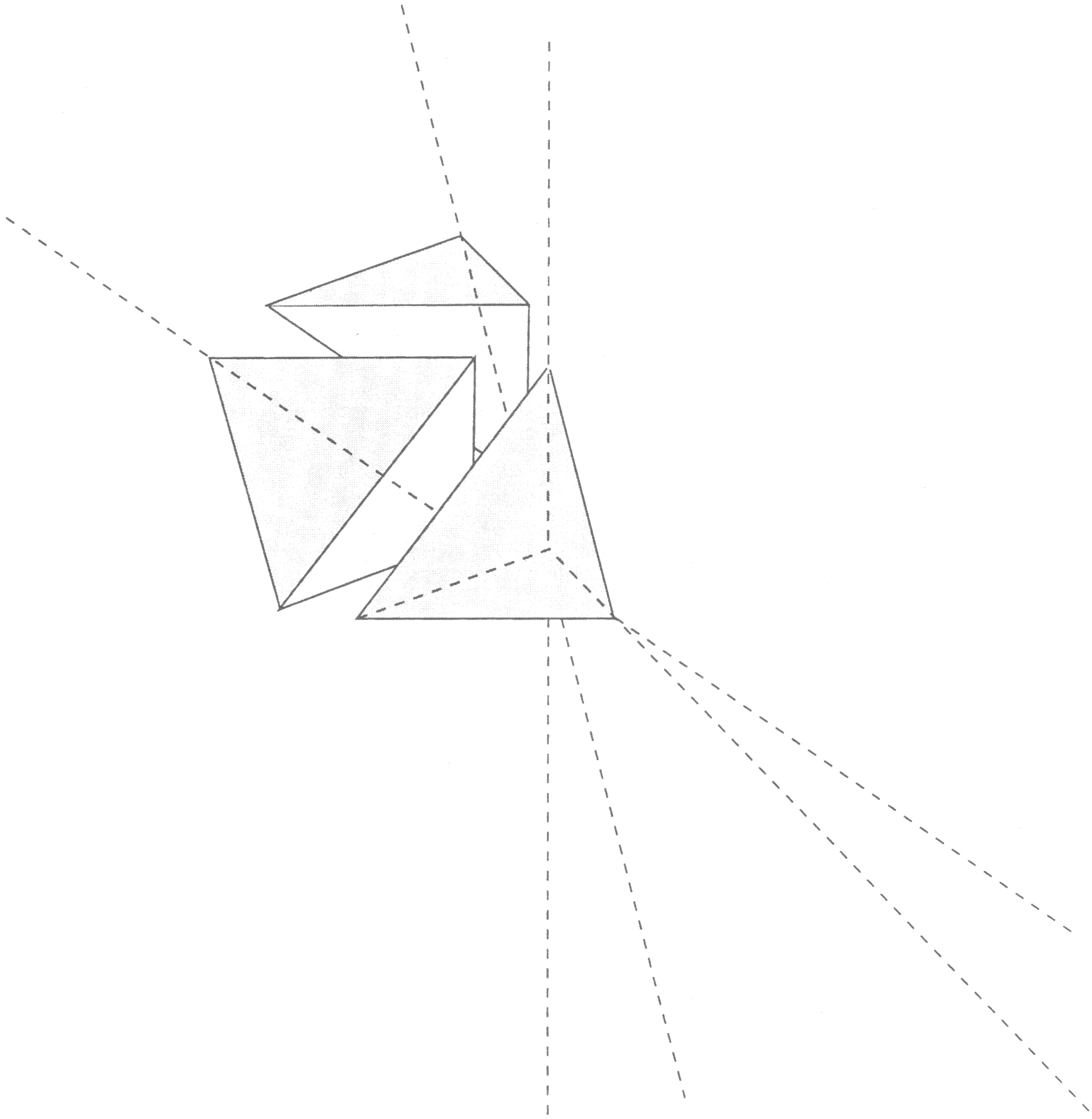
THREADS IN TIME

Historical Connections to School Mathematics



K. Subramaniam

P. K. Sahoo

Technical Report No. 45

Contents

Preface

When we see young children going to school, little do we realize that the bundle on their backs has taken shape over long periods in history. We take so much of the tools of elementary mathematics for granted. Children learn to say the numbers up to hundred even before they enter school. The basic arithmetic operations are taught to primary school children in predigested capsules and we grow angry with them if they fail to learn such elementary things. Yet humankind struggled with the notation for numbers for millennia before they could perform the arithmetic operations with relative ease. Some of the formulas that children in high school routinely use, were great challenges to the best mathematicians across the civilizations of the world, in the West and in the East.

This booklet attempts to follow some of the threads in history that connect to school mathematics. It is difficult to unravel the innumerable threads that have contributed to the weaving of a well-knit fabric. What we have attempted is to trace the central connections and mention the work of some main contributors. We have also tried, at various places in the booklet, to fill out the logical context of a problem so that its history is appreciated better. A few of the brilliant insights and arguments of the past have been elaborated in some detail. Many historical expositions go over these too briefly to allow the reader to appreciate their rich intellectual content. We have chosen to elaborate in particular those topics which are central to school geometry, but whose fascinating history will be completely lost to school students once they begin to solve these problems from the viewpoint of calculus.

Insight and excellence in mathematics are not the prerogative of any single culture. Indeed, as we learn more about the past of various civilizations, we are struck by the deep similarities across civilizations. In this booklet too we hope to communicate this idea. Some important threads in the history of mathematics meander over many regions of the globe and we have attempted to follow at least some of them. It may be too much to expect this small booklet to create an interest and a taste for mathematics, but we do hope that this will happen once in a while with some of the readers.

This booklet was written during P.K. Sahoo's stay at HBCSE as a teacher fellow between February and June 1999. The authors thank Venola Fernando for her extensive help with preparing the booklet. She has drawn most of the beautiful figures in the booklet. She also helped with the typing, typesetting and proof-reading.

We also wish to thank V.V. Murthy, teacher fellow, especially for help with making a model of Archimedes' derivation of the volume of a sphere. He also participated in the presentation sessions during the preparation of the booklet and gave valuable support throughout the period of the project.

V.N. Purohit typed a major portion of the manuscript. We are thankful to her. We admire the speed and accuracy of her work, despite her unfamiliarity with LaTeX. N.S. Thigale provided reliable and speedy support at the finishing point as usual and we thank him.

We thank Meena, Garima and Nitin for helping with computer related jobs. We also take the opportunity to thank Dr. Nagarjuna and R.S. Patwardhan for the excellent and sophisticated computer network facilities at the centre which made our job easier and allowed us to keep to deadlines.

We thank our colleagues at HBCSE for help and support in many ways. Several of them attended an initial presentation by P.K. Sahoo and provided useful feedback.

We wish to thank AEES and HBCSE for the teacher fellowship offered to P.K. Sahoo at HBCSE. P.K. Sahoo also wishes to acknowledge the encouragment of Mr. A.K. Hazra, Principal of AECS (Eng. med.), Jaduguda. We thank Dr. H.C. Pradhan and Dr. Arvind Kumar for their encouragement and enthusiastic support as always. Finally, we wish to thank the members of our respective families for their unflinching support.

K. Subramaniam

HBCSE, Mumbai **P. K. Sahoo**

1 Numbers and numerals in history

The development and use of a number system is the first step in developing mathematics. When was this step taken? Who took this step? These are difficult questions to which we can find only indirect answers. As far as we can make out humans used numbers probably tens of thousands of years ago. The Ishango bone found in Africa, which shows an interesting pattern of tally marks on it, is as old as 20,000 BC. The grouping of the marks shows that the humans who made them were aware of patterns in addition. If this knowledge was so much earlier than the rise of civilizations, does it imply that numbers are a part of our biology, part of the way our brains are wired up? If our number sense is an outcome of our brain structure then more questions arise. For we are not the only species with brains, or even large brains. Bottle-nosed dolphins, for example, have brains of a very impressive size, about 14% larger than the human brain. So we now begin to wonder, do animals too understand numbers? Do they have a number sense?

We have already raised many difficult questions. Let us tackle them one by one. First let us ask whether animals really understand what numbers are. Notice that we are not asking whether animals can recognize what is meant by the symbol '5'. If you showed the symbol to Aristotle, who by all accounts was quite intelligent, he would be at a complete loss. The symbol would be entirely unfamiliar to him, but he certainly would understand numbers. So to avoid confusion we usually call the symbol '5' a **numeral**. Numerals are signs, usually written signs, which stand for numbers. Over time different cultures have evolved different ways of showing or representing numbers. The modern system for representing numbers originated in India more than 1200 years ago. Spoken number words like written symbols are also signs. *Paanch* in Hindi is a word which sounds very different from 'five', but refers to the same number. The number itself is an abstract entity and is different from the sign, spoken or written. One way to define number is to say that it is a property of a collection of things, it says how many things there are in the collection.

The distinction between numerals and numbers is a clear one. However we will use the term 'number system' in this book with a little more freedom. When we say 'number system', it will mean systems that humans have evolved to represent numbers. They could be spoken words or written symbols. The term 'numeral system' will have a narrower meaning – we will use it only to refer the written symbols which represent numbers.

Do animals understand numbers?

Do animals then understand numbers? There have been reports from time to time of clever animals that solve sums in arithmetic. All such claims so far have turned out to be empty claims. One well-known example is from Germany at the turn of the 20th Century. A horse called 'Clever Hans' could, it was claimed, solve sums written and presented to it on a blackboard. The horse would look at the sum and tap out the correct answer with his hoof. If the correct answer was 12, he would tap exactly twelve times. When someone did careful trials it turned out that the horse was actually reading unconscious signals from the master. When the master did not see the sum Clever Hans failed to get the answer. The horse was indeed clever not because he could do arithmetic but because he could sense signals from his master, which were probably both unintentional and unconscious.

