UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2016:44

Origins of Integration

Olle Hammarström

Examensarbete i matematik, 15 hp
Handledare: Vera Koponen
Examinator: Veronica Crispin Quinonez
Juni 2016

Department of Mathematics
Uppsala University

Origins of Integration

Olle Hammarström

Supervisor: Dr. Vera Koponen

Uppsala University

Sweden

2016

Contents

Chapter 1

Introduction

There can be no doubt that integration is one of the most important mathematical concepts ever conceived throughout the entire history of mathematical research; not only does integration, in the framework of calculus, appear as an indispensable part of essentially any subcategory of applied mathematics, but it also appears in seemingly unrelated fields - it really has attained, along with differentiation, a status as a basic tool of mathematical understanding. While it, or methods like it, have been studied for thousands of years, and are continually being advanced even today, it has retained, at its core, its basic nature: integration is, fundamentally, the study of areas, volumes; of the sizes of enclosed spaces.

Viewed in this light, it should hardly come as a suprise that the originators of this concept were the geometers of the ancient world; the giants of Ancient Greece, whose findings are still releveant today, over two millenia later. However, the Greeks were not the first to approach the subject. Thales of Miletus, often hailed as the first mathematician, owed a great deal to the Egyptians and their mastery of practical geometry; he elevated their results by taking the measurements performed by the Egyptians and turning them into general truths - into mathematical theorems - and in so doing ushered in the golden age of Greek mathematics [10].

During this time, there were many attempts at solving problems which resembled, to some extent, the way we now work with integral theory. Among the earliest of these was when Democritus of Thrace found the formula[1] for the volume of a cone or pyramid, which he accomplished by considering these shapes as being made up of an innumerable amount of layers [10]. However, Democritus found himself pondering the curious situation that this layered approach gave rise to: if the layers are differently sized (i.e. if they taper off towards a point), then the shape must be irregular, and if they are the same size they will form a cylinder rather than a cone - in either case, it seems impossible to form the desired shape. We now know that we can overcome this problem through the use of infinitesimal calculus and limiting values, but how far Democritus took this early foray into the field is unknown [10].

The way the Greeks found to attempt to resolve the troubling situation above is now known as the method of exhaustion, and is the central topic of Chapter 2. It was pioneered by Eudoxus of Cnidus, who through it managed to create a solid foundation for the study of irrational numbers, and thus fill in the blanks in many of the older Greek theorems whose proofs had been questioned after the discovery of the irrational [10]. How the method was used in practice would differ from case to case, but its central notion was that of approximating a complicated shape by a simpler one in such a way that the simpler shape could be iterated in some fashion, such as adding sides to a polygon, in order to have it resemble the complicated shape more closely [2].

While the method of exhaustion was used to great effect in Ancient Greece,

[1] Democritus discovered the formula, but Eudoxus was the first to provide a rigorous proof [8].

it, and mathematics at the time, had its limitation; for one, the method needed to be adapted to each new shape - there was no general procedure that could be applied with minimal modifications to any problem. Furthermore, the framework of Greek mathematics had no ability to deal with the infinite, infinite sums, or limit processes. Because they could not be considered in any sense that would allow for mathematical rigour, they were taboo in the mathematical world of Ancient Greece [4]. Even though Archimedes and others came close to these concepts in their work with the method of exhaustion, they were still limited in their ability to progress the subject and approach the modern concept of an integral [4].

The problems the Greeks faced have been most famously illustrated by Zeno in his paradox regarding Achilles and the Tortoise. In essence, Zeno argued that if a Tortoise is given a head start, then Achilles, in order to overtake the Tortoise, must first reach the point where the Tortoise started. However, by that time the Tortoise will have moved some distance further, and hence Achilles must now reach a new point before he can overtake the Tortoise. This process continues indefinitely, seemingly implying that Achilles can never overtake the Tortoise. However, in reality he of course can do so, which creates a paradox. The problem, then, is not in understanding *that* Achilles overtakes the Tortoise, but rather *how* he does so [10].

Eventually, the glory days of Greek mathematics passed, and Europe, overcome by wars, started to enter the period of scientific decline in the early Middle Ages known as the Dark Ages. Even after passing through to the more prosperous period of the Middle Ages, substantial knowledge had to be reacquired by the new mathematicians. Because of this, there would be some time before significant progress was made in the theory of integration - but it was not solely a negative thing. While the Greek mathematical tradition was mostly lost in

Europe, only surviving through its preservation in Arab works, discussions on the infinitely large and small were no longer hampered by the strict notion of mathematical rigour previously imposed, and many advancements in this area were made [4]. While there were, at times, significant holes in the logic employed, this did not hinder the appearance of several important results, such as the analytical geometry of Descartes and Fermat, nor did it prevent the concept of the infinite to start entering mathematical discourse - something which would be used to great effect in the development of the calculus in the late 17^{th} century [4]. We also have appearing during this time new notational advancements through Viète, based in work by Diophantos, allowing for easier transcription of mathematical ideas [4].

In this new era of mathematical research, the first major advances in the theory of integration came with Johannes Kepler and Bonaventura Cavalieri; of these, Cavalieri's work was the most profound, and hence will be the focus of Chapter 3. Kepler's contributions to the development of integration were by no means minor, as anyone familiar with his work would know, however, his method was also more *ad hoc* in nature - a product of the mathematical atmosphere at the time - as can be illustrated by the fact that his second law of planetary motion was arrived at through a series of errors, which fortunately served to cancel each other out [4].

Cavalieri's contribution was through the method of indivisibles, and lives on today through Cavalieri's principle [1]. The method will be discussed in detail in Chapter 3, but in short it relies on comparing segments on figures, rather than the figures themselves, to conclude areas of new shapes based on knowledge of other shapes. In this regard it is similar to the method of exhaustion of Ancient Greece, and was indeed inspired by it [1], but differs in the regard that Cavalieri did not need to construct any polygonal approximations, nor did he need to rely on a *reductio ad absurdum* argument [1]. Furthermore, through its use of representation of figures, rather than the figures themselves, Cavalieri attempted somewhat successfully to extend the method of exhaustion and allow it to more easily deal with a more vast and complicated body of problems [1].

Unfortunately, Cavalieri's own work was both lengthy and difficult to follow, leading to a situation where it was mainly disseminated among the mathematical community in Europe through the use of third parties - and these third parties sometimes failed to properly communicate, or understand, the subleties and details of his work [1]. Most notable of these was Evangelista Torricelli, who popularised a method he referred to as "Cavalieri's method of indivisibles", even though it relied on a different, and less rigorous, foundation; Cavalieri's original method explicitly avoided infinite summation procedures due to the inability at the time to make such methods rigorous, but Torricelli modified it to a state similar to numerical integration, making explicit use of these infinite sums [1]. It was, largely, Torricelli's version, and others similar to it, that spread throughout Europe, causing some undue criticism to be directed at Cavalieri at the time [1].

This misinformation did have some beneficial effects as well, since it helped to advance the discussion on the kind of processes that would eventually crystallize into the familiar methods in the calculus of Newton and Leibniz [1]. During and after Cavalieri's time, rapid advancements were being made regarding infinite processes and series, which eventually produced a rigorous system of analysis by way of infinite series towards the latter half of the 17^{th} century [4]. With this, the world was set for the emergence of calculus, which will be the topic of Chapter 4.

The two principal actors on the calulus stage were, without a doubt, Isaac Newton and Gottfried Wilhelm Leibniz. Although their work differed in many regards, they can both be credited with discovering calculus as we know it today [4]. While there were already many well-known and efficient methods for dealing with the problems of computing tangents and areas, they were special methods applied to particular problems [4]. What Newton and Leibniz did was to create a system, called calculus, which allowed for these specialised methods to become general algorithmic procedures [4], and forever linked tangent problems and quadratures via the fundamental theorem of calculus.

Since this paper is focused on the history of integration, and not calculus as a whole, discussion on the development of tangent problems has been, and will continue to be, mostly omitted. However, as the fundamental theorem of calculus shows, these two topics are inextricably linked. It is this theorem that, during the 17$^{\text{th}}$ century onwards, allowed a veritable explosion of mathematical progress by taking problems and questions regarding specific instances and fitting them into a general framework, which we now call calculus [4].

In order to formalize this system, a rigorous and systematic treatment of limits was necessary. This is what was, to a large extent, provided by Newton and Leibniz, the former in his *Principia Mathematica* and the latter in various letters and essays, detailed in the *Historia et origo calculi differentialis*. Their approaches to the calculus differed greatly, with Newton favouring a time-based interpretation through his fluxions, and Leibniz establishing his theory based on characteristic triangles [4].

While Newton and Leibniz certainly created a systematic framework for the study of this class of problems, it was not without faults. Developments were made by Augustin-Louis Cauchy, Bernhard Riemann, and others, in the following centuries; these brought us the familiar definition of an integral as a limit of a summation procedure [4]. Furthermore, in order for the calculus to reach its full potential, a more rigorous theory of functions was required.

This was provided by Leonhard Euler in the 18th century - his work was so thorough and used such advanced notation, that it could almost be mistaken for a modern work - later improved by Henri Lebesgue [4]. Lastly, we were given a mathematical definition of area by Guiseppe Peano in the 19th century, and the doubts and problems surrounding the role of the infinite and infinitesimal in mathematics were finally dispelled through the work of Abraham Robinson in the 20th century [4].

Thus we see that, throughout history, integral problems have been studied due to their great importance in geometrical considerations and, later, as part of the framework of calculus, owing to their status as the anti-derivative. This study has intrigued some of the greatest minds in human history, leading to great revelations and advancements of human knowledge. This story remains active today, especially when considering integration in more complicated dimensions, along more complicated axes, and in their association with advanced differential equations. This exposition, albeit necessarily incomplete, will hopefully serve to illuminate where this study have its roots, and how it grew through the millenia to end up where it is today.

Chapter 2

Polygons

2.1 Exhaustion in Euclid's *Elements*

A significant portion of Book XII of Euclid's *Elements* is concerned with the method of exhaustion [8]. There are a few propositions in particular which stand out as good examples of the method used in practice. Among them is Euclid XII. 2., which states: *circles are to one another as the squares on the diameters* [8]. In order to prove this, Euclid[1] uses a *reductio ad absurdum* argument - a common occurrence in proofs using the method of exhaustion - to show that the ratio of the areas is neither less nor greater than the ratio of the diameters, and hence the proposition must hold. Furthermore, he relies of Euclid XII. 1., see A.7, which states the equivalent result for similar polygons inscribed in circles. It is in the effort of connecting this proposition with the investigation at hand that the method of exhaustion is used, as will be shown below.

We will follow Euclid's argument, however we will also take advantage of modern notation to simplify the expressions, starting with the more easily un-

[1] We here refer to the fact that Euclid documented the proof, not that he discovered it. The first instance of the proof most likely came from Eudoxus [8].

derstood rewrite of the proposition itself:

Proposition 2.1. *Let* a_1*,* a_2 *be the areas of two circles, and let* d_1*,* d_2 *be their respective diameters. Then the following holds:*

$$\frac{a_1}{a_2} = \frac{d_1^2}{d_2^2}.$$

Proof. In this proof we will make significant use of Figure 2.1. We let the area of Figure 2.1(a) be a_1, and the area of Figure 2.1(b) be a_2. Now, suppose the proposition does not hold. Then

$$\frac{a_1}{S} = \frac{d_1^2}{d_2^2} \tag{2.1}$$

must hold for some $S \neq a_2$. We begin by assuming that $S < a_2$, and consider Figure 2.1(b) showing a circle with points containing an inscribed square with corners at $EFGH$.

We then draw tangents to the circle at the points $EFGH$ to construct another square, this time circumscribed, as is shown in Figure 2.1(c). The outer square is clearly twice the size of the inner square, and it is also larger than the circle. Hence the area of the inner square must be greater than half the area of the circle. This is our initial polygonal approximation and establishes the starting point for our iterative construction.

Now let the arcs cut off by the chords EF, FG, GH, and HE be bisected at K, L, M, and N respectively. We then form triangles surrounding the square in Figure 2.1(b). Drawing a tangent to the circle at K and completing the parallelogram, we can use a similar argument to the above to show that the area of each of the triangles is half the area of the parallelogram formed in this manner, and since the parallelogram is greater than the size of the segment it encompasses we conclude that each triangle is greater than half the size of the segment it is housed in. This provides the foundation for stepping from one iteration to the next, as we shall see below.

In these steps we have taken a circle, inscribed a square, and found a way to double the number of nodes in such a way that each time we do it, the difference between the area of the circle and the area of the inscribed polygon decreases. Since a_2 and S are constant, we can continue this process of bisecting the arcs and eventually arrive[2] at some shape which satisfies

$$a_2 - p_2 < a_2 - S \implies S < p_2,$$

where p_2 is the area of the polygon inscribed in the circle $EFGH$.

