P = \{x_0, x_1, \cdots, x_n\}$ be a partition of $[a, b]$. We know that

$$\alpha(I_a^b(f, P)) \leq A_a^b f \leq \alpha(O_a^b(f, P)),$$

where

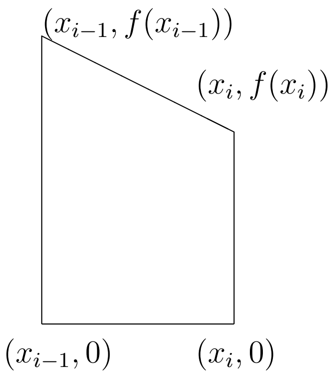
$$\alpha(I_a^b(f, P)) = \sum_{i=1}^{n} (x_i - x_{i-1}) f(x_i), \tag{5.83}$$

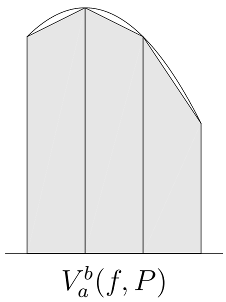
$$\alpha(O_a^b(f, P)) = \sum_{i=1}^{n} (x_i - x_{i-1}) f(x_{i-1}). \tag{5.84}$$

Let $V_a^b(f, P)$ be the average of $\alpha(I_a^b(f, P))$ and $\alpha(O_a^b(f, P))$, so

$$V_a^b(f, P) = \frac{\alpha(I_a^b(f, P)) + \alpha(O_a^b(f, P))}{2} = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \frac{f(x_i) + f(x_{i-1})}{2}.$$

Now $(x_i - x_{i-1}) \cdot \dfrac{f(x_i) + f(x_{x-1})}{2}$ represents the area of the trapezoid with vertices $(x_{i-1}, 0)$, $(x_{i-1}, f(x_{i-1}))$, $(x_i, f(x_i))$ and $(x_i, 0)$, so $V_a^b(f, P)$ represents the area under the polygonal line obtained by joining the points $(x_{i-1}, f(x_{i-1}))$ and $(x_i, f(x_i))$ for $1 \leq i \leq n$.





In the programs below, `leftsum(f,a,b,n)` calculates

$$\sum_{j=1}^{n} f\left(a + (j-1)\Big(\frac{b-a}{n}\Big)\right)\Big(\frac{b-a}{n}\Big) = \Big(\frac{b-a}{n}\Big)\sum_{j=1}^{n} f\left(a + (j-1)\Big(\frac{b-a}{n}\Big)\right),$$

which corresponds to (5.84) when P is the regular partition of $[a, b]$ into n equal subintervals, and `rightsum(f,a,b,n)` calculates

$$\sum_{j=1}^{n} f\left(a + j\Big(\frac{b-a}{n}\Big)\right)\Big(\frac{b-a}{n}\Big) = \Big(\frac{b-a}{n}\Big)\sum_{j=1}^{n} f\left(a + j\Big(\frac{b-a}{n}\Big)\right).$$

which similarly corresponds to (5.83). The command `average(f,a,b,n)` calculates the average of `leftsum(f,a,b,n)` and `rightsum(f,a,b,n)`.

The equation of the unit circle is $x^2 + y^2 = 1$, so the upper unit semicircle is the graph of f where $f(x) = \sqrt{1-x^2}$. The area of the unit circle is 4 times the area of the portion of the circle in the first quadrant, so

$$\pi = 4A_0^1[\sqrt{1-t^2}].$$

Also

$$\ln(2) = A_1^2\Big[\frac{1}{t}\Big].$$

My routines and calculations are given below. Here `leftsum`, `rightsum` and `average` are all procedures with four arguments, `f,a,b,` and `n`.

`f` is a function.

`a` and `b` are the endpoints of an interval.

`n` is the number of subintervals in a partition of `[a,b]`.

The functions `F` and `G` are defined by `F`$(x) = 1/x$ and `G`$(x) = \sqrt{1-x^2}$. The command

```
average(F,1.,2.,10000);
```

estimates $\ln(2)$ by considering the regular partition of $[1,2]$ into 10000 equal subintervals. and the command

```
4*average(G,0.,1.,2000);
```

estimates π by considering the regular partition of $[0,1]$ into 2000 equal subintervals.

```
> leftsum :=
>       (f,a,b,n) -> (b-a)/n*sum(f( a +((j-1)*(b-a))/n),j=1..n);
```

$$leftsum := (\,f, a, b, n\,) \to \frac{(\,b-a\,)\left(\sum_{j=1}^{n} f\left(a + \frac{(j-1)\,(\,b-a\,)}{n}\right)\right)}{n}$$

```
> rightsum :=
>       (f,a,b,n) -> (b-a)/n*sum(f( a +(j*(b-a))/n),j=1..n);
```

$$rightsum := (\,f, a, b, n\,) \to \frac{(\,b-a\,)\left(\sum_{j=1}^{n} f\left(a + \frac{j\,(\,b-a\,)}{n}\right)\right)}{n}$$

```
> average :=
>       (f,a,b,n) -> (leftsum(f,a,b,n) + rightsum(f,a,b,n))/2;
```

$$average := (\,f, a, b, n\,) \to \frac{1}{2}\,\mathrm{leftsum}(\,f, a, b, n\,) + \frac{1}{2}\,\mathrm{rightsum}(\,f, a, b, n\,)$$

```
> F := t -> 1/t;
```

$$F := t \to \frac{1}{t}$$

```
> leftsum(F,1.,2.,10000);
```

$$.6931721810$$

```
> rightsum(F,1.,2.,10000);
```

$$.6931221810$$

```
> average(F,1.,2.,10000);
```

$$.6931471810$$

```
> ln(2.);
```

$$.6931471806$$

```
> G := t -> sqrt(1-t^2);
```

$$G := t \to \mathrm{sqrt}(\,1 - t^2\,)$$

```
> 4*leftsum(G,0.,1.,2000);
```

$$3.142579520$$

```
> 4*rightsum(G,0.,1.,2000);
```

$$3.140579522$$

```
> 4*average(G,0.,1.,2000);
```

$$3.141579521$$

```
> evalf(Pi);
```

$$3.141592654$$

Observe that in these examples, `average` yields much more accurate approximations than either `leftsum` or `rightsum`.