There have been some successful attempts to teach animals to identify simple numbers. Several experiments have tested animals for their ability to identify how many objects there are in a collection and have come up with surprising findings. One of the most recent of these studies reports experiments conducted with rhesus monkeys. These monkeys were trained to arrange sets of objects in ascending order for small numbers (up to 5). This itself is remarkable because the monkeys could recognize the **numerosity** or 'how-many-ness' of the collection even if the objects in a collection were different. The monkeys could also identify the numerosity of collections containing strange objects which they had not seen before, and different from the ones presented to them earlier. But the monkeys did better. They were able to generalize the ability that they had picked up during training to bigger sets of objects containing up to 9 objects. This is a clear indication that animals are able to distinguish, identify and understand the numerosity of a collection.

Do humans understand numbers in a way similar to animals? Studies from a new and exciting branch of science, neuroscience, throw light on this question. As it turns out the human sense of numbers is very different and very much more complex than what animals can manage. Cognitive neuroscience, which is a branch of neuroscience, studies how the human mind and its capacities and functions are related to the human brain. One of the tools it uses to explore the mind-brain relationship is the careful study of patients who have suffered brain damage. Such studies have revealed how different parts of the brain have very specialized abilities. It appears, for instance, that reading and recognizing a number, writing a number, comparing numbers, having an idea of how big a number is, adding and subtracting small numbers, doing vertical addition on paper all are specialized functions managed by different parts of the brain. We know that different parts of the brain are involved because sometimes one of these abilities is impaired leaving the others intact.

Stanislas Dahaene, a mathematician turned neuroscientist who has studied several patients with brain damage, reports an almost unbelievable loss of capability in one of his patients, Mr. M. This patient could read numbers perfectly well and could write them too without any difficulty. He could also solve simple addition sums, but was totally unable to subtract or say which one of two single digit numbers was the larger. For example, Mr. M during one of his conversations with Dahaene told him that 5 was larger than 6. When asked for a number between 3 and 5, he suggested 3 and then 2 – complete nonsense. Between 10 and 20, he placed 30 and then corrected his answer to 25 saying 'I do not visualize numbers very well.' It seems almost paradoxical then that Mr. M was able to understand dates such as 1789 or 1815 perfectly well. He not only understood these dates but could lecture for hours on the historical events that took place in these years. (Mr. M was French; 1789 is the year of the French Revolution and 1815 is the year in which the major modern European states formed themselves.)

People like Mr. M are rare because they suffer only limited brain damage to specific areas in the brain. So only a very specialized function is disabled while leaving other functions intact. In most cases of brain damage a whole set of functions, probably linked to adjacent areas in the brain are lost. However the few cases of specific brain damage give us enormous insight into how a normal brain functions. The ability that we call number sense, and indeed higher mathematical abilities, call for the coordination of so many different specialized regions of the brain. Our understanding of numbers is a complex ability built up out of a number of skills, many of which are unconscious and automatic processes in the brain. So one cannot give a brief or simple definition of number sense in humans.

This tells us that the understanding of numbers that animals are capable of developing and human number sense are vastly different. Animals probably manage to acquire only a few specific abilities which form part of the complex human ability to understand numbers. Understanding the numerosity of small collections, identifying which collections are larger are processes which can be localized to areas of the brain's right hemisphere. The right hemisphere, it appears, is able to do approximate simple calculations, but is unable to do exact calculations. Language ability, the ability to comprehend written and spoken language, are located in the left brain. Clearly human number sense is intimately tied to the human ability to understand symbols and to manipulate them. Apparently human number sense involves functions which are similar to or rely upon our ability to develop and understand language.

When did humans develop number systems?

Despite the capacities of human brains to develop a number sense, much of this number sense is culturally acquired. It is not an intrinsic ability wired into our brains like perhaps our ability to recognize different faces or voices. Even language appears to be more intrinsic and was developed in nearly all its complexity much before any significant development of our number sense. A well developed and sophisticated language is universal among all human communities and probably played a crucial role in the evolution of the human species. In contrast, studies of many different tribes show that it is not necessary for humans to use numbers or develop mathematics in order to survive. The environments in which many tribes live and upon which they depend for their food, and the lifestyles and social organization that they have evolved, do not require the development of number systems. There are many tribes which till recently had no number words other than one, two and many, sometimes even one and many. Numbers, numeral systems and mathematics arose after humans had developed settled agriculture and established civilizations.

This of course presents a puzzle from a modern biological point of view. The capacity of the human mind or brain to develop mathematics is obvious and undeniable – one only has to look at the magnificent edifice of modern mathematics. Yet the capacity did not develop because it conferred an evolutionary advantage to humans. How then do we explain the capacity of the human brain to develop number systems and mathematics? One explanation is that the human brain is not designed (or has not evolved structures) specifically for mathematics. Rather structures in the brain that have evolved for other functions were taken over for learning mathematics once those functions became redundant as societies progressed. However these are still hypotheses awaiting confirmation in different fields of science.

Our interest here is mainly in how number and numeral systems first developed and what were the major stages in their development. There are three contexts which have provided fertile ground for the growth of mathematics in the Ancient world. These are time-keeping, which is also linked to astronomy, commerce, and ritual connected with religion or magic. The first written numeral system developed in Babylon (Mesopotamia) and Egypt in the contexts of commerce and religious ritual. The Babylonian and Egyptian numeral systems developed fairly independent of each other. In fact, the development of numeral systems provides excellent examples of how human thought proceeds along similar lines when circumstances are similar. The place value notation for numbers, which we consider to be a great step in human intellectual development, took place independently at least thrice. **Positional number systems** use the same sign to show different values depending on the place. Our own number system is positional. When we write '333', there are three '3's and all of them have different values (3, 30 and 300) since they occur in different positions.

The first positional number system developed in Babylon in the third millennium BC. The second is the Chinese rod numeral system which was used by them at least from the second century BC. This system is so different from the Babylonian that clearly the Chinese must have evolved the positional notation independently. The third positional system for numerals was developed by the Mayan civilization which reached a high point between the third and tenth centuries AD. Again we have evidence of independent development because there was virtually no contact between the Mayan civilization in Central America and the 'Old World' civilizations of Africa, Asia and Europe till after the time of Columbus. Despite this there is a striking similarity in the numerals used by the Mayans and the Chinese. It shows that human minds can strike upon the same ideas through quite independent paths.

Do numbers first appear in spoken language or in written language?

There are many tribes that have developed systems of number words but have no written language. These number word systems are based on different principles. As we mentioned earlier, some tribes have very few number words, only for one and two. It is of course possible to use these number words to indicate bigger numbers. An Australian aboriginal tribe, the Gumulgal, have only two basic number words which are repeated to form larger numbers. Table 1.1 shows some 2-count number systems. Two of them have only two basic number words, while the third is slightly modified, but is essentially a 2-count system.

This system probably reminds you of binary numbers. Indeed it is a numeral system which uses 2 as the base. There are two distinct number words to indicate

Number	Gumugal (Australia)	Bakairi (Central Brazil)	Bushman (Southern Africa)
1	urapon	tokale	xa
2	ukasar	ahage	t'oa
3	ukasar-urapon	ahage tokale	'quo
4	ukasar-ukasar	ahage ahage	t'oa t'oa

Table 1.1: Examples of counts with base 2

Number	Aboriginal tribe (Australia) 3-count	Toba (Paraguay) 4-count
1	mal	nathedac
2	bularr	cacayni or nivoca
3	guliba	cacaynilia
4	bularr-bularr	nalotapegat
5	bularr-guliba	nivoca-cacaynilia
6	guliba-guliba	cacayni-cacaynilia

Table 1.2: Examples of counts with bases 3 and 4

1 and 2. Bigger numbers are formed through combinations of these words. Of course, the number words become more cumbersome as the numbers get bigger and so a system like this is severely limited. There are also systems which use 3 and 4 as bases of which some examples are in Table 1.2 . A system which uses 4 as base would have four independent number words to indicate the numbers 1, 2, 3 and 4. The number 5 would be spoken as 4 and 1 (or as 3 and 2 in the table), the number 6 as 4 and 2 (or as twice three in the table) and so on.

The use of 5 as base is more common than these relatively rare bases of 2, 3 and 4. But the most widespread base was the number 10, probably because it is natural to associate numbers with the fingers of the hand. Many major number systems of both the Old World and the Americas were decimal, that is, they adopted the base 10; the Egyptian, the Indian, the Chinese and the Inca numeral systems are examples. Another common base was 20, again linked to the fingers and toes. The Mayan numeral system was vigesimal (base 20) and was also a system based on place value. Other societies which used base 20 numbers were the Yorubas in West

1	ookan	6	eefa	11	ookanlaa	16	eerindinlogun
2	eeji	7	eeje	12	eejilaa	17	eetadinlogun
3	eeta	8	eejo	13	eetalaa	18	eejidinlogun
4	eerin	9	eesan	14	eerinlaa	19	ookandinlogun
5	aarun	10	eewaa	15	aarundinlogun	20	ogun
						21	ookanlelogun

Table 1.3: The Yoruba counting system with base 20. Notice the sub-base of 10

Africa and the Aztecs in Mexico. Many vigesimal number systems did not have independent words or signs for all the first 20 numbers. Up to 20 the words could be formed using base 5 or base 10. But for bigger numbers the two bases would be combined. An example is shown in Table 1.3 .

The sexagesimal system which has 60 as base is also an important system although the only well-developed example of such a system is the Babylonian. It would be surprising if the sexagesimal system had 60 independent signs or words for the first 60 numbers. That would have made the number words difficult to memorize. Rather the base 10 is used up to the first 60 numbers. Then 60 becomes the base. Moreover the system is positional, so the written symbol for 60 is the same as the symbol for 1. We will see more details of this system soon.