Assume we have continued the mentioned process until we have arrived at the situation where $S < p_2$, and for simplicity assume that this situation is the one displayed in Figure 2.1(b), such that the polygon with area p_2 has vertices $EKFLGMHN$. We then inscribe a similar polygon $AOBPCQDR$, with area p_1, in the circle $ABCD$, shown in Figure 2.1(a). Hence we have

$$\frac{p_1}{p_2} = \frac{d_1^2}{d_2^2} \qquad \text{(Euclid XII. 1., A.7)}$$

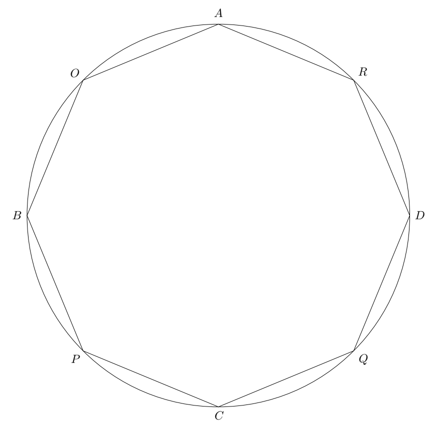
Furthermore, we have also assumed that

$$\begin{aligned} \frac{a_1}{S} &= \frac{d_1^2}{d_2^2} = \frac{p_1}{p_2} \\ \implies \frac{a_1}{p_1} &= \frac{S}{p_2} \end{aligned}$$

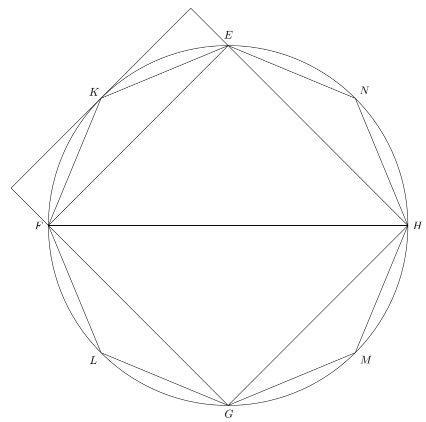
Here Euclid relied on A.3 and A.4, whereas we would recognize it as familiar algebraic manipulation of fractions.

Now, we know that $a_1 > p_1$, since p_1 is inscribed in a_1, which by the above

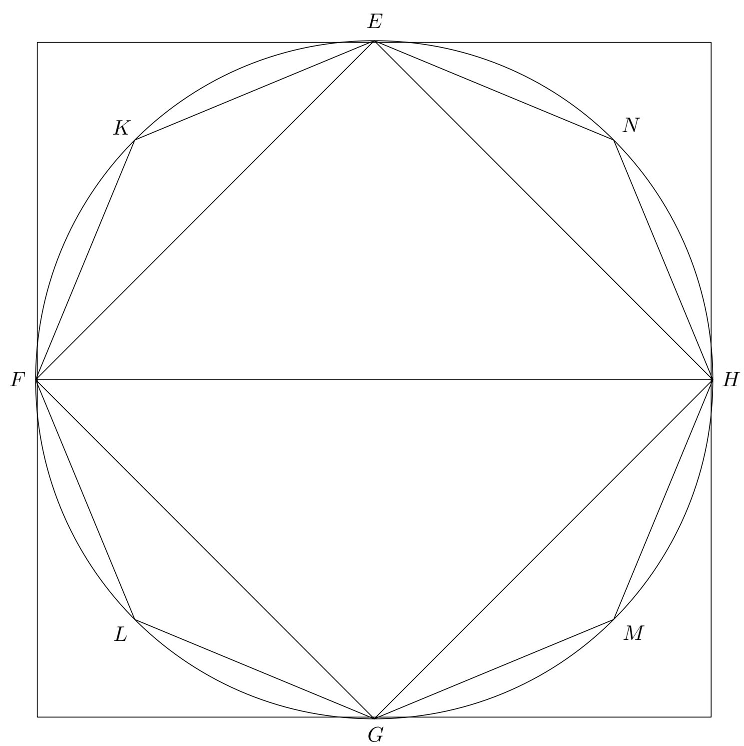
[2]See A.6.



(a) Circle with area a_1.



(b) Circle with area a_2.



(c) Circle with circumscribed square.

Figure 2.1: Figures used in the proof of Proposition 2.1.

implies that $S > p_2$: a contradiction, since we already assumed that $S < p_2$. Hence (2.1) does not hold for any $S < a_2$.

We next assume that $S > a_2$. By inverting the ratios, we have

$$\frac{d_2^2}{d_1^2} = \frac{S}{a_1}.$$

To proceed, we first need to establish the following lemma:

Lemma 2.1. *If $S > a_2$, then*

$$\frac{S}{a_1} = \frac{a_2}{T},$$

where $T < a_1$.

Proof. Rearraning, we have

$$\frac{S}{a_2} = \frac{a_1}{T}.$$

However, $S > a_2$, which implies that $a_1 > T$. Thus we have the desired result. □

Using Lemma 2.1, we can write

$$\frac{S}{a_1} = \frac{a_2}{T} = \frac{d_2^2}{d_1^2},$$

where $T < a_1$. This is analogous to the first case, which we have already proved to be impossible. Thus we have shown that S can be neither larger nor smaller than a_2, which implies that $S = a_2$ and the proposition holds. □

The method of exhaustion is crucial to this proof, since it provides the bridge between the comparatively simple result of Euclid XII. 1. to this more complicated situation. More specifically, the important step is showing that a circle can be "exhausted" - that we can inscribe a polygon in it and through an appropriate method increase its number of sides in such a way that we at each step cut away more than half of the remaining area.

2.2 Approximating π

Having established the general concept of the method of exhaustion, we now turn to the most prominent example of its use; computing an accurate approximation of π. While this task has been present for most of the history of mathematics - certainly as far back as there is surviving documentation - it was Archimedes of Syracuse who first produced a systematic approximation of π containing both a lower and an upper bound on the value [2], namely that

$$3 + \frac{10}{71} < \pi < 3 + \frac{1}{7}.$$

2.2.1 The Area of a Circle

Archimedes documented his efforts in the book *Measurement of the Circle*, wherein he proves three propositions about the circle [2]. The third of these concerns the bounds given above, but we will begin by focusing our efforts on the first, where Archimedes gives a specific expression for the area of any circle as

$$\text{Area } = \frac{rC}{2},$$

where r is the radius and C the circumference of the circle [2]. Knowing that $\pi = \frac{C}{2r}$, we can rewrite this as the more familiar

$$\text{Area } = \pi r^2.$$

We previously discussed how the Greeks worked by comparing areas of objects, and would therefore not produce any statement as the above. Archimedes was no different: his proposition reads

Proposition 2.2. *The area of any circle is equal to a right-angled triangle in which one of the sides about the right angle is equal to the radius, and the other to the circumference, of the circle [5].*

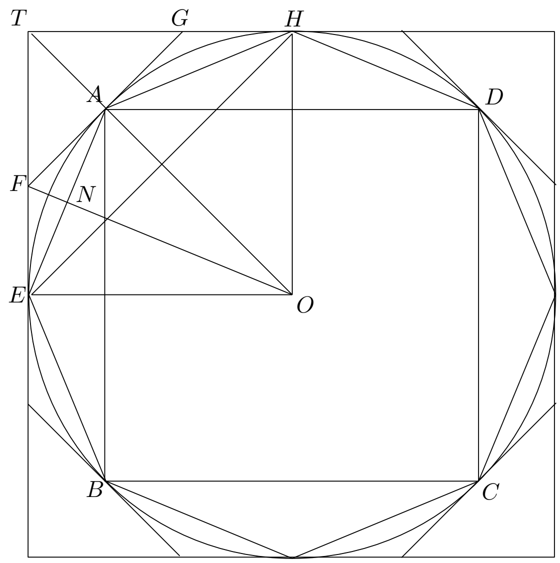
Archimedes constructs the mentioned triangle through clever use of the method of exhaustion, after which he shows that the area of the circle is neither less nor more than the area of the triangle, and hence must be equal to it [5]. His method is similar to the one Euclid used in Section 2.1, but uses both an inscribed and a circumscribed polygon.

Proof. To start, we present the illustrations in Figure 2.2 to help clarify the discussion. Here Figure 2.2(a) shows the circle in question with both inscribed and circumscribed polygons, as well as some additional lines which will be used in the proof, while Figure 2.2(b) is the mentioned triangle, of area K, with side lengths equal to the radius, r, and the circumference, c, on either side of the right angle. For simplicity, we let the area of the circle be denoted by a.

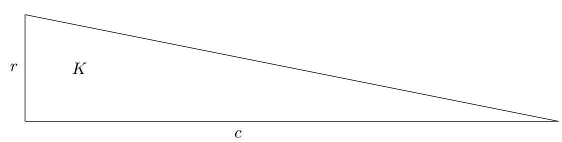
We start by noting that if the area of the circle is not equal to K, then it must be either less or greater than it. Assume that $K < a$, and inscribe a square in the circle, shown in Figure 2.2(a) as $ABCD$. Then bisect the arcs AB, BC, CD, and DA, forming the eight-sided polygon shown in Figure 2.2(a), and continue this process until the difference between the area of the circle and that of the inscribed polygon is less than the difference between that of the circle and the triangle[3]. In other words, letting p_i be the area of the inscribed polygon,

$$a - p_i < a - K \implies K < p_i.$$

[3]This is the same process as was used in the proof of 2.1 in Section 2.1.



(a) Circle with inscribed and circumscribed polygons.



(b) Triangle with area K

Figure 2.2: Figures used in the proof of Proposition 2.2.

Without loss of generality, we can assume that p_i is the inscribed regular eight-sided polygon in Figure 2.2(a). Consider the side AE, and construct ON perpendicular to this side. ON is by design less than the radius of the circle, and hence less than the side r in Figure 2.2(b). Furthermore, the perimeter of p_i is less than the circumference of the circle, that is, less than the side c in Figure 2.2(b). The area of the polygon is computed by summing the areas of the triangular segments, which by their regularity is given by

$$p_i = n * \frac{1}{2} * AE * ON,$$

where n is the number of triangular segments in the polygon. Since the area of the triangle is

$$K = \frac{1}{2} rc,$$

and $rc > n * AE * ON$, we see that

$$p_i < K.$$

However, this contradicts our conclusion that $K < p_i$, and hence we conclude that $K \not< a$.

Next we assume that $K > a$. We circumscribe a square, shown in Figure 2.2(a), which has a corner in T, and bisect the arcs such that OT passes through the midpoint of EH, marked by A, and FAG forms the tangent to the circle at A. Thus the angle TAG is a right angle, which implies that $GT > AG$.

Since $AG = GH$ by construction, we have the following chain of arguments:

$$\begin{aligned} & GT > GH \\ & \implies \frac{GT}{2} > \frac{GH}{2} \\ & \implies GT > \frac{GT}{2} + \frac{GH}{2} = \frac{HT}{2} \\ & \implies GT > \frac{HT}{2}. \end{aligned}$$

We then note that the area of the triangle AGT is given by $\frac{GT*h}{2}$, where h is the height of the triangle, and the area of the triangle AHT is similarly given by $\frac{HT*h}{2}$. Combining this with the above, we have that the area of AGT is greater than half the area of AHT, and by symmetry we conclude that the area of FGT is greater than half the area of $AETH$.

We can similarly bisect the arc AH and repeat the above argument to show that we again can cut off more than half of the area with a tangent at the point of bisection; this process is independent of the number of sides of the circumscribed polygon, which means that we can repeat it until we arrive at a situation where the difference between the circumscribed polygon and the circle is less than the difference between K and the circle. In other words, letting p_c denote the area of the circumscribed polygon, we can arrive at

$$p_c - a < K - a \implies p_c < K.$$

However, the area of the circumscribed polygon is given by

$$p_c = n * \frac{1}{2} * AO * GF,$$

where n is the number of sides in the polygon, and since AO is equal to the radius, and $n * GF$ is greater than the circumference, of the circle, we conclude that

$$p_c > a \implies a < p_c < K \implies a < K,$$

which contradicts our initial assumption. Hence K is not greater than the area of the circle, and since it is not less either we have established the proposition;

$$a = K = \frac{1}{2} r * c = \pi r^2.$$

□

It should be noted here that the use of a circumscribed polygon in addition to the inscribed one required a significant amount of new work, which hints at the fact that the method of exhaustion requires adaptation each time it is used for a new task. This, as was discussed earlier, is one of the fundamental problems of the method. That said, it is still a powerful tool for the right task, as we shall see next.