The sexagesimal system, formed the basis for the very impressive computational skill of the Babylonians. The base 60 system has advantages in representing fractions as sexagesimal numbers. Sexagesimal numbers are similar to our decimal numbers and are in fact superior to them in some ways. This was very useful in obtaining the accuracies that were needed for astronomical calculations. Ptolemy who is the greatest of the Greek astronomers used sexagesimals for all astronomical computations in his writings and used the more common decimal system while presenting numbers for better comprehension. Ptolemy lived in Alexandria in Egypt. Alexandria was the capital of a later Greek civilization, established after Alexander's conquests. This period saw great developments in mathematics and science and is referred to as the Hellenistic period to distinguish it from Classical Greece. Ptolemy, like much of Alexandrian science, represents a fertile amalgamation of Greek with other older traditions.

We have mentioned examples of several number systems with different bases some of which were written numeral systems and some of which were only spoken.

Clearly written numeral systems came much later than spoken number systems. Sophisticated developments in using numbers could only arise on the basis of a written numeral system. But this does not mean that spoken number systems did not show any sophistication at all. The Yoruba system stands as an example of how a refined number sense can be incorporated within a purely spoken system.

The Yoruba number system, which is a base 20 system, uses a sub-base of 10. There are distinct words for the numbers from 1 to 10. If you look at Table 1.3 you will be able to make out that the numbers from 10 to 14 are formed on the pattern of '1 more than 10', '2 more than 10', and so on. However the numbers from 15 to 20 use the subtraction principle, they are in the pattern '20 less 5', '20 less 4' and so on. The system becomes more complicated as the numbers proceed beyond 20. Here are examples of how some number words are formed.

$$\begin{aligned} 45 &= (3 \times 20) - 10 - 5 \\ 50 &= (3 \times 20) - 10 \\ 300 &= 20 \times (20 - 5) \\ 525 &= (200 \times 3) - (20 \times 4) + 5 \end{aligned}$$

This complicated system is probably related to the common practice among the Yorubas of using cowrie shells for counting. The cowrie shells would be gathered into convenient sized piles of 5, 20 or 200, reflecting the structure of the number words. The Yorubas were able to perform a fair amount of computation with their number words. Even expressing numbers properly required a sound number sense and number manipulation skills. However the system makes computation with larger numbers very cumbersome besides being difficult to learn.

When and how did written numerals originate?

The first written numerals were simple tally marks and these are more than 30,000 years old. Indeed the more sophisticated numeral systems first arose as modifications of the simple tally marks. Tally sticks were usually of wood or bone and were used in many parts of the Old World and the Americas. The Latin word for tally (*talea*) means 'cut twig'. Even till the 19th Century tally sticks were used to keep tax records in Britain. The oldest tally stick that has been found is a thigh bone of a baboon discovered in a cave in Swaziland in Southern Africa which is dated to about 35,000 BC. The bone has 29 notches indicating a record perhaps of the lunar month, the time between two full moons. The bone bears some resemblance to the 'calendar sticks' used by some communities in Namibia even today. Another bone found in Czechoslovakia which contains 57 deep notches is dated to about 30,000 BC. The most interesting of the tally mark bones however is the Ishango bone dated to about 20,000 BC which was excavated from the shores of Lake Edward in Africa.

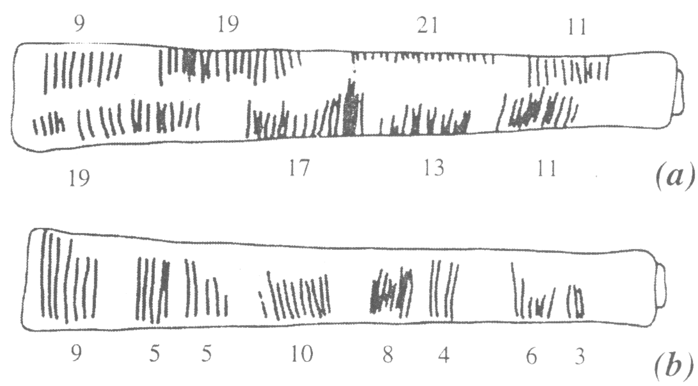


Figure 1.1: The Ishango bone looked at from the front *(a)* and from the back *(b)*

Figure 1.1 shows two views of the Ishango bone. The bone is some kind of tool and has a small piece of quartz fixed at one end. Notice the pattern of numbers, especially the bottom numbers on the front side. They are all the prime numbers between 10 and 20. Both the rows on the front side add to 60. The pattern of the marks also suggests the use of a decimal count system. Considering the interesting pattern on it and its antiquity, the Ishango bone has given rise to much speculation. Some have suggested that it may represent an arithmetical game of some sort. Some have suggested a link to the phases of the moon based on finer markings on the bone that are visible only through a microscope. We may never know what the bone was used for, but it clearly indicates a significant development of number sense even at that early date.

The next important step in the development of written numerals occurs in Babylon. The interesting side to this story is that the development of the first numerals and the first written language appear to take place simultaneously. This is what the clay tablets of Babylon which go back to the 3rd millennium BC tell us.

BABYLONIAN NUMBERS

Mesopotamia or the land between the two rivers Tigris and Euphrates is one of the most ancient civilizations of the world. Around the middle of the 4th millennium BC, the first city-states in Sumer in Southern Mesopotamia grew out of the small agricultural villages in this fertile region. The cities of Ur, Nippur and Lagash were the most powerful of these and attained a high level of cultural development. The cities naturally attracted outsiders looking for plunder and conquest and were attacked numerous times. Around 2400 BC, a large empire of the Akkadians was established in Sumer for about three centuries. The city states then became independent again for a while till another powerful kingdom was established for another three centuries. This was the Old Babylonian empire with its capital at Babylon which lasted from 1900 to 1650 BC. One of the most famous kings of this empire was Hammurabi who lived from 1792 to 1750 BC. Most of the mathematical records of Mesopotamia are clay tablets from the Old Babylonian empire.

Mesopotamia is a region where kingdoms change every few hundred years or sooner. So the Mesopotamian civilization absorbed many influences and currents. The Sumerians, the Akkadians, the Hittites, the Assyrians and the Chaldeans all ruled in succession till the Persian invasion in 539 BC. After this the history of Mesopotamia is no longer independent, it becomes integrated into the other main currents of the history of the period centered around Persia, Greece and Rome.

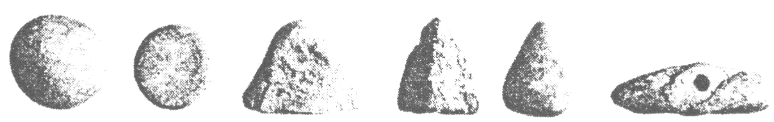


Figure 1.2: Clay tokens from Mesopotamia, each about 1.5 centimeters across

As the city states in Mesopotamia grew so did commerce in these cities. We find a peculiar object associated with this commerce – clay envelopes or hollow clay balls which are sealed from the outside but contain clay tokens inside. The clay tokens appear to be records of goods bought or sold in the ancient world. Clay tokens have been found in several places in Mesopotamia, in the Indus valley and in Africa. The oldest clay tokens go back to 8000 BC, the beginning of the Neolithic period. Figure 1.2 shows some of these tokens from Mesopotamia. Since the clay tokens were placed inside sealed clay envelopes one could not see what tokens were inside. So an impression of the tokens was made on the surface of the clay envelope. This must have led to the realization that the tokens were not really necessary, the marks were enough as records of what was sold.

Thus begins the age of clay tablets of which over half a million have been excavated. These tablets which were sun dried or baked are some of the best preserved archeological records. The clay tablets are in different sizes – from a square inch to the size of a newspaper. They also come from different periods.

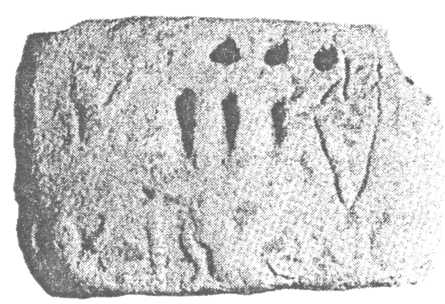
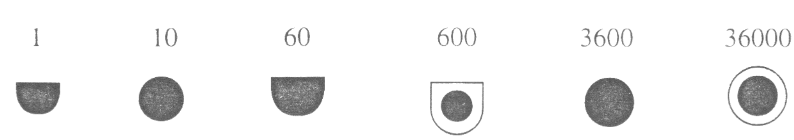
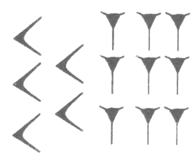


Figure 1.3: Impressions on a clay tablet showing 33 oil jars. The long marks are numerals for 1.

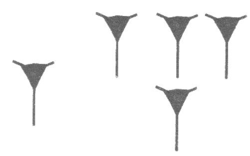
Even in the early tablets the practice of making multiple impressions to indicate a number of objects was replaced by a combination of a numeral and a mark for an object. That is, instead of showing three 'oil-jar-marks' to show three oil jars, the mark for the number 3 would be put next to the mark for the oil jar. Figure 1.3 shows the use of this principle. The early Babylonian signs from about 3000 BC did not show a fully developed positional system. The symbols were



This numeral system developed gradually to the full sexagesimal positional system by 2000 BC. The later system used only two symbols: 𒁹 for 1 and 𒌋 for 10. In this system 4 would be shown as 𒁹𒁹𒁹𒁹 and 40 as 𒌋𒌋𒌋𒌋. The symbol for 59 was the following.

59 

For the next number 60 the Babylonians used the same symbol as for 1 . Thus 240 would be written as . The symbol  was also used for 3600 which is the next power of 60 (60^2). So the number 3604 would be written as

3604 

As you are reading this you are probably wondering whether the Babylonians made mistakes while reading , because they did not know if it was 1 or 60. We don't know whether they made any mistakes, but it is true that the system was ambiguous. For example, the symbol  could mean 90 ($60 \times 1 + 30$) or 3630 ($60^2 \times 1 + 30$) or even 5400 ($60^2 \times 1 + 60 \times 30$). What was the problem? Clearly they needed a symbol to indicate that there was an empty place between  and  while writing the number 3630 ($60^2 \times 1 + 60 \times 0 + 30$). In other words they needed a zero. But a zero did not develop till much later in Mesopotamia. Till then they had to understand the number from the context. In fact the confusion could have been even worse because the Babylonians used the same symbols to show fractions, numbers less than 1.