2.2.2 Bounds for π

The symbol for π is a relatively recent construction; it did not appear in mathematical work until the 18^{th} century, before which one would have to refer to the quantity by its property as the ratio between the circumference and the diameter of a circle [2]. The lack of a symbol did not, however, discourage mathematicians from attempting to find a value for the ratio, and Archimedes was no exception to this. Having established a formula to compute the area of a circle, he had found that it depended on the already well-studied ratio between the cicumference and diameter[4]. Naturally, he set about attempting to zero in on what the value of this ratio actually was, and his efforts are documented in the third proposition of *Measurement of the Circle* [5], given below

[4]Or, rather, the radius, though the only difference is a scaling factor of 1/2.

Proposition 2.3. *The ratio of the circumference of any circle to its diameter is less than* $3 + \frac{1}{7}$ *but greater than* $3 + \frac{10}{71}$.

The proof relies yet again on the method of exhaustion; it involves the use of two 96-sided polygons, one inscribed and one circumscribed, and includes quite a bit of manual computation, as well as fractional approximations for both $\sqrt{3}$ and several square roots of large numbers. The origin of these approximations will not be discussed, but it will be mentioned that it was a non-trivial task to compute them. For clarity, the proof will be divided into two lemmas, each proving the case of one of the given bounds.

Lemma 2.2. *The ratio of the circumference of any circle to its diameter is less than* $3 + \frac{1}{7}$.

Proof. For an illustration to clarify the steps below, please refer to Figure 2.3.

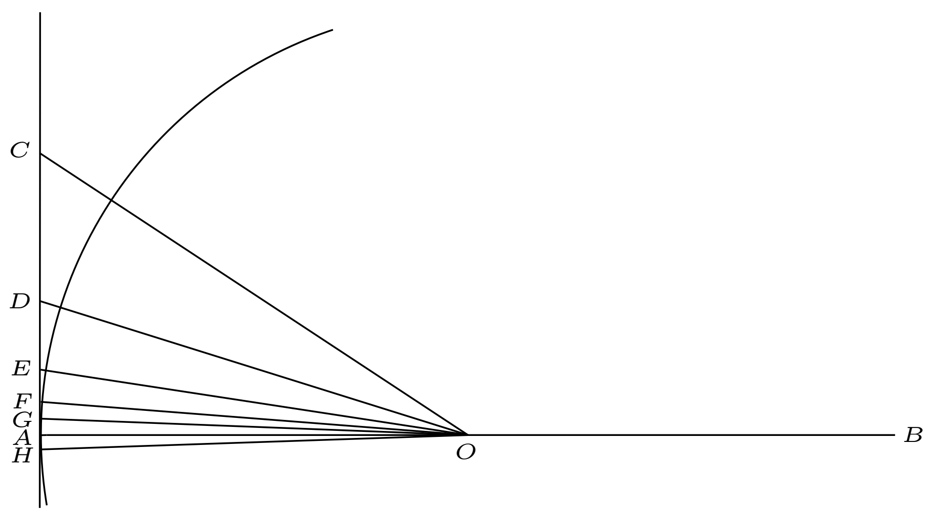


Figure 2.3: Illustration for the proof of Lemma 2.2

Let a circle with diameter AB centered at O be defined, and let AC be the tangent to the circle at A. Choose C such that $\angle AOC = 30°$. Then the triangle ACO is a $30 - 60 - 90$ triangle. Letting the length of the side AC be denoted by ℓ, and using the approximation $\sqrt{3} \approx \frac{265}{153}$, we have

$$\frac{AO}{AC} = \frac{\sqrt{3}\ell}{\ell} > \frac{265}{153} \tag{2.2}$$
$$\frac{CO}{AC} = \frac{2\ell}{\ell} = \frac{306}{153} \tag{2.3}$$

Now bisect $\angle AOC$, and let the line meet AC in the point D. We have

$$\begin{aligned} \frac{CO}{AO} &= \frac{CD}{AD} && \text{(Euclid VI. 3., A.5)} \\ \implies \frac{CO+AO}{AO} &= \frac{CD+AD}{AD} = \frac{AC}{AD} && \text{(Componendo, D.1)} \\ \implies \frac{CO+AO}{AC} &= \frac{CO}{AC} + \frac{AO}{AC} = \frac{AO}{AD} \end{aligned}$$

Applying (2.2) and (2.3), the above yields

$$\frac{AO}{AD} > \frac{571}{153} \tag{2.4}$$

We then have that

$$\begin{aligned} \frac{DO^2}{AD^2} &= \frac{AO^2+AD^2}{AD^2} && \text{(Pythagorean Theorem)} \\ &> \frac{571^2+153^2}{153^2} && \text{(From (2.4))} \\ \implies \frac{DO}{AD} &> \frac{591+\frac{1}{8}}{153}, && (2.5) \end{aligned}$$

where $\frac{591+\frac{1}{8}}{153}$ is an approximation to $\sqrt{\frac{571^2+153^2}{153^2}}$.

Continuing in the same fashion, we let EO bisect $\angle AOD$ such that it meets

AD in E. Then

$$\begin{aligned}
&\frac{DO}{AO} = \frac{ED}{AE} && \text{(Euclid VI. 3., A.5)}\\
&\implies \frac{DO + AO}{AO} = \frac{ED + AE}{AE} && \text{(Componendo, D.1)}\\
&\implies \frac{DO + AO}{AD} = \frac{AO}{AE} \\
&\implies \frac{AO}{AE} > \frac{591 + \frac{1}{8} + 571}{153} = \frac{1162 + \frac{1}{8}}{153} && \text{(From (2.4) and (2.5))}\\
&\implies \frac{EO^2}{AE^2} = \frac{AO^2 + AE^2}{AE^2} && \text{(Pythagorean Theorem)}\\
&\implies \frac{EO}{AE} > \frac{1172 + \frac{1}{8}}{153}, && (2.6)
\end{aligned}$$

where $\frac{1172+\frac{1}{8}}{153}$ is an aproximation to $\sqrt{\frac{(1162+\frac{1}{8})^2+153^2}{153^2}}$.

From here on we will omit some of the detail since the process is completely analogous to the preceding. Bisecting $\angle AOE$ by FO such that the line meets AE in F, we compute

$$\begin{aligned}
&\frac{AO}{AF} > \frac{1162 + \frac{1}{8} + 1172 + \frac{1}{8}}{153} = \frac{2334 + \frac{1}{4}}{153} && (2.7)\\
&\implies \frac{FO^2}{AF^2} > \frac{(2334 + \frac{1}{4})^2 + 153^2}{153^2} && \text{(Pythagorean Theorem)}\\
&\implies \frac{FO}{AF} > \frac{2339 + \frac{1}{4}}{153} && (2.8)
\end{aligned}$$

For the fourth and last time, bisect $\angle AOF$ by the line GO such that it meets AF in G, yielding

$$\begin{aligned}
\frac{AO}{AG} &> \frac{2334 + \frac{1}{4} + 2339 + \frac{1}{4}}{153} \\
&= \frac{4673 + \frac{1}{2}}{153}. \qquad (2.9)
\end{aligned}$$

We have now bisected $\angle AOC$ four times, which means the angle is

$$\frac{30^\circ}{2^4} = \frac{30^\circ}{16} = \frac{90^\circ}{48}.$$

If we then choose a point H on the other side of the diameter AB such that $\angle AOH = \angle AOG$ we have that

$$\angle GOH = \frac{90^\circ}{24} = \frac{360^\circ}{96},$$

which cleary shows that GH is one side of a regular polygon with 96 sides circumscribed about the circle.

Noting that $AO = \frac{1}{2}AB$ and $AG = \frac{1}{2}GH$, we use (2.9) to write

$$\begin{aligned}\frac{AO}{AG} &= \frac{\frac{1}{2}AB}{\frac{1}{2}GH}\\ &= \frac{AB}{GH} > \frac{4673 + \frac{1}{2}}{153}.\end{aligned}$$

Defining the perimeter of the circumsribed polygon as $P_c = GH * 96$, we get

$$\begin{aligned}&\frac{AB}{P_c} > \frac{4673 + \frac{1}{2}}{153 * 96} = \frac{4673 + \frac{1}{2}}{14688}\\ &\implies \frac{P_c}{AB} < \frac{14688}{4673 + \frac{1}{2}}\\ &\implies \pi < 3 + \frac{667 + \frac{1}{2}}{4673 + \frac{1}{2}} && (\pi < \tfrac{P_c}{AB})\\ &\implies \pi < 3 + \frac{667 + \frac{1}{2}}{4672 + \frac{1}{2}}\\ &\implies \pi < 3 + \frac{1}{7} && (2.10)\end{aligned}$$

Note that $\pi < \frac{P_c}{AB}$ follows from the fact that the perimeter of the circumscribed polygon is greater than the circumference of the circle.

□

Having found the upper bound for π, we now turn to the problem of finding the lower one. The method is very similar, but rather than using a circumsribed polygon we will be using an inscribed one instead.

Lemma 2.3. *The ratio of the circumference of any circle to its diameter is greater than* $3 + \frac{10}{71}$.

Proof. Please refer to Figure 2.4 for a pictoral representation of the below.

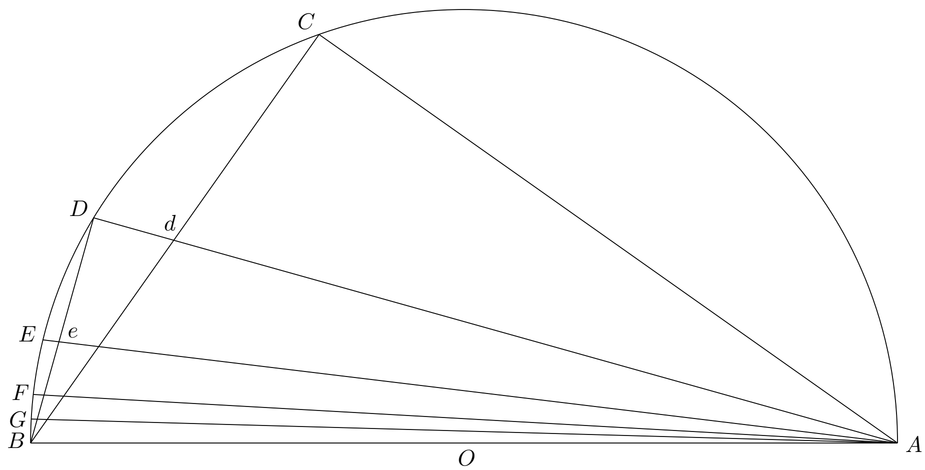


Figure 2.4: Illustration for the proof of Lemma 2.3

Let O be the center of a circle with diameter AB. Define a point C on the circumference of the circle such that $\angle CAB = 30°$. Then the $\angle ACB = 90°$ by Thales' Theorem (or the Inscribed Angle Theorem, D.2). Hence we have a $30 - 60 - 90$ triangle yet again which implies that

$$\frac{AC}{BC} = \frac{\sqrt{3}}{1} < \frac{1351}{780} \tag{2.11}$$

$$\frac{AB}{BC} = \frac{2}{1} = \frac{1560}{780} \tag{2.12}$$

where an approximation for $\sqrt{3}$ has been used.

Bisect $\angle CAB$ by a line meeting the circle in D and BC in d. Then $\angle ADB = 90°$ by Thales' Theorem and $\angle BAD = \angle dAC = \angle dBD = 15°$ by construction, which means that the three triangles ADB, ACd, and BDd are similar.

Because they are similar, we have the following

$$\begin{aligned} &\frac{AD}{BD} = \frac{BD}{Dd} = \frac{AC}{Cd} && \text{(Similarity)} \\ &\frac{AC}{AB} = \frac{Cd}{Bd} && \text{(Euclid VI. 3., A.5)} \\ \implies &\frac{AC}{Cd} = \frac{AB}{Bd} \\ \implies &\frac{AD}{BD} = \frac{AB+AC}{Bd+Cd} && \text{(D.1, property 4)} \\ &\quad = \frac{AB+AC}{BC}, \end{aligned}$$

which shows that $\frac{AD}{BD} = \frac{AB+AC}{BC}$.

From (2.11) and (2.12), we have

$$\frac{AD}{BD} = \frac{AB}{BC} + \frac{AC}{BC} < \frac{2911}{780} \tag{2.13}$$

$$\implies \frac{AB^2}{BD^2} < \frac{2911^2 + 780^2}{780^2} \qquad \text{(Pythagorean Theorem)}$$

$$\implies \frac{AB}{BD} < \frac{3013 + \frac{3}{4}}{780}, \tag{2.14}$$

where an approximation has been used for the square root of the right hand side.