You might have learnt about both fractions and decimals in primary school. Did you wonder which came first – fractions or decimals? The answer is surprising. Both appear at about the same time in history. The fraction symbol arises in Egypt by the 2nd millennium BCand the sexagesimal (which is the equivalent of the decimal number in a system with base 60) around the same time in Mesopotamia. The Egyptian fractions were somewhat underdeveloped since they had a symbol to show only unit fractions. Unit fractions are fractions which have 1 as the numerator like $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{15}$, $\frac{1}{20}$, etc. In contrast the equivalent of decimal numbers, the sexagesimal numbers were fully developed. The Babylonians had a full fledged numeral system to show numbers less than 1.

In the decimal system to show $\frac{1}{10}$ we use the same symbol as for 1, only we put a decimal point – 0.1. The symbol 1 is also used for $\frac{1}{10}$, $\frac{1}{100}$, etc. Similarly in the sexagesimal system, the symbol 𒁹 is used for $\frac{1}{60}$. The same symbol is also used for $\frac{1}{3600}$. Thus the number $2\frac{1}{2}$ would be written as 𒈫 𒌍 $(2 \times 1 + 30 \times \frac{1}{60})$. Notice that this is no different from the symbol for 150.

The Babylonians did not have a symbol for a decimal point. Because of this the ambiguity in their numerals was even more. The numeral for $2\frac{1}{2}$ could be read as 150 or even $\frac{150}{3600} = 0.04166667$! If you now check the symbol for 90 which we discussed, 𒁹 𒌍, we now realize that this could have meant 90 or 3630 or 5400 or $1\frac{1}{2}$ $(1 + 30 \times \frac{1}{60})$ or even more numbers.

This did not however prevent the Babylonians from carrying out their business and even doing some impressive computations. One of their clay tablets, for example, shows the value of $\sqrt{2}$ as 𒁹 𒎙𒐉 𒐐𒁹 𒌋 (see Figure 1.4). Let us find out how much this is

$$1 + 24 \times \frac{1}{60} + 51 \times \frac{1}{60^2} + 10 \times \frac{1}{60^3}$$

$$\approx 1 + 0.4 + 0.01416667 + 0.0000463$$

$$= 1.41421297$$

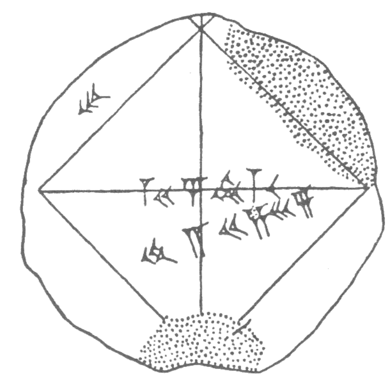


Figure 1.4: A clay tablet from Mesopotamia showing the value of $\sqrt{2}$

The correct value of $\sqrt{2}$ to the same number of decimal places is 1.41421356, which makes the Babylonian value correct to 5 decimal places. This was in the 2nd millennium BC. There are two things which we come to know from this impressive tablet. One of course is the power of the sexagesimal number system in carrying out computations. In fact the Babylonians could carry out the 4 basic operations, and also square root extraction as we have seen, with ease. The other point made by the tablet about Babylonian mathematics is that their knowledge and use of the

Pythagoras theorem that the square on the hypotenuse is equal to the sum of the squares on the other two sides.

Who invented the zero?

Every Indian feels proud about the contribution of Indian mathematicians in inventing the zero. The story of the invention of zero is not such a simple one - no Indian mathematician or philosopher woke up one day having dreamt at night of nothingness and announced the invention of the zero. Like many other developments in mathematics, this story is a long one and proceeds in stages. The first stage is the development of positional notation. We have already seen that this arose first in Babylon. In a positional notation ambiguities in reading numbers are unavoidable. If there were no zero in the decimal system that we use for example, we could not make out if the numeral '67' was the number 67 or 607 or 6007 or 670 to consider only four possibilities. As we have seen in the old Babylonian period of around 1800 BC, this ambiguity is present and there is no symbol for zero.

One can imagine that the need to avoid ambiguity would have been sufficiently strong for the Babylonians to naturally arrive at a symbol for an empty place holder. Indeed this seems to have happened. When we come to the Selucid period in Babylon, after the conquest of the area by Alexander in 312 BC, we find that the Babylonians had begun to use the symbol  to show an empty place in the number. The same symbol was also used sometimes like a full-stop – as a separation mark between sentences. Thus the number 3620 would be written as  .

However, even during the Selucid era, there is no clear example among the Babylonian clay tablets of the symbol for zero occurring at the end of a number. The numbers 2, 120 and 7200 would all be written as  . So the use of the sign for zero at the end of a number is the next stage in the invention of the zero. The first instance of this occurs in the writings of Ptolemy, the astronomer whom we mentioned earlier. Ptolemy lived in the 2nd Century AD. Interestingly Ptolemy used the Greek letter omicron *o* as the symbol for zero.

Indian numerals with the use of the zero symbol in the middle as well as at the end of a number are seen on rock and metal inscriptions from the eighth century AD onwards. Given the frequency with which the zero symbol occurs over a widespread area in India from this period onwards, it is quite likely that the use of zero as a symbol stems from an earlier period. There are some references to the use of the dot as a symbol for zero (*sunya-bindu*) while writing numbers in literary texts which are older, but there are no written records from earlier periods.

The Mayan numeral system that we mentioned earlier was a positional system and had a symbol for zero. In the Mayan numbers the zero symbol occurs both at the end as well as in the middle. Generally the numerals are written vertically and the place value increases as we go up. Two factors made the Mayan system extremely economical. First it used only three different signs - a dot signifying 1, a bar signifying 5 and a shell shaped symbol for zero. Other numerals between 2 and 19 were shown with combinations of the bars and dots. The second important factor is that the Mayan numeral system has 20 as base. So very large numbers can be expressed with fewer places than in the decimal system. Unfortunately we have very few written texts extant from the Mayan civilization since most of them were destroyed by the Spanish conquerors. We do not know how exactly the Mayans did computations with their numbers.

As we have seen, a symbol for zero occurs independently in different cultures across the world. The Indian contribution to the story really lies in treating zero as a number in its own right, not merely in using it as a symbol for an empty place. Halstead captures this picturesquely when he says that the contribution of the Indians lies in "giving to airy nothing, not merely a local inhabitation and a name, a picture, a symbol, but also a helpful power". The Indian mathematician Brahmagupta who lived in the 6th Century AD, recognized that zero is a number which has a place in arithmetic operations. Sridhara, a mathematician who lived around 900 AD, gives the rules for operations with zero except for division. Nearly all Indian mathematical texts from this period onwards mention the rules for operations with zero. The great mathematician Bhaskara II (12th Century AD), specifies the correct rules for all operations including dividing by zero. There are also remarks by Bhaskara which show subtle distinctions between the number zero and an infinitesimally small quantity.

We have traversed many civilizations and time periods in our account of how numbers and numeral systems developed. There are still many themes that we have left out. The bibliography at the end of the book will give you leads to follow up on this story. In the remaining chapters we will look at another area of mathematics that has fascinated humans from early times - geometry. Here too we will unravel threads that connect these ancient periods to modern school mathematics.

2 The story of π

π is one of the most important and intriguing numbers in mathematics. Nearly every school student is familiar with this universal geometrical constant. It is the ratio of the circumference to the diameter of any circle on a flat plane. π also figures in the formulas for the surface area and volume of spheres. You might recall that the surface area of a sphere is $4\pi r^2$ and the volume of a sphere is $\frac{4}{3}\pi r^3$. In fact π finds a place in the formulas for the surface areas and volume of spheres even in dimensions higher than three.

π also turns up in the most unexpected places. One such place is the Buffon's needle experiment. Suppose you have a ruled paper with parallel lines drawn at equal intervals say a distance d apart. Let the paper lie flat on the table. Take a needle of length d and let it fall randomly on the paper. When the needle lands it may or may not cross a line. The probability that it does cross a line is $\frac{2}{\pi}$ which is equal to $0.63662\ldots$.

You might argue that since the orientation of the needle is chosen randomly from $360°$, the probability is likely to involve the circle and hence π. Here then is another probability measure involving π, this time involving whole numbers. Choose two integers entirely at random. What is the probability that they are co-prime, that is, that they do not have any common factors except 1? The probability turns out to be $\frac{6}{\pi^2}$ which is equal to $0.60793\ldots$.

These are, of course, developments in modern mathematics. The ancients did not explore the properties of π as a number. π is a letter of the Greek alphabet but the Greeks did not use the symbol π for the ratio of the circumference to the diameter. In fact, in the Greek system of numerals the symbol 'π ' would have stood for the number 80! The first use of 'π ' to denote the ratio of circumference to diameter was in 1706 by an Englishman William Jones. Euler, the great Swiss mathematician, adopted it in 1737 and made it popular. The ancient mathematicians did not always pose their problems in the way we do in modern times. They did not ask what the value of π is. They posed problems somewhat differently. What is the ratio of the area of a circle to the square on its diameter? How do we construct a square exactly equal in area to a circle? If a circle has a diameter of so many units, what is the length of its circumference? We will look at some of these questions below.

We have used the common school textbook definition of π – the ratio of the circumference of a circle to its diameter. Many school books also give the value of pi as $\frac{22}{7}$. It is important to notice that these statements come with some qualifications. The first is that the value $\frac{22}{7} = 3.142857\ldots$ is an approximate value for π correct only to two decimal places. Zu Zongshi, a Chinese mathematician, obtained a value in 480 AD which is more accurate than that. His value was correct to six decimal places ($3.1415926 < \pi < 3.1415927$). Today with the use of some powerful mathematical techniques and computers, the value of π has been calculated to more than 50 billion decimal places! It is no longer of any practical significance to find a more accurate value of π. But some of the mathematics associated with the techniques for calculating pi can be quite fascinating.