Now bisect $\angle BAD$ with a line AE, such that the line meets the circle in E and BD in e, and join BE. Then we have a situation analogous to the preceding one, and we continue in much the same fashion. We first note that $\angle BAE = \angle eAD = \angle eBE$ from the construction through bisection. This implies that the triangles BAE, eAD, and eBE are similar. We then have

$$\frac{AE}{BE} = \frac{BE}{Ee} = \frac{AD}{De} \qquad \text{(Similarity)}$$

$$\frac{AD}{AB} = \frac{De}{Be} \qquad \text{(Euclid VI. 3., A.5)}$$

$$\implies \frac{AD}{De} = \frac{AB}{Be}$$

$$\frac{AE}{BE} = \frac{AB + AD}{Be + De} \qquad \text{(D.1, property 4)}$$

$$= \frac{AB + AD}{BD}$$

$$< \frac{3013 + \frac{3}{4} + 2911}{780} \qquad \text{(From 2.13 and 2.14)}$$

$$= \frac{1823}{240} \tag{2.15}$$

$$\frac{AB^2}{BE^2} < \frac{1823^2 + 240^2}{240^2} \qquad \text{(Pythagorean Theorem)}$$

$$\implies \frac{AB}{BE} < \frac{1838 + \frac{9}{11}}{240} \tag{2.16}$$

Continuing, we bisect $\angle BAE$ by AF meeting the circle in F. Since the situation is exactly analogous to the two preceding cases, we omit some detail here. We then have

$$\frac{AF}{BF} = \frac{AB + AE}{BE}$$

$$< \frac{1007}{66} \tag{2.17}$$

$$\implies \frac{AB^2}{BF^2} < \frac{1007^2 + 66^2}{66^2} \qquad \text{(Pythagorean Theorem)}$$

$$\implies \frac{AB}{BF} < \frac{1009 + \frac{1}{6}}{66} \tag{2.18}$$

For the fourth and last time: bisect $\angle BAF$ by AG, such that AG meets the circle in G. Then it follows in the same way that

$$\begin{aligned}
\frac{AG}{BG} &< \frac{2016+\frac{1}{6}}{66} && \text{(From (2.17) and (2.18))}\\
\frac{AB}{BG} &< \frac{2017+\frac{1}{4}}{66} && \text{(Pythagorean Theorem)}\\
\implies \frac{BG}{AB} &> \frac{66}{2017+\frac{1}{4}} && (2.19)
\end{aligned}$$

Since the $\angle BAG$ has been created from four bisections of a $30°$ angle, we have

$$\angle BAG = \frac{30°}{16}.$$

From the Inscribed Angle Theorem, see D.2, we then have that

$$\angle BOG = 2\angle BAG = \frac{30°}{8} = \frac{90°}{24}.$$

This means that the line BG is the side of an inscribed, regular, 96-sided, polygon. Letting the perimeter of the polygon be denoted by P_i, we have from (2.19) that

$$\begin{aligned}
\frac{P_i}{AB} &> \frac{6336}{2017+\frac{1}{4}}\\
&> 3+\frac{10}{71}
\end{aligned}$$

Since P_i is necessarily smaller than the circumference of the circle, we have that

$$\pi > \frac{P_i}{AB} > 3+\frac{10}{71}.$$

Hence we have found a lower bound for the value of π.

□

Having proved both Lemmas 2.2 and 2.3, we have established the proof for

Proposition 2.3, and can state that

$$3 + \frac{10}{71} < \pi < 3 + \frac{1}{7}.$$

2.3 Summary

We have introduced the method of exhaustion as it was used in Ancient Greece through some of its most prominent examples of application. In Section 2.1, we first established a baseline understanding of the process by the relatively simple example, from Euclid's Elements, of linking the areas of circles to their diameters. It was found that the critical step in the use of the method of exhaustion is finding the appropriate way to construct a more detailed polygonal approximation based on a previous iteration, so that the polygon does indeed move closer in area to the desired shape.

In Section 2.2, we took this idea even further to establish the formula for computing the area of a circle by comparing it to a triangle. This comparison was made possible due to the special relationship between the triangle and the segments which made up the polygonal approximation to the circle, where each segment was an isoceles triangle with side length equal to the radius. Yet again, this illustrated the importance of choosing an appropriate polygonal construction, and how the entire method relies on it.

Lastly, we involved ourselves with a more practical example, and one of incredible historical importance, when we followed Archimedes' argument to establish both a lower and upper bound for π - something unprecedented in Archimedes' time [2]. In fact, his construction was of such significance that even two millenia later, his was still the preferred method of computing more accurate values of π, since by simply continuing the process one can continually refine the approximation [2].

Clearly, the method of exhaustion was a critically important development in Greek mathematics. It was, nonetheless, not without its flaws, the most glaring of which was that it had to be adapted to each situation; one polygonal construction may work for the circle, but it would have little luck in the problem of defining the area of a cone, where instead a pyramid would have had to be used. Furthermore, one has to define an iterative process in such a way that the approximation approaches the target shape, which may be difficult for non-convex shapes. Indeed, almost all examples of the method of exhaustion are of circles, cones, and cylinders, which all lend themselves to similar constructions as the ones in this chapter.

In fact, Democritus had already hinted at the necessary steps to overcome these difficulties when he pondered the curious situation of the volume of a cone. However, it would take over a millenia before Cavalieri finally resolved the situation through his method of indivisibles, surviving in modern form as Cavalieri's principle, allowing for a rigorous treatment of infinitesimal values in mathematics [1].

Chapter 3

Infinitesimals

3.1 Medieval Developments

When the works of the mathematicians of Ancient Greece, as well as those of their counterparts in the Arabic world, were translated to Latin in the 12$^{\text{th}}$ century, much of the content was too sophisticated for it to bear fruit right away - rather, it took until the 16$^{\text{th}}$ and 17$^{\text{th}}$ centuries until significant work was done based on these works [4]. However, there was still important work being done in medieval Europe, albeit of a different nature. The primary contributions were more philosophical in nature, and concerned speculations on the properties of continuity and variability [4].

These discussions did not serve to make many rigorous mathematical developments, but to make concepts such as infinitesimals more acceptable in the common mind, and no doubt helped foster the attitude of discovery in favour of rigour that led to so many important discoveries in the 16$^{\text{th}}$ and 17$^{\text{th}}$ centuries [4].

One of the most significant contributions of the period was made by Thomas

Bradwardine, Richard Swineshead, and other members of a group of natural philosophers and logicians at Merton College in Oxford. They provided the groundwork for the study of variability and motion, and produced the Merton Rule of uniform acceleration, also known as the mean speed theorem [4]. If we let d be the total distance traveled, v_o is the initial velocity, v_f is the final velocity, t is time, and we assume that acceleration is uniform over this time interval, then this can be denoted by the following equation:

$$d = \frac{1}{2}(v_o + v_f)\,t$$

These ideas spread to other parts of Europe, leading to the French scholar Nicole Oresme to consider the importance of graphical representations in these studies [4]. His proposition was to represent the intensity of a quality, in the above case the velocity, by means of a perpendicular line segment, as demonstrated in Figure 3.1, where v_m has been added to indicate the graphical depiction of the mean velocity.

Oresme assumed, without proof, that the area of the trapezoid was equal to the total distance traveled, possibly by considering the shape to be made up of the infinitely many line segments produced in this manner [4]. This idea of shapes being made up of lines was far from established, but still spread in the mathematical community in this era of less-than-rigorous mathematical development. It would have significant impact, though its effects would not be fully felt until the 17$^{\text{th}}$ century when Cavalieri constructed a much more fundamental framework of working with shapes as made up of other shapes.

Another area of mathematics that saw significant development during this time period was that of the summation of infinite series. Swineshead and Oresme were the main contributors, and successors generally failed to improve significantly on their work in the following two centuries [4]. The results obtained

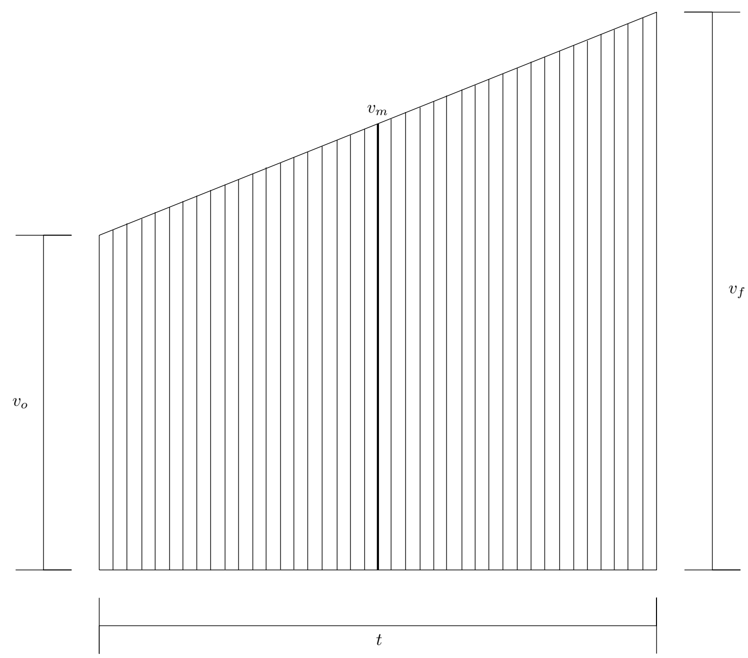


Figure 3.1: Oresme's graphical representation

during this period where not as important as the effect they had in encouraging the study of problems involving infinite series, promoting an acceptance of the infinite as something that could genuinely be studied through mathematical means [4].

The final touch to set the stage for serious advancements in infinitesimal mathematics came by the hand of Descartes[1] and Fermat in the 17$^{\text{th}}$ century, who through their essays titled *Geometry* and *Introduction to Plane and Solid Loci*, respectively, served to lay the foundation for the field of analytic geometry [4]. What they did was to establish the central idea of a correspondance between an equation and a locus, in such a way that the locus consists of all the points whose coordinates satisfy the equation with respect to a set of two coordinate axes. For example, consider Figure 3.2.

[1]Also known as Cartesius, from where we get the Cartesian coordinate system [10].

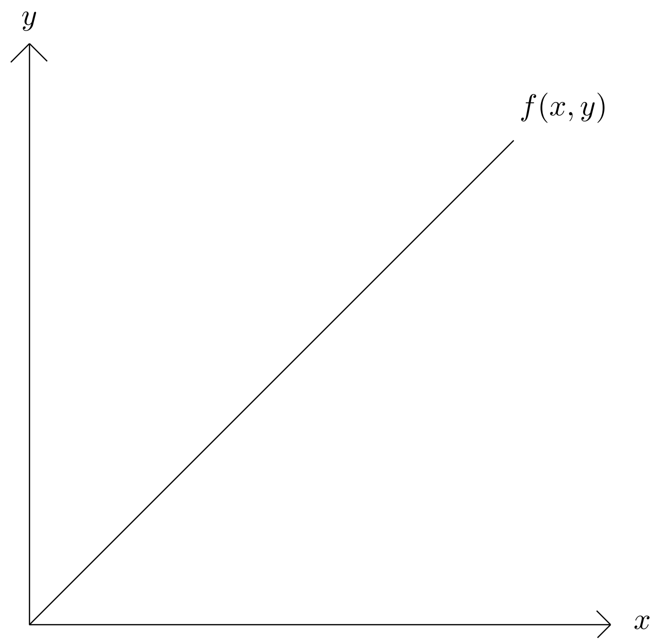


Figure 3.2: $f(x, y)$ given by $y - x = 0$

In this simple case, we have chosen two perpendicular straight lines as axes and illustrated the locus $y - x = 0$ with respect to these axes. It is worth noting that while the idea of constructing graphs in this manner originated with Descartes and Fermat, they did not always employ the modern standard of using two perpendicular straight lines as axes [4]. While Descartes preferred to take a geometrical problem and translate it to algebraic form through this approach, Fermat tended to take the opposite route of beginning with some algebraic equation and produce the geometrical representation from it [4]. Something that they both put great focus on, however, was the study of indeterminate and continuous variables, whereas their predecessors, such as Viète, had only studied determinate variables [4]. Taken together, their collective work served to massively expand the study of geometry. Mathematicians were no longer bound by the relative sparcity of available curves; now a new curve could be constructed just as easily as a new equation could be written down. Furthermore, Descartes and Fermat had provided powerful new tools to analyse any geometrical construction, and had created a bridge between the fields of algebra and geometry.

3.2 Cavalieri's Indivisibles

The first systematic treatment of area and volume problems to significantly improve on the works of Ancient Greece came at the hand of Bonaventura Cavalieri in the 16^{th} century. Through two texts, the seven books of the *Geometria*[2] of 1635 and the six books of the *Exercitationes*[3] of 1647, he constructed a mathematical framework with the intent of bridging the gap between the method of exhaustion and the recent developments in the theory of continuity, and the called this the method of indivisibles [1]. His method may have had a less than rigorous foundation in certain parts, but it did regardless manage to overcome some of the problems that the method of exhaustion faced, in particular those relating to limit processes of the kind that Democritus had pondered.