The other point that one must notice about π is that it is irrational. Although we define π as a ratio between the circumference and diameter, these are actually **incommensurable**. This means that it is not possible to find a unit length of which the circumference and the diameter are exact multiples. This is impossible however small we make the unit length. You must remember that we are talking about ideal geometrical objects, not those actually drawn on paper. It is possible that if you used a unit of length as small as, say, the size of an atom, you might find that the diameter and the circumference of an actual circle printed on paper are exact multiples of this unit length. But this is not relevant. Ideal circles are different from those drawn on paper, and for these even the tiniest units will not exactly measure out both the circle and the diameter. Another way of saying that the circumference and diameter are incommensurable is to say that it is impossible to express π as a rational number, that is a number of the form $\frac{p}{q}$ where p and q are integers.

Lambert in the 18th century proved that π is irrational. However the ancient mathematicians probably realized the irrationality of π. The idea is nicely captured in this quotation from around 1500 AD, from a commentary written on the *Aryabhatiya*, an important astronomical text in Ancient and Medieval India. Nilakantha Somayaji of the Kerala school of mathematics, who is the author of the commentary, writes,

> Why is only the approximate value (of circumference) given here? let me explain. Because the real value cannot be obtained. If the diameter can be measured without a remainder, the circumference measured by the same unit (of measurement) will leave a remainder. Similarly, the unit which measures the circumference without a remainder will leave a remainder when used for measuring the diameter. Hence, the two measured by the same unit will never be without a remainder. Though

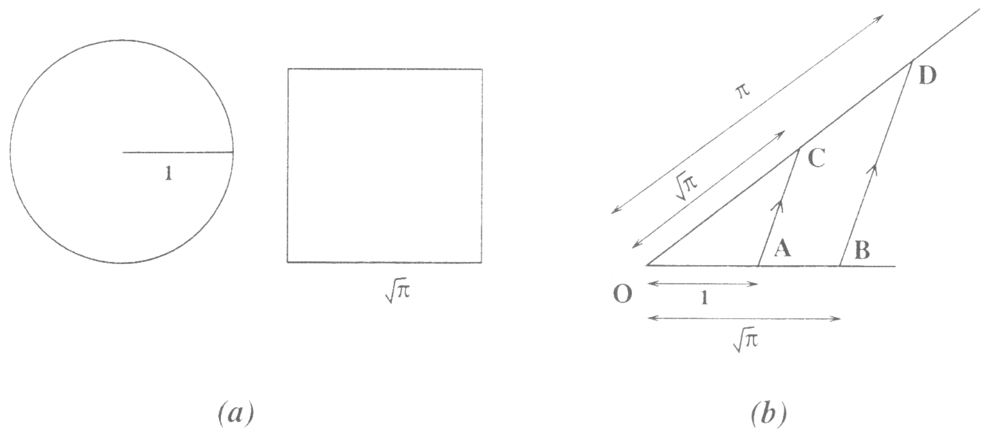


Figure 2.2: (a) Squaring a circle is the same as constructing a length $\sqrt{\pi}$ (b) Constructing a length π from $\sqrt{\pi}$

The construction is shown in Figure 2.2 (b). First make any angle that you like. In the figure O is the vertex of the angle. On one of the arms of the angle (OB) mark the points A and B which are at a distance of 1 and $\sqrt{\pi}$ respectively from O. On the other arm mark a point C at a distance $\sqrt{\pi}$ from O. Join A and C. At the point B draw a line parallel to AC. (A school textbook would describe how to draw a line through a point parallel to a given line.) Let the point where it cuts the other arm of the angle be D. The distance OD will be equal to π. You can verify this with the help of the theorem about parallel lines in a triangle. Compare $\triangle$OBD with $\triangle$OAC

$$\frac{OD}{OC} = \frac{OB}{OA} \quad \text{or} \quad OD = OC \times \frac{OB}{OA}$$

$$\text{or} \quad OD = \sqrt{\pi} \times \frac{\sqrt{\pi}}{1} = \pi$$

Now let us ask about the reverse problem. Suppose that we have a method of constructing a line segment of length π, given a unit length. Can we square a given circle? Again if we assume that the circle has unit radius, we need to construct a square whose area is equal to $\pi \times 1^2 = \pi$. The problem is the same as constructing a line segment of length $\sqrt{\pi}$, given a line segment of length π.

The construction is shown in Figure 2.3. Draw a line segment AB of length π. Extend the line segment so that BC has length equal to 1. Let the midpoint of AC be D. Now draw a semi-circle with AC as diameter. (The centre of the semi-circle is at D.) At B draw a line perpendicular to AC. It meets the semi-circle at E. BE has a length equal to $\sqrt{\pi}$. This is easy to prove if we join AE and CE. The triangle AEC, ABE and EBC are all right angle triangles. Each triangle also shares a vertex angle with one of the other two triangles. hence $\triangle$AEC $\sim$ $\triangle$ABE $\sim$ $\triangle$EBC.

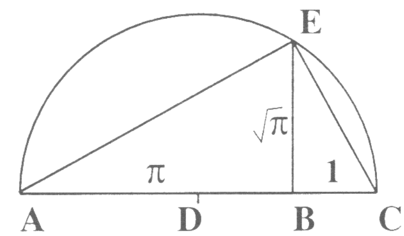


Figure 2.3: Constructing a length $\sqrt{\pi}$ from π

So we have

$$\frac{BE}{AB} = \frac{BC}{EB}$$
$$\text{or} \quad BE^2 = AB \times BC = \pi \times 1 = \pi$$
$$\text{or} \quad BE = \sqrt{\pi}$$

Thus we get a straight line segment equal to $\sqrt{\pi}$ starting from a line segment of length π. So if we are able to square a circle we can also draw a line segment of length π and vice versa. We conclude that the two problems are therefore equivalent. Of course, as Lindemann showed, π is transcendental and hence, both the problems, of squaring a circle and drawing a line of length π cannot be solved.

Notice that the two constructions we have used are general constructions. The first shows that given any line segment of length x, we can construct a line segment of length x^2 using straight edge and compass. The second shows that given any length x we can construct a length $\sqrt{x}$.

Why is π the same for all the circles?

There is a more basic fact about π that we assumed so far without questioning. How do we know that π is the same for all circles on the Euclidean or flat plane? How can we be sure that the ratio of the circumference to the diameter is the same for all circles?

The ancient civilizations seem to have been aware of the constancy of π. However we come across the proof of this fact in Euclid's book the *Elements*, probably the most famous and the most printed book after the Bible. At one time it was printed into more copies than the Bible! We do not know if the older civilizations had a proof or an argument but it is likely that they did. The proof of the constancy of π as it appears in the *Elements* is attributed to a great Greek geometer Eudoxus who lived in the 4th Century BC, a little before Euclid. Eudoxus' proof is interesting because it involves very powerful technique that plays a central role in mathematics. This is the so-called **method of exhaustion**.

A very simple instance of the use of the method of exhaustion is the approximation of a circle by a many sided regular polygon. If you draw a square inside a circle so that its corners touch the circle exactly, (this is called inscribing the square) then some of the area of the circle will be covered by the square and some of the area will be left out. If you inscribe a regular pentagon inside a circle, less area of the circle will be left out of the pentagon, with a hexagon even less. (Figure 2.7 shows a regular hexagon inscribed in a circle.) If we go on increasing the number of sides then less and less areas will be left out. In the limit, that is when the number of sides in the inscribed polygon is infinitely large, the area of the polygon will be equal to the area of the circle.

Of course nothing that we actually construct corresponds to increasing the number of sides to infinity. The concept of a limit covers the cases when an operation can be extended or repeated as many times as we wish and the process has a definite endpoint. For example, if you add up all the fractions in the sequence

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \ldots$$

the sum approaches 1. This means that the more the number of terms in the sequence that you consider the closer you get to 1. The difference from 1 becomes smaller and smaller as you increase the number of terms. We say that the sum of this series of fractions **converges** to 1 or that 1 is the limit of the sum of this series. Similarly the area of the inscribed regular polygon converges to the area of the circle as we increase the number of sides. So the limit of the area of the polygon as the sides increase in number is the area of the circle.

This is strictly modern mathematical jargon. Formalizing the properties of limits and applying the general concept of the limit was a modern development which occurred with the development of calculus after the 17th century. The ancient mathematicians were aware of the basic idea although they did not have a generalized approach to problems involving the limit. They applied the idea and the arguments afresh to every problem.

We will look at two instances where Euclid (or Eudoxus, to be faithful to history) applies the method of exhaustion. The first is to the proof that the area of the circle is proportional to the square of its diameter. We will discuss the second example, where he uses the method to prove that pyramids with equal heights are proportional to the areas of their bases, in the next chapter. In both cases we see that the pattern of argument is the same. However no general principles or properties of a limit are used. Rather the exhaustion is carried out in each proof afresh.

Our problem of the constancy of π is somewhat different from the proposition proved by Euclid. Euclid does not prove that the circumference is proportional to the diameter for any given pair of circles. In the *Elements*, at least in the version that has come down to us, Euclid does not deal with the length of arcs or circles or other curves. Euclid proves, in Proposition 2 of Book XII of *Elements*, that the areas of the two circles are proportional to the squares on the diameter. That is he proves that

$$\text{Area of any circle} \propto d^2$$

If we know the relation between the area of the circle and its circumference, then we can show that the circumference is proportional to the diameter. We can relate the area of a circle to its circumference by a simple argument.

Let the circumference of the circle be c. We can cut up the circle into parts by drawing many radii which are at an equal angular distance from each other as in Figure 2.4 . Each of these parts is like a triangle whose base is a part of the circumference and whose sides are of length r. The base is of course, curved and not a straight line. But we can divide the circle into very large number of parts, so that the bases can be approximated by a straight line. Also we can assume that the height of the triangle is equal to r. Then the area of each triangle is equal to $\frac{1}{2} \times base \times r$. The combined area of all the triangles is equal to $\frac{1}{2}$ × sum of all the bases $\times\ r = \frac{1}{2} \times c \times r$.

We have effectively used a method of exhaustion in this argument without being very rigorous. We will see soon how the Greeks ensured that their proofs involving the method of exhaustion did not compromise in rigour. For the time being we will use the relation we have found to show that saying that the areas of circles are proportional to the squares on the diameter is the same as saying that π is constant.

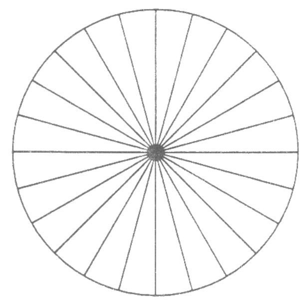


Figure 2.4: The relation between circumference and area

Suppose that the area of a circle $\propto d^2$

That is, the area of a circle = kd^2 where k is some constant.

Then we have,

$$\frac{1}{2} \times c \times r = kd^2 \quad \text{or} \quad \frac{1}{4} \times c \times d = kd^2$$
$$\text{or} \quad C = 4kd$$

Hence we see that $C \propto d$ or π is a constant.

Euclid proves that areas of two circles are proportional to the squares on their diameters in two steps. He first proves that similar polygons which are inscribed in two circles have an area which is proportional to the squares on their respective diameters. This is Proposition 1 of Book X11 in the *Elements*. This can be easily proved by using the fact that similar triangles have areas which are proportional to the squares of their sides. Any polygon can be split into triangles by joining one of the corners of the polygon to all other corners. Similarly the polygons inscribed in the circle can be split into triangles and one can show that their areas are proportional to the squares of the diameters of the circles. We will leave the details of this proof to the reader. Let us however state this proposition.

Proposition 2.1 Similar polygons which are inscribed in circles have areas which are proportional to the squares on the diameters.

Now Euclid inscribes regular polygons with an equal number of sides in both the circles. He proceeds to show that the circle can be exhausted by these polygons. From Proposition 2.1 it follows that the areas of the two circles are proportional to he squares on their diameters. In the course of the proof he also uses the method of *reductio ad absurdum*, that is he assumes a fact and derives a contradiction thereby showing that the assumption was false. Let us now look at the proof in more detail.

Let c_1 and c_2 be the two circles. Let their diameters be d_1 and d_2 respectively. We will denote the respective areas of the circles by C_1 and C_2. We need to prove that the areas of the circles are proportional to the squares on the diameters. That is, we need to prove that

$$\frac{C_1}{C_2} = \frac{{d_1}^2}{{d_2}^2}$$

Suppose that $\frac{C_1}{C_2} \neq \frac{{d_1}^2}{{d_2}^2}$, then $\frac{C_1}{C_2}$ could be less than or greater than $\frac{{d_1}^2}{{d_2}^2}$. Let us examine both these possibilities.

Assumption 1

$$\frac{C_1}{C_2} < \frac{d_1^2}{d_2^2}$$

The ratio on the left is less than the ratio on the right. To increase the ratio on the left we could either increase the numerator or decrease the denominator so that the ratios become equal. We will decrease the denominator, that is, we will choose C_2' such that

$$\frac{C_1}{C_2'} = \frac{d_1^2}{d_2^2} \quad \text{and} \quad C_2' < C_1 \tag{2.1}$$

There exists a circle whose area is equal to C_2'. Let us call the circle c_2'. To illustrate the idea, we have drawn c_2' inside c_2. Of course, the circle c_2' could be much closer to the size of c_2 than we have shown. But clearly c_2' is wholly within c_2 since its area is smaller.

Now inscribe a square in c_2 as shown in the Figure 2.5. Bisect the arcs of c_2 on each side of the square. The midpoints of the arcs along with the corners of the square form a regular octagon inscribed in the circle. Now bisect the arcs of c_2 on each side of the octagon, and obtain the vertices of a 16-gon.

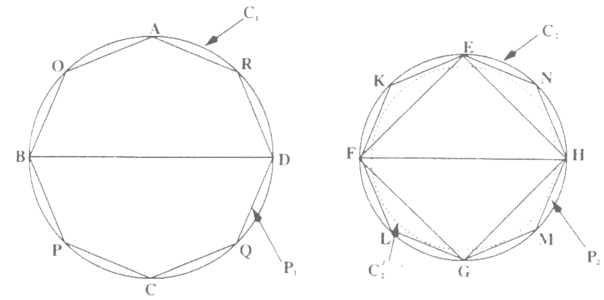


Figure 2.5: Examining *Assumption 1*: $\frac{C_1}{C_2} < \frac{{d_1}^2}{{d_2}^2}$

We can continue this process to get polygons inscribed in c_2 with greater and greater number of sides. At some point we will have the area of the inscribed polygon greater than C_2'. Although this is intuitively clear, Euclid does not rely on intuition alone. He uses propositions proved earlier to show that the area of the circle left over after inscribing the polygon can be made as small as we want and hence can be made smaller than the difference between C_2' and C_2. Therefore we have a polygon inscribed in c_2 whose area is larger than C_2'. Call this polygon p_2 and its area P_2.

(2.2) $$P_2 > C_2'$$

Inscribe a polygon p_1 similar to p_2 in c_1. Let the area of p_2 be P_2. Now from Proposition 2.1 and (2.1) above we have

$$\frac{P_1}{P_2} = \frac{{d_1}^2}{{d_2}^2} = \frac{C_1}{C_2'}$$

We know that $P_1 < C_1$ since P_1 is inscribed in C_1. Hence

(2.3) $$P_2 < C_2'$$

But (2.3) contradicts (2.2) above. Hence we conclude that *Assumption 1* was wrong.

Let us examine the second possibility, namely,

Assumption 2

$$\frac{C_1}{C_2} > \frac{{d_1}^2}{{d_2}^2}$$

Now we make the ratios equal by decreasing the numerator, that is, we choose a C_1' such that

$$\frac{C_1'}{C_2} = \frac{{d_1}^2}{{d_2}^2} \quad \text{and} \quad C_1' < C_1 \tag{2.4}$$

There exists a circle c_1 whose area is equal to c_1'. Figure 2.6 shows c_1' drawn inside c_1. As we did before, we will inscribe in c_1 a square, and then an octagon and then a 16-gon and so on. At some point the inscribed polygon, say p_1' will an area P_1' which is larger than c_1'.

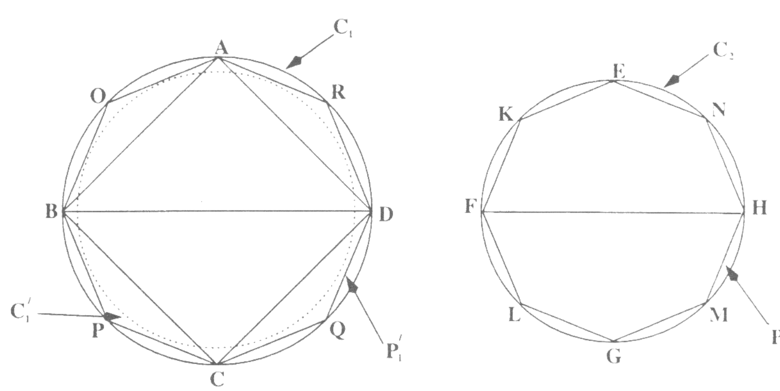


Figure 2.6: Examining *Assumption 2*: $\frac{C_1}{C_2} > \frac{{d_1}^2}{{d_2}^2}$

Hence

$$P_1' > C_1' \tag{2.5}$$

Now inscribe a polygon p_2' similar to p_1' inside c_2. Let its area be P_2'. By Proposition 2.1 and (2.4) above we have

$$\frac{P_1'}{P_2'} = \frac{{d_1}^2}{{d_2}^2} = \frac{C_1'}{C_2}$$

We know that $P_2' < C_2$ since the polygon p_2' is inscribed in c_2. Hence it follows that

$$P_1' < C_1' \tag{2.6}$$

But (2.6) contradicts (2.5) above. Hence we conclude that *Assumption 2* was wrong.

If both *Assumptions 1 and 2* are wrong, the only possibility that we are left with is

$$\frac{C_1}{C_2} = \frac{{d_1}^2}{{d_2}^2}$$

That is, the area of a circle is proportional to the square of its diameter.

As we mentioned earlier, this proposition can be combined with the relation $area = \frac{1}{2} \times c \times r$ to show that the circumference of a circle is proportional to the diameter or that π is a constant. The question that now arises is how did the ancients estimate the value of π. One way to do this would be by actual measurement. Since we know that all circles have the circumference to the diameter ratio π, we could measure the circumference and diameter for many circles and take the average ratio.

Can one find π by making actual measurements?

The only way to answer this question is to try to make the measurements. You could measure the circumference of a circle drawn on paper. Draw the circle with the help of a compass as carefully as you can and let it be as large as possible. Stick some pins (as many as you can) on the circumference. Pass a thread round the pins and measure the circumference. The pins help the thread to stay in place. Measure the diameter and find the ratio. You could do this with different circles and find the average ratio.

What value do you get for π? Compare it with the values given at the beginning of this chapter. How many decimal places did you get right? Think of ways to improve the accuracy of your measurement. Is it possible to get the value of π correct to say 4 decimal places? You could try other variations of this experiment like wrapping a thread around a bangle or a powder tin to measure the circumference. Or you could roll a bangle for one full turn on some rough surface so that it does not slip. Of course we are assuming here that the bangle and the cross-section of the powder tin are examples of perfect circles.

How did the ancient mathematicians find the value of pi?

It is probably clear to you by now that actual measurements give a value of π with very limited accuracy. The oldest texts typically use only rough approximations for π since they were sufficient for practical purposes. In the Bible and in other places we find that the ratio of the circumference to the diameter is taken to be 3. Figure 2.7 shows a regular hexagon inscribed in a circle. The side of the hexagon is equal to the radius of the circle r. So the perimeter of the hexagon is $6r$ or 3 times the diameter. So if we take the value of π to be 3 we are approximating the circumference of the circle by the perimeter of an inscribed regular hexagon.