Cavalieri's main work on the method of indivisibles was featured in the first six books of the *Geometria*, wherein he developed a system to work directly with the areas of two different figures by ascribing magnitudes (in the Greek sense) to them. There was also a later development, possibly inspired by the work others had done with his method, of comparing two different figures by way of their cross-sections, appearing in book seven of the *Geometria*, which more closely relates to the modern version of Cavalieri's principle: [1].

Theorem 3.1 (Cavalieri's Principle)**.** *If two solids have equal altitudes*[4]*, and if*

[2]Full title: *Geometria indivisibilibus continuorum nova quadam ratione promota* [1].
[3]Full title: *Excertationes geometricae sex* [1].
[4]Two solids have equal altitudes if we can find some direction such that the maximum length of the intersection between the solid and all lines traveling along this direction is equal for both solids.

sections made by planes parallel to the bases and at equal distances from them are always in a given ratio, then the volumes of the solids are also in this ratio [4].

A thorough treatment of Cavalieri's method goes beyond the scope of this paper. Instead, the main features of the method will be discussed along with an example of its use in practice, with some omissions and simplifications of the sake of brevity. For a more complete treatment of Cavalieri and his methods, please refer to Andersen's 1985 paper on the topic, [1].

3.2.1 The Method of Indivisibles

The essence of Cavalieri's collective method of indivisibles is the creation of a type of magnitude which can be used to represent the area of a figure. In order to do so, Cavalieri required quite a few assumptions and theorems, the most significant of which can be found in Appendix C.

One of the most crucial aspects of Cavalieri's work was its attempt to emulate Greek mathematics. Most significantly, the similarity between the method of indivisibles and the method of exhaustion is that they both rely on comparing figures. In Chapter 2 we saw comparisons between circles and regular polygons, and the fact that they are both plane figures is important since the Greek understanding of mathematics only allowed the comparison of magnitudes of the same type. Cavalieri adopted the same thinking, but attempted to extend the concept of a magnitude [1]. Perhaps the most fundamental definition for the application of his method was that of *omnes lineae*, or "all the lines", given here in somewhat simplified form.

Definition 3.1 (All the lines)**.** *If through opposite tangents to a given plane figure two parallel and indefinitely produced planes are drawn perpendicular to the plane of the given figure, and if one of the parallel planes is moved toward*

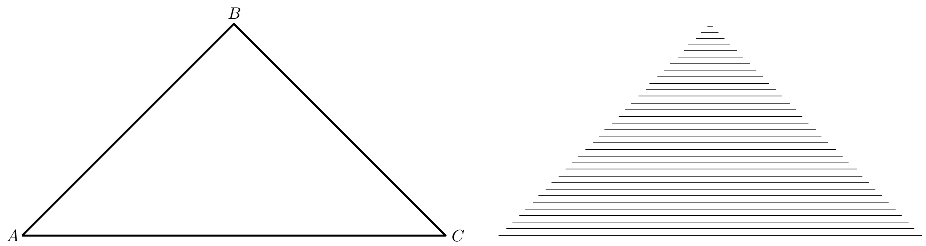


Figure 3.3: All the lines of a triangle ABC

the other, still remaining parallel to it, until it coincides with it; then the single lines which during the motion form the intersections between the moving plane and the given figure, collected together, are called all the lines, recti transitus, of the figure taken with one of them as regula[5] *[1].*

These lines are referred to as the indivisibles of the figure with respect to the given regula, and hence is where the method gets its name [1]. It should be noted that Cavalieri also accounted for the possibility that the planes are drawn with an incline, and thus not perpendicular, to the plane of the given figure, and called these "all the lines, obliqui transitus" [1]. However, these constructions were used to a much lesser extent, and we will therefore focus on the recti transitus case.

To help illustrate the concept of all the lines, consider Figure 3.3. Here the rightmost figure represents the collection of all the lines[6] created from the triangle ABC with regula AC.

From this, it is a natural conclusion that the collection of lines can be used to represent the area of the triangle, in the sense that general conclusions drawn about the collection of lines also holds for the area of the triangle. This, in essence, is Cavalieri's collective method of indivisibles, and he expressed this in

[5] A line parallel to a tangent to the figure, used as reference for the construction of the line segment intersections of the figure. See C.1.

[6] In this paper, "all the lines", "the collection of lines", "omnes lineae", and variations of these terms are to be understood to refer to the same thing.

Geometria via the following theorem.

Theorem 3.2 (Book II, Theorem 3)**.** *The ratio between the areas of two figures equal the ratio between their collections of lines taken with respect to the same regula [1].*

Cavalieri's proof of this theorem relies on several assumptions he made regarding collections of lines, listed below:

A. 1 The collections of lines of congruent figures are congruent[7].

A. 2 Collections of lines can be ordered. That is, if we have two collections of lines, A and B, then either $A > B$, $A = B$, or $A < B$ holds.

A. 3 Collections of lines can be added.

A. 4 A smaller collection of line can be subtracted from a greater collection of lines.

A. 5 If one figure is greater than another, then its collection of lines is greater than the other's.

A. 6 *The ut-unum principle:* As one antecedent is to one consequent so are all the antecedents to all the consequents[8].

Furthermore, we will need two lemmas.

Lemma 3.1. *Two collections of lines can form a ratio.*

Proof. Suppose two figures F_1 and F_2 have equal altitudes h. Choose some distance from a chosen regula, and denote the line segments intersecting the figures at this distance by ℓ_1 and ℓ_2, respectively. Now, ℓ_1 can be multiplied to exceed ℓ_2, and since these are chosen at arbitrary distances this will hold for any

[7] If two congruent figures are placed so that they coincide, then each line in their respective collections of lines will match up, and hence the collections of lines will be congruent.

[8] This assumption gives rise, almost directly, to Cavalieri's principle, mentioned earlier.

pair of corresponding line segments, and we can conclude that the collections of lines can be put in a ratio.

Now suppose that the two figures have different altitudes h_1 and h_2, respectively, and further assume, without loss of generality, that $h_1 > h_2$. We then split h_1 into two parts, h_3 and h_4, such that $h_3 = h_2$. For the sake of simplicity, we also assume that $h_4 < h_2$, since if this was not true we could simply repeat the process until it was.

We then draw a horizontal line through F_1 at the height h_3, creating two new figures. Denote the figure lying beneath this line by F_3 and the one above it by F_4, such that $F_1 = F_3 + F_4$. Since F_3 and F_2 have equal heights, we can apply the same principle as earlier to conclude that any line l_2 in F_2 can be multiplied to exceed its corresponding lines l_3 and l_4 in figures F_3 and F_4, respectively. Hence all the lines of F_2 can be put in a ratio with all the lines of F_3 and F_4. From the assumed additive property (A. 3) we have that all the lines of F_3+ all the lines of F_4 = all the lines of F_1, and it thus holds that the collection of lines of F_2 can be put in a ratio with F_1. □

In the above, Cavalieri entirely avoids the question of the existence of a minimum multiple m so that $m\ell_1 > \ell_2$, and hence his proof is rather loosely put together. Regardless, the lemma is significant, since it essentially underpins the entire construction of the method; if collections of lines can not form ratios, then the rest of the method fails.

Lemma 3.2. *If two figures have the same area*[9]*, then their collection of lines are equal.*

Proof. Suppose two figures, F_1 and F_2, have the same area. Superimpose F_1 on F_2, such that F_3 is the part of F_1 lying inside or on F_2, and F_4 is the residual

[9] Cavalieri developed his fundamentals using mainly area concepts, however, he also applied similar methods to volume problems [1].

part of F_1 outside of F_2. Repeat this process until all the residual parts have been distributed in or on F_2. There are now a series of figures F_3, F_4, ..., such that their sum $F_3 + F_4 + \dots$ is congruent with F_2. Then, from assumptions A. 1 and A. 3, we conclude that the collection of lines of F_1 is equal to the collection of lines of F_2. □

This proof is more problematic than the previous one. Specifically, the problem lies in sentence ”repeat this process until all the residual parts have been distributed in or on F_2”, since this process carries a risk of being infinite if there is no sufficiently nice way to superimpose the residuals. This would then lead to the introduction of infinite sums in Cavalieri's method, which would be unfortunate since he specifically wanted to avoid such things [1]. Cavalieri seemed to believe that a rigorous application of the method of exhaustion could redeem his proof in this event, although he never expanded on the idea [1]. In any case, we are now ready to present the proof of Theorem 3.2.

Proof. Let two figures be denoted F_1 and F_2, with their respective collections of lines, taken with the same regula, being denoted by O_1 and O_2. In order to prove the theorem in style of the Ancient Greeks, specifically with reference to Euclid V. 5 (A.2), Cavalieri needed to show the following:

$$n \times F_1 > m \times F_2 \implies n \times O_1 > m \times O_2 \tag{3.1}$$

$$n \times F_1 = m \times F_2 \implies n \times O_1 = m \times O_2 \tag{3.2}$$

$$n \times F_1 < m \times F_2 \implies n \times O_1 < m \times O_2 \tag{3.3}$$

(3.1) and (3.3) follow directly from the assumptions A. 3 and A. 5: the additive property and the ordering of collections of lines based on the figures giving rise to them. Furthermore, (3.2) can be derived from A. 3 and Lemma 3.2. □

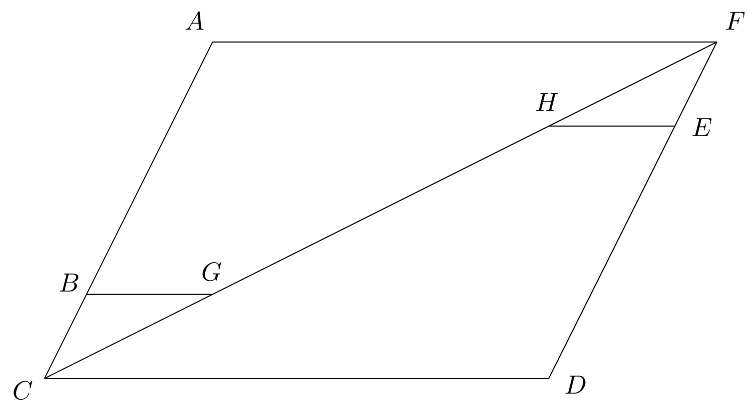


Figure 3.4: Illustration for Theorem 3.3

Having now established the fundamental support for the method of indivisibles, we turn to an example of its use. Cavalieri's first goal with his new method was not to find new results, but rather to show the validity of his method by computing already well-known results. One of the simpler examples is the following theorem.

Theorem 3.3. *If a diagonal is drawn in a parallelogram, the parallelogram is the double of each of the triangles determined by the diagonal [1].*

Proof. Consider the parallelogram $ACDF$ in Figure 3.4, having diagonal CF. For sake of clarity, let $\square ACDF$ denote the area enclosed by $ACDF$, and similarly let ΔACF be the area of the triangle ACF. We can then restate Theorem 3.3 in terms of this parallelogram as

$$\square ACDF = 2\Delta FAC = 2\Delta CDF$$

Cavalieri then proceeded to draw the arbitrary pair of corresponding line segments BG and EH. From the fact that the pair of lines are corresponding, we know that $BC = EF$, and we can use Euclid I. 26. (See A.1 in the appendix) to conclude that the triangles BCG and EFH are congruent[10], which in turn implies that $BG = EH$.

To apply Cavalieri's method, we then consider the regula to be the line CD, which is parallel to the lines BG and EH. This means that BG and EH are part of the collection of all the lines corresponding to this choice of regula. Furthermore, this choice of regula gives CD and AF as opposite tangents in the figure. Now, we have already shown that, at corresponding distances BC and EF from their respective tangents, $BG = EH$. We can then apply Theorem 3.2 to state that, since the ratio of the collection of lines is $1 : 1$, the ratio of the areas of the two triangles is also $1 : 1$. In other words, $\Delta ACF = \Delta CDF$, and since these two triangles by construction make up the parallelogram $ACDF$, we have the desired result. □

Another example application can be found in the computation of the volume of a circular cone. In this case we will use the later development, in chapter 7 of the *Executationes*, of Cavalieri's principle, to show that the volume of a cone of height h and base radius r is $\pi r^2 \frac{h}{3}$.

Proof. The essence of Cavalieri's principle is the comparison of something unknown with something known. For this reason, we will use the fact, known at Cavalieri's time, that the volume of a pyramid with square unit base is

$$\frac{h}{3},$$

where h is the height of the pyramid [4].

We then consider a cone of height h and base radius r. Now, instead of considering all the lines created from the intersection of a plane figure with

[10] The alert reader may here object, since the proof would be considerably simpler by using the same theorem from Euclid to conclude that the triangles ACF and CDF are congruent. However, as was mentioned earlier, Cavalieri's intent here was not to produce new results, but rather to illustrate his method.

perpendicular planes, as we did in Definition 3.1, we look at all the planes created when intersected the cone with planes perpendicular to its base. This gives us a concept of all the planes of the figure, which we use analogously to the collection of lines[11].

The cross-section of the pyramid at a distance d from its apex is given by

$$\frac{d^2}{h^2}.$$

Similarly, noting that the base of the cone is given by πr^2, the cross-section of the cone at a distance d from its apex is given by

$$\pi r^2 \frac{d^2}{h^2}.$$

Using these results, we note that the ratio of the cross-sections is $\pi r^2 \ : \ 1$, we use Cavalieri's principle, see Theorem 3.1, to conclude that the ratio of their volumes is also $\pi r^2 \ : \ 1$.

Since the volume of the pyramid is given by $\frac{h}{3}$, we thus conclude that the volume of the cone is

$$\frac{\pi r^2 h}{3}.$$

□

3.2.2 Evolution and Aftermath

It is important to point out that Cavalieri did not consider a figure to be made up of its indivisibles in the same sense that, for example, Kepler did. While Kepler made a similar construction and proceeded to "add up" the lines in an *ad hoc*, and less formal, fashion to compute the area of a figure [4], Cavalieri

[11] If this seems a bit *ad hoc*, it does so for good reason. Cavalieri himself focused his work on developing the theory for all the lines, and more or less assumed that this theory was valid for collections of other shapes, such as planes, as well [1].

sought to avoid any infinite summations in order to keep the mathematical rigour of his method intact [1]. Instead of using sums, Cavalieri compared the properties of the collections of lines of two figures directly, in much the same way that the Ancient Greeks did before him with the method of exhaustion.

However, Cavalieri's method proved difficult to disseminate in the mathematical community. This was likely due, in no small part, to the fact that the *Geometria* was 700 pages long, very difficult to follow, and that Cavalieri did not take advantage of the advances made in algebra and mathematical notation to simplify the discussion; rather, Cavalieri relied on long discussions, meaning that some theorems which could today be proven using one page and a figure took Cavalieri 90 pages to complete [1]. Thus it came to be that Cavalieri's work was distributed through third parties, one of the most notable being Evangelista Torricelli.

These third parties did not always get the content matter exactly right, leading to a situation where, while Cavalieri had strived to avoid any and all uses of infinite sums in his constructions, mathematicians outside Italy came to know the method of indivisibles as one specifically dealing with a figure as an infinite sum of line segments, popularised by Torricelli, or infinitesimals [1].

However, this confusion was not entirely a negative thing for mathematics at the time. While it is true that the infinite summation present in some of the interpretations made much of the foundation of the method questionable, and certainly distanced it from the ideal rigour of the method of exhaustion, it served to promote serious study in various methods of integration [1]. Hence, while Cavalieri's method of indivisibles, as he constructed it, did not directly give rise to modern calculus, it, and the interpretations other mathematicians made of it, helped stimulate research in the area, and was no doubt a significant development in the history of integration.

Chapter 4

Calculus

4.1 Early Calculus

Tangent constructions and instantaneous motion started to become a major part of mathematical investigations in the first half of the 17^{th} century, leading to many new and important results [4]. Similarly, many new theorems for quadratures of various functions were produced during this time [4]. However, the common theme in all this was specificity - the mathematicians of the day had yet to produce a framework, a generalised system, in which their specific results would fit. Some mathematicians came closer than others, with Evangelista Torricelli and Isaac Barrow often being credited as the first to seriously approach the fundamental theorem of calculus, at least on an intuitive level [10]. However, it was not until Isaac Newton and Gottfried Wilhelm Leibniz that this was made more rigorous, although it would still be some time before it was thoroughly understood and attained its modern form, given below [4].

Theorem 4.1 (The Fundamental Theorem of Calculus [11])**.** *If f is real on*

$[a, b]$ *and if there is a differentiable function* F *on* $[a, b]$ *such that* $F' = f$*, then*

$$\int_a^b f(x)dx = F(b) - F(a)$$

While Newton was the first to develop his ideas, he neglected to publish them until after Leibniz - a fact that led to much dispute, and created a rift between the mathematical communities of Britain and mainland Europe, to the detriment of the former [4]. This contributed to making Leibniz's work the most well-known, but it was not the sole reason. Leibniz notational system was so sophisticated and well-designed, that it made working with these problems significantly easier when compared to Newton's convoluted method of fluxions [4]. The genius of Leibniz' notational system is expressed by Edwards in the following quote [4]:

> His infinitesimal calculus is the supreme example, in all of science and mathematics, of a system of notation and terminology so perfectly mated with its subject as to faithfully mirror the basic logical operations and processes of that subject. It is hardly an exaggeration to say that the calculus of Leibniz brings within the range of an ordinary student problems that once required the ingenuity of an Archimedes or a Newton. Perhaps the best measure of its triumph is the fact that today we can scarcely discuss the results of Leibniz' predecessors without restating them in his differential notation and terminology

Besides notational differences, the works of Newton and Leibniz also differed in their respective foundations. Newton relied on an intuitive concept of continual motion and made explicit use of limit concepts, leading to his definition of a fluxion as what we would now call a derivative, whereas Leibniz' approach was more geometrical in nature and closer to the method of exhaustion in style, but effectively hid the role of limit concepts in its derivation [4]. As a result of this, Newton considered integrals in their indefinite form and as the inverse to rate-of-change problems, while Leibniz saw them as infinite sums of differentials [4]. To keep in the spirit with the geometrical focus of the preceding chapters, we will here focus on Leibniz' version of the calculus, as it lends itself more favourably to that approach.

One of the most important discoveries made by Leibniz was his transmutation method, through which he could derive almost all the previously known results on quadratures [4]. His approach was similar to that of Cavalieri, which can most effectively be shown through this excerpt of a letter he wrote to Newton:

> The basis of the transformation is this: that a given figure, with innumerable lines [ordinates] drawn in any way (provided they are drawn according to some rule or law), may be resolved into parts, and that the parts - or others equal to them - when reassembled in another position or another form compose another figure, equivalent to the former or of the same area even if the shape is quite different; whence in many ways the quadratures can be attained [4].

In other words, Leibniz conceived of a principle of transformation whereby two figures are divided into indivisible segments. If there then is a one-to-one correspondance between the segments of one with the segments of the other, then it is said that the latter is derived from the former, and that their areas are equal. This is very similar, albeit it more general, to Cavalieri's principle. In practice, however, the most important improvement Leibniz made over Cavalieri was that of using triangular indivisibles, rather than rectangular ones, and it allowed him to produce the following transumation theorem [4]

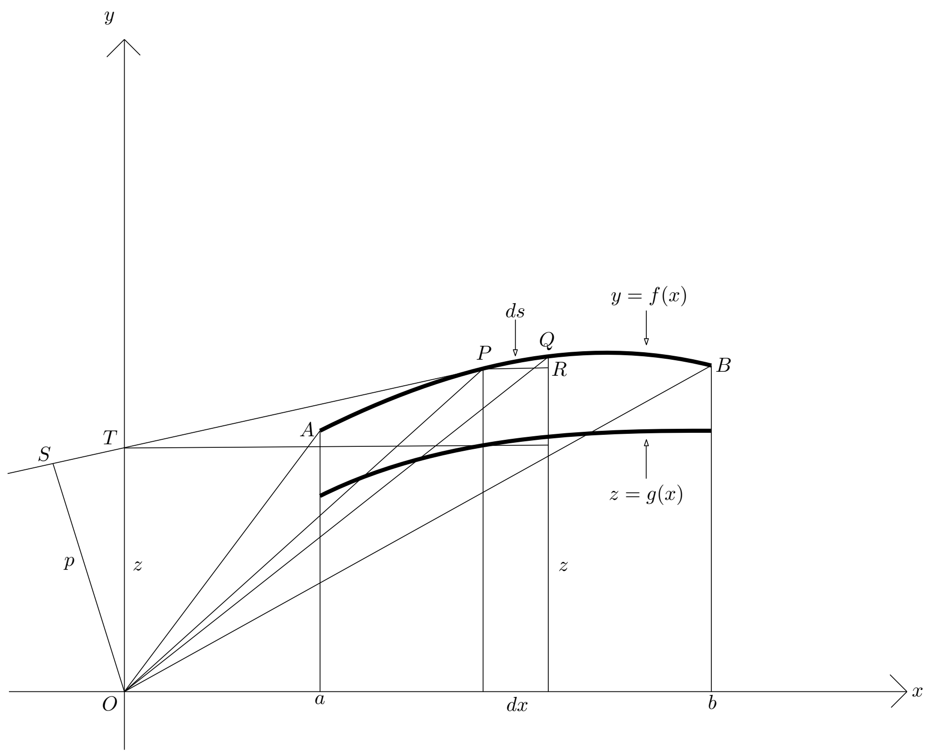


Figure 4.1: Figure for Theorem 4.2

Theorem 4.2 (Leibniz' Transmutation Theorem)**.** *Let $y = f(x)$ for $x \in [a, b]$, and let $z = g(x) = y - x\frac{dy}{dx}$. Then*

$$\int_a^b ydx = \frac{1}{2}\left([xy]_a^b + \int_a^b zdx\right)$$

Proof. Define the points $P = (x, y)$ and $Q = (x + dx, y + dx)$ on the curve $y = f(x)$. Let the tangent to the infinitesmal arc ds intersect the y-axis at $T = (0, z)$, and construct a line OS of length p perpendicular to it, meeting the tangent line in S. Denoting a triangle by Δ, we then have that ΔOST is similar to ΔPQR, since $\angle OST = \angle PRQ$ and $\angle OTS = \angle RQP$, and consequently we get

$$\frac{dx}{p} = \frac{ds}{z} \qquad \Longrightarrow \qquad zdx = pds \tag{4.1}$$

Letting $a(\cdot)$ denote area, we have

$$\begin{aligned}
a(OQS) &= \frac{p}{2}QS \\
a(OPS) &= \frac{p}{2}PS \\
a(OPQ) &= a(OQS) - a(OPS) = \frac{p}{2}QS - \frac{p}{2}PS \\
&= \frac{p}{2}(QS - PS) = \frac{p}{2}ds
\end{aligned}$$

. From (4.1), we then have

$$a(OPQ) = \frac{p}{2}ds = \frac{z}{2}dx$$

Now consider the sector ABO as being subdivided into infinitesimal triangles such as ΔOPQ. We then have

$$a(ABO) = \int_a^b \frac{z}{2}dx = \frac{1}{2}\int_a^b zdx$$

We also have

$$\begin{aligned}
\int_a^b ydx &= \frac{1}{2}\left(bf(b) - af(a)\right) + a(ABO) \\
&= \frac{1}{2}\left[xy\right]_a^b + a(ABO) \\
&= \frac{1}{2}\left(\left[xy\right]_a^b + \int_a^b zdx\right)
\end{aligned}$$

as required. □

One of the more significant instances where Leibniz made use of this theorem was in his derivation of the series

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots \tag{4.2}$$

He started by considering the equation

$$y = \sqrt{2x - x^2}$$

which has derivative

$$\begin{aligned} \frac{dy}{dx} &= \frac{1-x}{y} \\ \implies z &= y - \frac{1-x}{y} = \sqrt{\frac{x}{2-x}} \\ \implies x &= \frac{2z^2}{1+z^2} \end{aligned}$$

We are now ready to apply the transmutation formula from Theorem (4.2):

$$\begin{aligned} \frac{\pi}{4} &= \int_0^1 y dx \\ &= \frac{1}{2}\left(\left[x\sqrt{2x - x^2}\right]_0^1 + \int_0^1 z dx\right) && \text{(Transmutation formula)} \\ &= \frac{1}{2}\left(1 + \left(1 - \int_0^1 x dz\right)\right) && (4.3) \\ &= 1 - \int_0^1 \frac{z^2}{1+z^2} dz \\ &= 1 - \int_0^1 z^2(1 - z^2 + z^4 + \dots) dz && \text{(Geometric series expansion)} \\ &= 1 - \left[\frac{1}{3}z^3 - \frac{1}{5}z^5 + \frac{1}{7}z^7 \dots\right]_0^1 && \text{(Integrating term-wise)} \\ &= 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots \end{aligned}$$

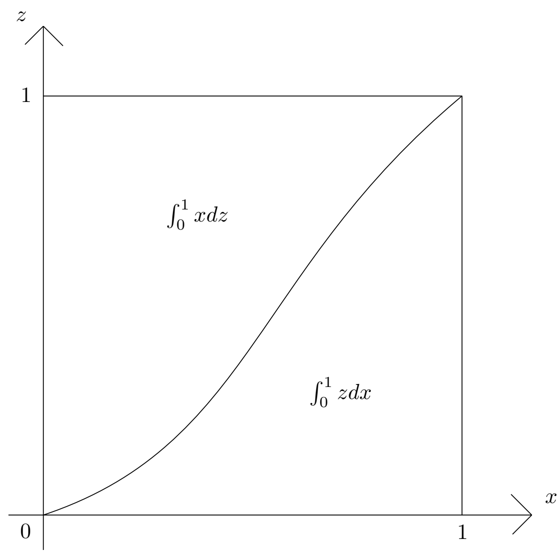


Figure 4.2: Clarifying (4.3)

To make the step in (4.3) clearer, consider Figure 4.2; clearly $\int_0^1 zdx$ is the area of the square, which is 1, minus $\int_0^1 xdz$.