A better value of π is found in the Ahmes papyrus (also called the Rhind papyrus) from Egypt from about 1650 BC. The Ahmes papyrus is a collection of 87 problems and their solutions. There are several practical problems in arithmetic and mensuration as well as examples probably chosen because they were suitable for teaching. One of the problems, Problem 48 has to do with estimating the area of a circle. This problem is the only problem in the Ahmes papyrus which is accompanied by a diagram. Figure 2.8 (a) shows the diagram as found in the papyrus. The diagram is redrawn in Figure 2.8 (b). The symbol in the centre of the diagram is the numeral for 9 and indicates that the side of the square is 9. The writing which appears below the diagram in the papyrus, calculates the area of the circle as 64.

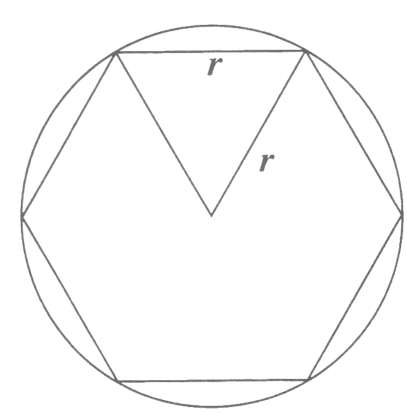
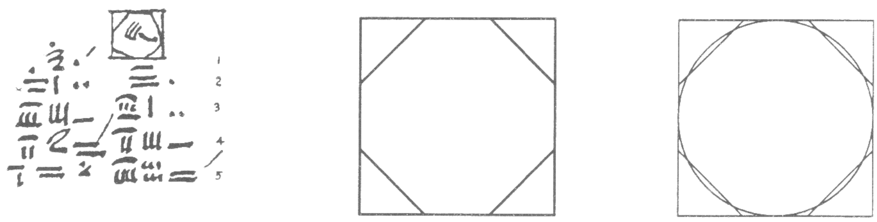


Figure 2.7: Approximating π to 3



(a) (b) (c)

Figure 2.8: The Egyptian estimation of π

Divide each side of a square of side 9 units into three equal parts and obtain an octagon by joining the points as in Figure 2.8 (b). Then the area of this octagon is a rough approximation to the area of the circle. You can get an idea of how close the approximation is by looking at Figure 2.8 (c). Note that the octagon is not a regular octagon. The four triangles which have been cut off at the corners each have an area of $\frac{1}{2} \times 3 \times 3 = 4.5$. The four triangles together have an area of 18. So the area of the octagon is *area of the square - area of the four triangles* = $9 \times 9 - 18 = 63$. This is quite close to the value 64 calculated in the papyrus. The implicit value of π here is about 3.1605, which is correct to one decimal place.

In the Indian *Sulvasutras* (8th to 6th Century BC) we find approximate procedures to construct a circle equal to a given square and a square equal to a given circle. The implicit values for π here are 3.088 and 3.004. Another value for π used in Jaina mathematics around the first few centuries AD is $\sqrt{10} \approx 3.16$. These values are also correct to one decimal place.

Contrast this with Aryabhata who in the 5th Century AD uses an implicit value of π as 3.1416 which is correct to 4 decimal places. Although we do not know how Aryabhata determined the ratio of the circle and diameter, it is quite likely that he used the method of exhausting the circle through regular polygons. Liu Hui, a Chinese mathematician who lived about 2 centuries earlier than Aryabhata obtained a similar value (3.1416). He too used the method of inscribing regular polygons in a circle. Indeed, around the same time as Aryabhata, Zu Zongshi, another Chinese mathematician, found the value of π correct to six decimal places. We have mentioned this at the beginning of this chapter. Zu Zongshi essentially used the same method as Liu Hui, but carried the computation further.

It was Archimedes in 250 BC who first applied the method of inscribing polygons to find the ratio of the circumference to the diameter. The idea was not new even at that time. Euclid (or Eudoxus), as we have seen, used the idea in his proof of the proportionality of the area of the circle to the square of the diameter. In fact, the idea probably originates from Antiphon who lived in the 5th Century BC. Antiphon thought that he had found the solution to the problem of squaring the circle. Begin with a square, he said, which is inscribed in the circle (or possibly the hexagon which is easy to construct). Bisect the arcs to obtain a polygon with double the number of sides. Continue this till you find a polygon which is equal in area to the circle. Since all polygons can be squared, this polygon can be squared. hence the circle can be squared.

Antiphon was of course wrong in believing, if he actually did believe, that we will find a polygon whose area is exactly equal to the area of the circle. it is possible only to approximate the area of the circle, although we can make the error as small as we want. Nevertheless, Antiphon's idea contains the germ of the idea of exhaustion later used by Eudoxus. Archimedes too uses Antiphon's idea to actually compute the ratio of the circumference to the diameter.

What was Archimedes method to find the value of π ?

We will describe the method used by Archimedes to find the value of π briefly since it is followed by so many mathematicians after him. The idea is a natural one, if one thinks of approximating a circle by a polygon. So many of the approaches after Archimedes are probably independent of him. The Chinese mathematicians, whom we mentioned, appear to be unaware of Archimedes work, although they solved some of the problems first solved by Archimedes.

Archimedes first begins with a square inscribed in a circle. He then bisects the arcs above the sides of the square to obtain 4 more points. The eight points are not the vertices of an inscribed octagon. The arcs above the sides of the octagon can again be bisected to obtain the vertices of a 16-gon and so on. We have described this procedure while going over our proof. You can also see Figure 2.5 .

Suppose that we have inscribed an n sided polygon in the circle. We not bisect the arcs above the sides of the polygon to obtain a $2n$ sided polygon. Figure 2.9 shows a portion of the circle – the arc above one of the sides of an n sided polygon. Assume that the circle has a radius of 1. Let the length of the line segment AB, one of the sides of the inscribed n-gon be a. Let the length of the line segment AC, which is one of the sides of the inscribed $2n$-gon, be b. Let the length of the line

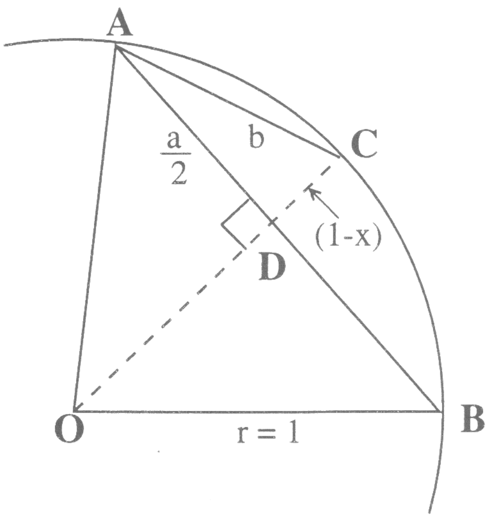


Figure 2.9: Deriving the length of the side of a $2n$-gon inscribed in a circle

segment OD be x.

In Figure 2.9, △ADO is a right angled triangle where ∠D = 90°. So by the Pythagoras theorem,

$$\begin{aligned} OA^2 &= AD^2 + OD^2 \\ 1 &= \frac{a^2}{4} + x^2 \\ x^2 &= 1 - \frac{a^2}{4} \\ x &= \sqrt{1 - \frac{a^2}{4}} \end{aligned}$$

Therefore,

$$CD = OC - OD = \left(1 - \sqrt{1 - \frac{a^2}{4}}\right)$$

In the △ADC

$$\begin{aligned} AC^2 &= AD^2 + CD^2 \\ b^2 &= \left(\frac{a}{2}\right)^2 + \left(1 - \sqrt{1 - \frac{a^2}{4}}\right)^2 \\ &= \frac{a^2}{4} + 1 + 1 - \frac{a^2}{4} - 2\sqrt{1 - \frac{a^2}{4}} \\ &= 2 - 2\sqrt{\frac{4 - a^2}{4}} \\ &= 2 - 2 \times \frac{\sqrt{4 - a^2}}{2} \\ &= 2 - \sqrt{4 - a^2} \end{aligned}$$

Therefore,

$$b = \sqrt{2 - \sqrt{4 - a^2}}$$

where a is the length of the side of a regular inscribed n-gon and b is the length of the side of a regular inscribed $2n$-gon.

By applying this formula and iterating, we can find the values of the sides of the inscribed polygons in each step. Table 2.1 shows the values obtained by iterating upto the 32-gon. If the radius of the circle is 1, then a square inscribed in the circle has a side of length $\sqrt{2}$. Hence we start with the square and put $a = 12$. b now is obtained from the formula above. Multiplying b by 8 $(2n)$ gives the perimeter of the octagon that is obtained. For the next step, we start with the octagon and substitute b for a and $2n$ for n. The basic steps of the iteration are

1. Start with the square, $a = \sqrt{2}$ and $n = 4$
2. Find $b = \sqrt{2 - \sqrt{4 - a^2}} = \sqrt{2 - \sqrt{4 - 2}} = \sqrt{2 - \sqrt{2}} \quad \approx \quad 3.0615$
3. For the next iteration put $a = \sqrt{2 - \sqrt{2}}$ and $n = 2 \times 4 = 8$ and go to step 1.

Iteration	n-gon	a	$2n$-gon	b
1	4 (square)	$\sqrt{2}$	8 (octagon)	$\sqrt{2-\sqrt{2}}$
2	8	$\sqrt{2-\sqrt{2}}$	16	$\sqrt{2-\sqrt{2+\sqrt{2}}}$
3	16	$\sqrt{2-\sqrt{2+\sqrt{2}}}$	32	$\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2}}}}$
⋮	⋮	⋮	⋮	⋮

Table 2.1: Obtaining the sides of a $2n$-gon inscribed in a circle

The perimeter of the $2n$-gon is obtained by $2n \times b$. For the 32-gon which we obtain in the third iteration, the expression for the perimeter is

$$32 \times \sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2}}}} \quad \approx \quad 6.2730969811$$

Since the diameter of the circle is 2, we obtain a value for π

$$\pi \quad \approx \quad \frac{1}{2} \times 6.2730969811 \quad \approx \quad 3.13654849$$

This value is correct to the second decimal place. Notice the interesting pattern with nested $\sqrt{2}$s in the expression for b. Go ahead and calculate taking as many sides of the polygon as you can.