4.2 New Integrals

In the time after Newton and Leibniz brought calculus to the world, integration, through the fundamental theorem of calculus, was genereally seen as anti-differentiation, rather than its own concept [4]. The notion of the integral as the limit of a summation process was certainly something that had been explored, as we have seen in the preceding section, but it was treated more as a convenient way to treat difficult integrals, rather than as a fundamental support for integration theory [4]. Furthermore, area itself was not yet a mathematically defined concept - it was still regarded as something self-evident that did not need clarifying - and integration was only applied to functions that were defined by a single and explicit analytical expression [4]. The need to expand the concept of integration was illustrated by Fourier's series, since its coefficients were determined by integrals which did not fit very well within the contemporary understanding [4].

4.2.1 Cauchy and Continous Integrals on Closed Intervals

Augustin-Louis Cauchy was the first to start to address these problems, by noting the necessity of producing general existence theorems for integrals of classes of functions, and more or less completed the theory behind integrals of continous functions on closed intervals in his *Resume des leçons donnees a l'Ecole Royale Poly technique sur Ie calcul infinitesimal* of 1823 [4].

To do so, Cauchy started with a function $f(x)$ that is continuous on some interval $[x_0, x_f]$. He then divides the interval into n subintervals through the points $x_0, x_1, \ldots, x_{n-1}, x_n = x_f$, and calls this a partition P of the interval. With this partition he then associates an approximating sum S, given by

$$S = \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1}),$$

which is a sum of the rectangles formed from the partition and the curve, with the rectangle in each interval having its height assigned from its left-most point on the curve - see Figure 4.3. Cauchy then wants to define the integral of $f(x)$ over the interval $[x_0, x_f]$ by the limit of S as the maximum of $|x_i - x_{i-1}| \to 0$.

First of all, Cauchy needed to prove that this limit exists. In order to do so, he made use of the following lemma:

Lemma 4.1. *If $\alpha_1, \ldots, \alpha_n$ are positive numbers, $a_1, \ldots, a_n$ are arbitrary numbers, and $\bar{a}$ is some number that lies between the largest and the smallest of*

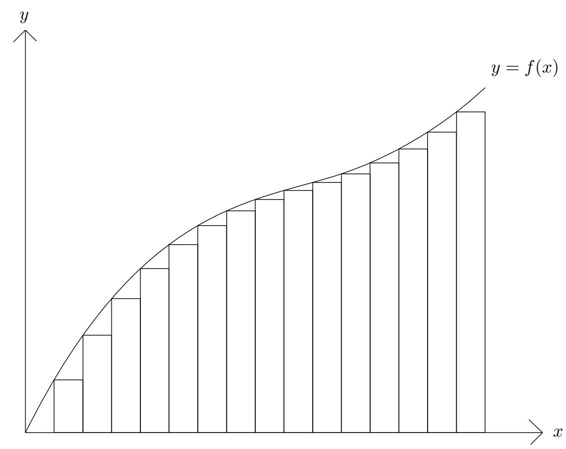


Figure 4.3: Example of Cauchy's approximating sum

$a_1, \ldots, a_n$, *then*

$$\sum_{i=1}^{n} \alpha_i a_i = \bar{a}(\alpha_1 + \ldots + \alpha_n)$$

We then set $\alpha_i = x_i - x_{i-1}$ and $a_i = f(x_{i-1})$, and note that, by the intermediate value theorem (see D.3), $f(x)$ will obtain $\bar{a}$ at some point in the interval, which we will define by $f(x_0 + \zeta(x_f - x_0)$, where $\zeta \in (0, 1)$. Putting all this in our approximating sum S, we have

$$\begin{aligned} S &= \bar{a}(\alpha_1 + \ldots + \alpha_n) \\ &= f(x_0 + \zeta(x_f - x_0))(x_1 - x_0 + x_2 - x_1 + \ldots + x_{n-1} - x_{n-2} + x_n - x_{n-1}) \\ &= f(x_0 + \zeta(x_f - x_0))(x_f - x_0) \end{aligned} \tag{4.4}$$

Now Cauchy considered a refinement of the partition P into P', such that each interval of P' lies in some interval of P. Then the sum for the refined intervals can be written as

$$S' = S'_1 + S'_2 + \ldots + S'_n$$

where S'_i is the sum of the the terms in S' that correspond to the intervals of P' lying in the ith interval of P. Applying (4.4) to these intervals, such that x_0 corresponds to x_{i-1} and x_f to x_i, we have

$$S'_i = f(x_{i-1} + \zeta(x_i - x_{i-1}))(x_i - x_{i-1}) \qquad \zeta_i \in (0, 1), \quad i = 1, \ldots n$$

This in turn gives

$$\begin{aligned} S' &= \sum_{i=1}^{n} S'_i \\ &= \sum_{i=1}^{n} f(x_{i-1} + \zeta_i(x_i - x_{i-1}))(x_i - x_{i-1}) \end{aligned}$$

We then let $\epsilon_i = f(x_{i-1} + \zeta_i(x_i - x_{i-1})) - f(x_{i-1})$, for $i = 1, \ldots, n$, giving

$$S' - S = \sum_{i=1}^{n} \epsilon_i(x_i - x_{i-1}) = \bar{\epsilon}(x_f - x_0),$$

where ϵ lies between the minimum and maximum of $\epsilon_1, \ldots, \epsilon_n$).

Now let P_1 and P_2 be two partitions of $[x_0, x_f]$, let P' be the common refinement formed by collecting the points of subdivision of P_1 and P_2, and let S_1, S_2, and S' be the corresponding approximating sums. Then

$$\begin{aligned} S' - S_1 &= \bar{\epsilon}_1(x_f - x_0) \\ S' - S_2 &= \bar{\epsilon}_2(x_f - x_0) \\ \implies S_1 - S_2 &= (\bar{\epsilon}_2 - \bar{\epsilon}_1)(x_f - x_0) \end{aligned}$$

Here Cauchy concludes that we can make the difference between S_1 and S_2 arbitrarily small by choosing P_1 and P_2 with sufficiently short subintervals. However, this is erroneous given his assumptions, since it only holds if f is uniformly continuous: that is, if given $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(x_a) - f(x_b)| < \epsilon$ for any two points $x_a, x_b \in [x_o, x_f]$ with $|x_a - x_b| < \delta$ [4]. Only if this property holds can the numbers ϵ_i be made arbitrarily small through a partition P with sufficiently small intervals.

Regardless of this overlooked fact, Cauchy goes on to define a definite integral as the limit of the above process, which he has now proven exists. He later goes on to show that the approximation for sums given in (4.4) is also valid for integrals, i.e. that

$$\int_{x_0}^{x_f} f(x)dx = f(x_0 + \zeta(x_f - x_0))(x_f - x_0) = f(\bar{x})(x_f - x_0) \tag{4.5}$$

as a consequence of their interpretation as the limit of a sum, and also that the following two properties hold:

$$\int_{x_0}^{x_f} (af(x) + bg(x))\, dx = a\int_{x_0}^{x_f} f(x)dx + b\int_{x_0}^{x_f} g(x)dx \tag{4.6}$$

$$\int_{x_0}^{x_f} f(x)dx = \int_{x_0}^{\bar{x}} f(x)dx + \int_{\bar{x}}^{x_f} f(x)dx \tag{4.7}$$

Cauchy's final contribution to the theory of integration for continuous functions on bounded intervals is his rigorous derivation of the fundamental theorem of calculus, wherein he uses his arithmetical definition of an integral as a sum to avoid using intuitive notions of area [4].

Given a continous function $f(x)$, defined on $x \in [x_0, x_f]$, Cauchy wanted to prove that $F(x)$ defined on $x \in [x_0, x_f]$ by

$$F(x) = \int_{x_0}^{x} f(t)dt$$

is an antiderivative of $f(x)$ on $[x_0, x_f]$. From (4.7) and (4.5), we then have

$$\begin{aligned}
F(x+\alpha) - F(x) &= \int_{x_0}^{x+\alpha} f(t)dt - \int_{x_0}^{x} f(t)dt \\
&= \int_{x}^{x+\alpha} f(t)dt \\
&= \alpha f(x + \zeta\alpha), \qquad \zeta \in [0,1] \\
\implies \frac{F(x+\alpha) - F(x)}{\alpha} &= f(x+\zeta\alpha) \\
\implies F'(x) = f(x) \iff \frac{d}{dx}\left(\int_{x_0}^{x} f(t)dt\right) &= f(x)
\end{aligned}$$

where the last step follows from the definition of a derivative and the continuity of f. We can also deduce the other familiar form of the theorem via the following: let $\hat{F}(x)$ be such that $\hat{F}'(x) = f(x)$ on $[x_0, x_f]$. If

$$\begin{aligned}
\omega(x) &= F(x) - \hat{F}(x) \\
\text{then } \omega'(x) &= F'(x) - \hat{F}'(x) = f(x) - f(x) = 0 \\
\implies \omega(x) &= \omega(x_0) + (x - x_0)\omega'(\bar{x}) = \omega(x_0) \qquad \text{(Mean Value Theorem)}
\end{aligned}$$

for all $x \in [x_0, x_f]$. We then have

$$\begin{aligned}
F(x) - \hat{F}(x) &= F(x_0) - \hat{F}(x_0) = -\hat{F}(x_0) \qquad (F(x_0) = 0) \\
F(x) &= \hat{F}(x) - \hat{F}(x_0) \\
\int_{x_0}^{x} f(x)dx &= \hat{F}(x_f) - \hat{F}(x_0)
\end{aligned}$$

which holds for any antiderivative $\hat{F}(x)$ of $f(x)$.

4.2.2 Riemann et al. Extends Cauchy's Concepts

Cauchy had almost completely defined the theory of integration for continuous functions on bounded intervals, and his work could easily be extended to piecewise continuous functions by dividing up the integral into many smaller ones through (4.7), where the function is continuous on each smaller interval. However, the problem of integrating a discontinous function had yet to be approached in a rigorous manner, likely due to its more esoteric nature and lack of direct physical analogies. It was Georg Friedrich Bernhard Riemann who first tackled the task of generalizing Cauchy's work, motivated by a theorem by Dirichlet which required piecewise continuity to hold specifically since it made use of integration [4].

Riemann's generalization was to use the arbitrary point $\bar{x}_i = x_{i-1} + \epsilon_i \delta_i$, rather than x_{i-1}, to define the height of the ith subinterval of the partition in Cauchy's approximate sum, so that his integral is given by

$$\int_a^b f(x)dx = \lim_{\delta \to 0} \sum_{i=1}^{n} f(\bar{x}_i)(x_i - x_{i-1}), \qquad \delta = \max_i \delta_i$$

where $\epsilon_i \in [0, 1]$, and $\delta_i = x_i - x_{i-1}$, $i = 1, \ldots, n$.

In order to investigate the properties of this new definition of the integral, he considers a bounded function f and a partition P of $[a, b]$. Letting $\delta_i = x_i - x_{i-1}$ and

$$D_i = \max_{x_{i-1} < x < x_i} f(x) - \min_{x_{i-1} < x < x_i} f(x),$$

that is, the difference between the maximum and minimum value of $f(x)$ in the ith interval, Riemann defines the total oscillation with respect to P as

$$D(P) = \sum_{i=1}^{n} D_i \delta_i,$$

and concludes that the integral exists if and only if

$$D(P) \to 0 \text{ as } \delta \to 0,$$

where δ is the norm of the δ_i. He here overlooks, or omits, the need for completeness of the real numbers for this to hold true [4].

Moving on, Riemann defines $\Delta = \Delta(d)$ to be the maximum value of $D(P)$ for all partitions P with $\delta \leq d$, implying that $\Delta(d)$ is a decreasing function of d and that f is integrable on $[a, b]$ if and only if

$$\lim_{d \to 0} \Delta(d) = 0.$$

Then, given some $\sigma > 0$ and some partition P, Riemann defines $s = S(\sigma, P)$ as the sum of the lengths δ_i of the subintervals in P where $D_i > \sigma$. Armed with all the above, he establishes the following existence theorem:

Theorem 4.3 (Existence of Integral of a Bounded Function)**.** *If $f(x)$ is bounded for $x \in [a, b]$, then $\int_a^b f(x)dx$ exists if and only if, given $\sigma > 0$ and $\epsilon > 0$, there exists a $d > 0$ such that for any partition P with norm $\delta < d$, the sum s of the lengths of the subintervals i of P where the oscillation D_i of $f(x)$ is greater than σ, is less than ϵ.*

Proof. Necessity: If $\delta < d$, then $D_i > \sigma$ on a some set of subintervals having total length s. Thus

$$\sigma s < \sum_{i=1}^{n} D_i \delta_i \leq \Delta(d),$$

and $s(\sigma, P) < \frac{\Delta(d)}{\sigma}$. Assuming that the integral exists, this, for some fixed σ, approaches 0 as d approaches 0.

Sufficiency: Let some $\sigma > 0$ and some $\epsilon > 0$ be given, and choose $d > 0$ as in the case for necessity above. If $\delta < d$, then the subintervals of $f(x)$ wherein the oscillations D_i of $f(x)$ is greater than σ will contribute an amount less than $D\epsilon$, where D is the oscillation of $f(x)$ on $[a, b]$, to $D(P)$, since their lengths add up to s which is less than ϵ. The other subintervals will contribute at most $\sigma(b-a)$ to $D(P)$, giving

$$D(P) < D\epsilon + \sigma(b-a).$$

We can then take ϵ and σ sufficiently small so as to make $D(P)$ as small as we like, satisfying the requirement that $D(P) \to 0$ as $\delta \to 0$ and proving that the integral exists. □

In the latter third of the 19[th] century, several authors independently introduced what is now referred to as upper and lower Riemann sums [4]. These were given, for some bounded function $f(x)$ defined on an interval $[a, b]$ with an associated partition P with n subintervals, by

$$U(P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}),$$
$$L(P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}),$$

where M_i is the maximum and m_i the minimum values of $f(x)$ on the ith interval $[x_{i-1}, x_i]$. An example of these sums is given in Figure 4.4. In the framework of the above discussion on similar sums, we state that these sums approach limits U and L, respectively, as $\delta \to 0$, where δ is the norm of P, and note that this is the case regardless of whether or not $f(x)$ is integrable.

The notation was modifed by Vito Volterra, who introduced upper and lower

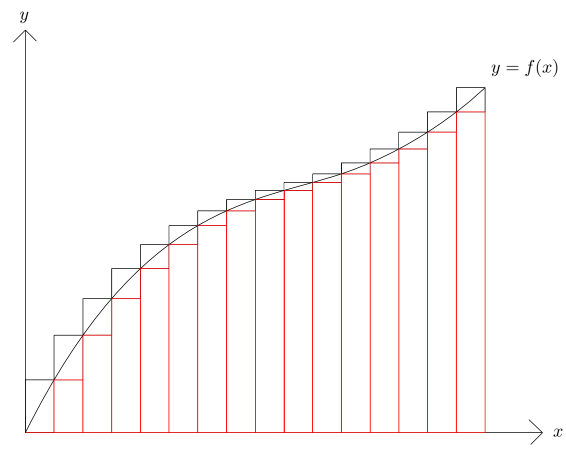


Figure 4.4: Example of upper (black) and lower (red) Riemann sums

integrals given by

$$U = \overline{\int_a^b} f(x)dx,$$
$$L = \underline{\int_a^b} f(x)dx,$$

and Giuseppe Peano noted that these upper and lower integrals could be defined as the greatest lower and least upper bounds for all partitions P of $[a, b]$ of the upper and lower Riemann sums, respectively [4]. These developments allowed for a new interpretation of the existence of an integral via the statement that $f(x)$ is integrable if and only if its upper and lower integrals are equal.

4.3 The Definition of Area

Throughout this paper, we have made many mentions of the close relationship between area and integration. However, it wasn't until late in the 19$^{\text{th}}$ century that the concept of area was actually given a mathematical definition by Peano, which finally allowed these comparisons to become rigorous [4].

Peano took the method of exhaustion as his starting point, adopting its style of defining an inner and outer area and finding the true area as the meeting point between the two. Starting with a figure S, he defined its inner area $a_i(S)$ as the least upper bound of the areas of all the polygons contained in S, and similarly defined the outer area $a_o(S)$ as the greatest lower bound of the areas of all the polygons containing S [4]. That the inner area is always less than or equal to the outer area is clear, but it could be that they are unequal - for example if S is the set of all points located on the unit square such that both its x- and y-coordinates are irrational. Then the outer area is 1, the area of the square, but the inner area is 0, since it contains only degenerate polygons.

Having defined area in the above manner, the lower and upper Riemann integrals naturally fit into the context of these concepts, which Peano noted by establishing that the upper integral is equal to the outer area, and the lower integral the inner area [4]. If the outer and inner area are equal, or equivalently, if the upper and lower integrals are equal, we then speak of *the* area and *the* integral. We have thus, finally, come full circle - we started with intuitive area concepts and created semi-rigorous procedures similar to integration, and after having defined integration through anti-differentiation we have now linked it back to the original area considerations.

Significant developments in the theory of integration were still being made after Peano, most notably by Jordan's generalization of the area concept to work in any dimension and Lebesgue's extension of the theory of integration, and the field continues to be advanced even today. The purpose of this paper, however, has been to illuminate the origins of integration, from its roots to its early modern form, and to this end this seems as fitting a place as any to stop.

Chapter 5

Summary

The story of the integral is rich and varied, and could never be told in its complete and unabridged form within the narrow confines of this paper. However, this text has hopefully served to introduce and explain some of the most important advancements made in the history of integration: the early method of exhaustion and its polygons, the method of indivisibles and its attempt to extend its predecessor to more complicated quadratures, and finally the calculus - this most incredible achievement - which brought a powerful new tool to the world that has helped shape mathematical development ever since.

Despite the fact that the methods described above differ greatly in their presentation, the core ideas at play remain the same - untouched, albeit perhaps overshadowed, by the increasingly advanced and complex concepts. For when Peano gave us area in the notion of a border, an interface between a least upper and greatest lower bound, we were given a link to the upper and lower Riemann integrals, which themselves grew from Riemann's own summation of rectangles - simple geometrical shapes giving rise to a more complicated one. Riemann's, and Cauchy's, sums were not the first of their kind, but owed to the work of

Leibniz' triangular indivisibles, which were in turn a generalization of the work Cavalieri had done with his planar intersections and collections of lines. Yet Cavalieri's motivation was a generalization of the method of exhaustion - the very method that started it all - after all, is the comparison of an upper and lower integral - the comparison of an upper and lower sum - not simply another polygonal construction, similar to those inscribed and circumscribed regular polygons used by the Ancient Greeks in their works on circles?

Thus we see that when we strip away the layers of terminology and abstraction in the theory of integration, what we arrive at is that most basic idea, that which persists throughout all human endeavours, of using a simple approximation to say something about a complicated situation.

Bibliography

[1] K. Andersen. “Cavalieri’s Method of Indivisibles”. In: *Archive for History of Exact Sciences* 31.4 (1985), pp. 291–367. ISSN: 00039519, 14320657. URL: `http://www.jstor.org/stable/41133738`.

[2] J. Arndt and C. Haenel. *Pi - Unleashed*. Springer, 2001. ISBN: 978-3540665724.

[3] Brilliant.org. *Componendo and Dividendo*. URL: `https://brilliant.org/wiki/componendo-and-dividendo/`.

[4] C. H. Edwards Jr. *The Historical Development of the Calculus*. Springer-Verlag, 1979.

[5] T. L. Heath. *The Works of Archimedes*. Cambridge University Press, 1897. URL: `https://archive.org/details/worksofarchimede00arch`.

[6] T.L. Heath. *The Thirteen Books of Euclid’s Elements*. Vol. 1. Cambridge University Press, 1908. URL: `http://www.wilbourhall.org/pdfs/heath_euclid_i.pdf`.

[7] T.L. Heath. *The Thirteen Books of Euclid’s Elements*. Vol. 2. Cambridge University Press, 1908. URL: `http://www.wilbourhall.org/pdfs/heath_euclid_ii.pdf`.

[8] T.L. Heath. *The Thirteen Books of Euclid's Elements*. Vol. 3. Cambridge University Press, 1908. URL: http://www.wilbourhall.org/pdfs/heath_euclid_iii.pdf.

[9] F. Lemmermayer. *Exhaustion: From Eudoxus to Archimedes*. URL: http://www.fen.bilkent.edu.tr/~franz/M300/exh.pdf.

[10] J. R. Newman. *The World of Mathematics*. Vol. 1. Simon and Schuster, 1956.

[11] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill Book Company, 1976.

Appendix A

Euclid

The following results are found in the appropriate volume of Heath's The Thirteen Books of Euclid's Elements*, unless stated otherwise.*

Proposition A.1 (Euclid I. 26.)**.** *"If two triangles have the two angles equal to two angles respectively, and one side equal to one side, namely, either the side adjoining the equal angles, or that subtending one of the equal angles, they will also have the remaining sides equal to the remaining sides and the remaining angle to the remaining angle."*

Proposition A.2 (Euclid V. 5)**.** *"If a magnitude be the same multiple of a magnitude that a part subtracted is of a part subtracted, the remainder will also be the same multiple of the remainder that the whole is of the whole."*

Proposition A.3 (Euclid V. 11.)**.** *"Ratios which are the same with the same ratio are also the same with one another". In other words, if* $\frac{a}{b} = \frac{c}{d}$ *and* $\frac{c}{d} = \frac{e}{f}$ *then* $\frac{a}{b} = \frac{e}{f}$*.*

Proposition A.4 (Euclid V. 16.)**.** *"If four magnitudes be proportional, they will also be proportional alternately". That is, if* $\frac{a}{b} = \frac{c}{d}$*, then* $\frac{a}{c} = \frac{b}{d}$*.*

Proposition A.5 (Euclid VI. 3.)**.** *"If an angle of a triangle be bisected and the straight line cutting the angle cut the base also, the segments of the base will have the same ratio as the remaining sides of the triangle; and, if the segments of the base have the same ratio as the remaining sides of the triangle, the straight line joined from the vertex to the point of section will bisect the angle of the triangle."*

Proposition A.6 (Euclid X. 1.)**.** *"Two unequal magnitudes being set out, if from the greater there be subtracted a magnitude greater than its half, and from that which is left a magnitude greater than its half, and if this process is repeated continually, there will be left some magnitude which will be less than the lesser magnitude set out".*

Proposition A.7 (Euclid XII. 1.)**.** *"Similar polygons inscribed in circles are to one another as the squares on the diameters".*

Appendix B

Archimedes

The following results are found in Heaths' The Works of Archimedes, *unless stated otherwise.*

Proposition B.1. *The area of any circle is equal to the right-angled triangle in which one of the sides about the right angle is equal to the radius, and the other to the circumference, of the circle.*

Proposition B.2. *The area of a circle is to the square on its diameter as* 11 *to* 14.

Proposition B.3. *The ratio of the circumference of any circle to its diameter is less than* $3 + \frac{1}{7}$ *but greater than* $3 + \frac{10}{71}$.

Appendix C

Cavalieri

The following results can be found in Andersen's Cavalieri's Method of Indivisibles, *unless stated otherwise.*

Definition C.1 (Book VII: Tangent)**.** *I say that a straight line touches a curve situated in the same plane as the line when it meets the curve either in a point or along a line and when the curve is either completely to the one side of the meeting line [in the case when the meeting line is a point] or has no parts to the other side of it [in the case when the meeting is a line segment].*

Note: the above definition is distinct from the common modern usage in that, for example, it admits a line touching the apex of a triangle as a valid tangent to the triangle.

Theorem C.1. *Suppose a closed plane figure is given along with a direction, which we call a* regula*. Then there exists two tangents to the figure which are parallel to the regula. Furthermore, any lines parallel to the regula situated between the two tangents will intersect the figure in line segments, and any lines parallel to the regula situated outside the two tangents will have no points of intersection with the figure.*

Appendix D

Various

Theorem D.1 (Componendo and Dividendo Theorems [3]). *Let a, b, c, and d be numbers such that b and d are non-zero and $\frac{a}{b} = \frac{c}{d}$. Then the following holds:*

1. $\dfrac{a+b}{b} = \dfrac{c+d}{d}$ (Componendo)
2. $\dfrac{a-b}{b} = \dfrac{c-d}{d}$ (Dividendo)
3. *For* $k \neq \dfrac{a}{b}$, $\dfrac{a+kb}{a-kb} = \dfrac{c+kd}{c-kd}$
4. *For* $k \neq \dfrac{-b}{d}$, $\dfrac{a}{b} = \dfrac{a+kc}{b+kd}$

Theorem D.2 (Inscribed Angle Theorem). *Let a circle be defined with center O, and let some points A, B, C be placed on the circumference. Then the angle ABC is half the angle AOC. The special case where A and C form a diameter of the circle is called Thales' Theorem.*

Theorem D.3 (Intermediate Value Theorem [11]). *Let f be a continuous real function on the interval $[a, b]$. If $f(a) < f(b)$ and if c is a number such that $f(a) < c < f(b)$, then there exists a point $x \in (a, b)$ such that $f(x) = c$.*