The Chinese mathematician, Liu Hui, obtained the value of π correct to 4 decimal places by starting with a hexagon and successively constructing polygons of 12, 24, 48 and 96 sides. Zu Zongshi obtained the value of π correct to 6 decimal places by going further than Liu Hui. He obtained a polygon of 24576 sides. (Try and find out which polygon he started with and how many times he bisected the arcs.) But the record for inscribing the polygon with the largest sides and computing π before 1500 AD is held by Al-Kashi, a Persian mathematician. In 1429 Al-Kashi computed the perimeter of a regular polygon inscribed in a circle with $3 \times 2^{28} = 805306368$ sides! He obtained a value of $\pi = 3.1415926535897932$, which is correct to 16 decimal places.

The formula for finding π derived by inscribing regular polygons gives better accuracies for π only very slowly. To obtain accuracy for π correct to 16 decimal places we have to take over 28 terms. However, the best methods to find π uses infinite series which are usually obtained by modifying trigonometric series. As far as we know, the first derivation of the value of π based on the infinite series method was done by an Indian mathematician, Madhava of Sangamagramma from Kerala around 1400 AD. The mathematicians of the Kerala school, showed a sophisticated understanding of infinite series and had some important results using the idea of the limit. In fact, they appear to be virtually knocking on the doors of calculus. However, to cover the work of the Kerala mathematicians in detail is beyond the scope of this booklet. We will now pass on the next dimension and consider the geometry of 3-dimensional solids. In the next chapter, we discuss how mathematicians tackled the problem of the volume of solid bodies many centuries ago, long before calculus was developed.

3 The Volume of Solid Bodies

Mensuration is a topic familiar to you from your middle and high school textbooks. For most students this is a topic full of formulas that need to be memorized. It does not seem particularly stimulating or exciting. Yet this area of mathematics was one of the great challenges faced by ancient mathematicians. Obtaining the correct mathematical conception of area and volume must itself have been a long and difficult task. Finding the areas and volumes of different geometrical objects posed a series of problems that occupied many mathematicians of the past. The response to these challenges and the solutions that were obtained by mathematicians from different civilizations are often stunning and brilliant.

You may know many formulas to find the volumes of various solid bodies. Have you wondered how these formulas were first discovered? Some of the formulas are easy to obtain. The formulas for the volumes of cubes and cuboids are simple and can be got directly from the definition. Of course, one needs to formulate the right definition. One way to define volume is to first define the volume of a simple body. The volumes of other bodies can be defined in terms of the volume of the simple body. A cube of unit side is taken to have unit volume. The volume of a cube or cuboid is then the number of unit volumes that it contains. We allow for fractional volumes too, by taking these as parts of the unit volume.

A simple way to measure the volume by experiment, although it is not very accurate in practice, is to use a measuring cylinder. The experiment also serves as a kind of operational definition. It communicates what is meant by volume. A sketch of the experiment is shown in Figure 3.1 We take a cylinder of some convenient diameter and height and make that our standard. Whole unit volumes are marked on the cylinder and each unit is divided into a convenient number of equal parts. We make the first few marks by immersing bodies whose volumes we know. Now the volume of any body can be measured by immersing the body in water. (If the body floats on water we attach a heavy sink to sink it.) We collect the water which overflows and measure it in the measuring cylinder that we have already marked.

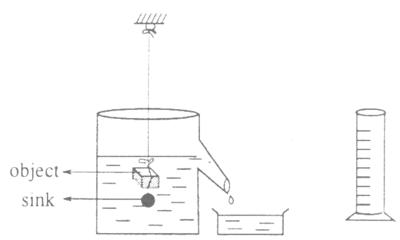


Figure 3.1: A simple way to measure volume

Get hold of a measuring cylinder from the school chemistry lab and try to measure the volume of some solids by this method. If you know the formulas, compare your measurement with the formula. You will probably find a lot of difference in the volume obtained experimentally and by the formula. This is because it is very difficult to avoid the errors which creep into the experiment at different stages. Think about what these errors might be.

Even if we tried to eliminate all these errors by getting expensive equipment and carry out a careful experiment, there would still be limitations. We cannot really think of a careful experiment to immerse the pyramids in Egypt in water to find their volume. Or an experiment to immerse the earth. So clearly it is useful to find exact formulas for the volumes of at least some simple shapes.

Finding the volumes of prisms

The simplest formulas are for the cube and the cuboid. They almost follow from the definition if we think of how many unit cubes or parts of unit cubes make up the given cube or cuboid. Let us write down the formulas for these solids.

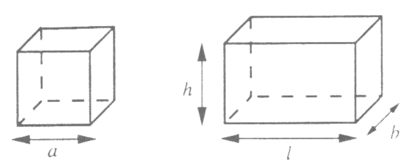


Figure 3.2: Volume of the cube and the cuboid

Volume of a cube = side × side × side = a^3

Volume of a cuboid = length × breadth × height

The simplest solids after the cuboid are the prisms. A prism is a solid composed of pairs of parallel faces and edges. From the formulas that we have just written down it is simple to find the volume of a prism whose base is a right triangle. We can cut any cuboid into two identical halves each of which is a prism of this sort (see Figure 3.3). So we know that the volume of such a prism is half the volume of a cuboid from which it is cut.

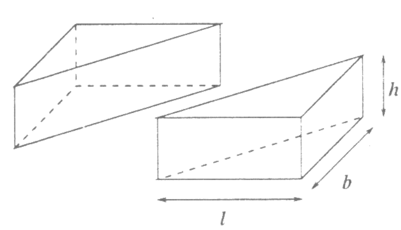


Figure 3.3: Volume of a prism with a triangle as base

Volume of a right prism with a right triangular base = $\frac{1}{2} \times l \times b \times h$

The term 'right' prism means that the prism is upright and not slanting.

We can now find the formula for any prism with a parallelogram as base. This we do by adding two prisms with right triangular bases to the given prism and getting a cuboid as in Figure 3.4. We will call this operation completing the given prism to a cuboid. The volume of the completed cuboid is $(l + x) \times b \times h$. The two right triangular prisms which have been added have a total volume of $x \times b \times h$. So the volume of the prism is $l \times b \times h$. Observe the parallelogram which forms the base of the prism. l and b are the length and height of the parallelogram. $l \times b$ is therefore the area of this parallelogram. So we can say that the volume of the prism is the area of the base × height.

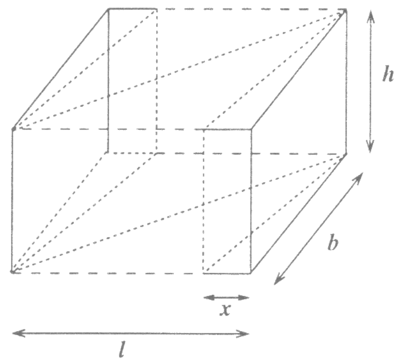


Figure 3.4: Volume of a prism with a parallelogram as base.

Two identical triangles can be put together to make a parallelogram. We can complete any right triangular prism to a prism with a parallelogram as base by doubling it. Figure 3.4 also shows the parallelogram-prism cut into two triangular prisms. So the volume of any upright triangular prism is half the volume of the parallelogram-prism obtained by doubling it.

Here too the formula reduces to: Volume = Area of base × height of the prism. If you check all the formulas for prisms, cubes and cuboids that we have obtained so far, you can see that in general, the volume of the prism is the area of the base multiplied by the height of the prism. What about a prism with any odd shaped polygon as base? Any polygon can be cut up into triangles. Hence any prism with a polygon as base can be cut up into triangular prisms. So we can write down a general formula.

Volume of any upright prism = Area of base × height.

Now we need to check if this formula is also valid for slanting prisms. Let us start with a slanting cuboid. As shown in Figure 3.5, a slanting cuboid can be completed to an upright cuboid by adding two triangular prisms. From the figure, the volume of the completed cuboid is $\left[(l+m)\times b\times h\right] - m\times b\times h = l\times b\times h$.

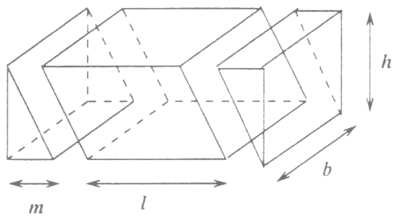


Figure 3.5: Completing a slanting cuboid into an upright cuboid

Now let us consider a slanting prism with a polygonal base. We follow the same steps as we did for upright prisms. The slanting prism with a polygonal base can be cut up into triangular prisms, each of which can be doubled to obtain parallelogram-prisms. These in turn can be completed into cuboids. Since the initial polygonal prism was slanting, the final cuboid and all the intermediate prisms will also be slanting. But the volume of all these slanting prisms is related to the slanting cuboid and is still area of base multiplied by the height. Hence we have a formula for the volume of any upright or slanting prism with polygonal base.

Volume of a prism = Area of base $\times$ height

We have seen that it is fairly simple to obtain the volume of cubes and cuboids once we have a correct concept of volume. It is also a simple matter to go further and obtain the volume of prisms. The volume of the cylinder too is given by the same formula : area of base $\times$ height. You can think of the cylinder as a prism whose base is a polygon with a very large, actually infinite, number of sides. We can assume that for the ancient mathematicians too the solids that we have considered so far did not pose any great problem. The first difficult problem is finding the volume of pyramids. The simplest pyramid is an upright pyramid with a triangular base.

Finding the volume of pyramids

Rather surprisingly, the procedure for finding the volume of a square pyramid appears to have been known to Egyptians as early as the 2nd millennium BC. You would of course know that the Egyptians were great builders of pyramids, which even today are imposing structures. We do not have much direct evidence of the level to which geometry had advanced in Ancient Egypt. However we do find

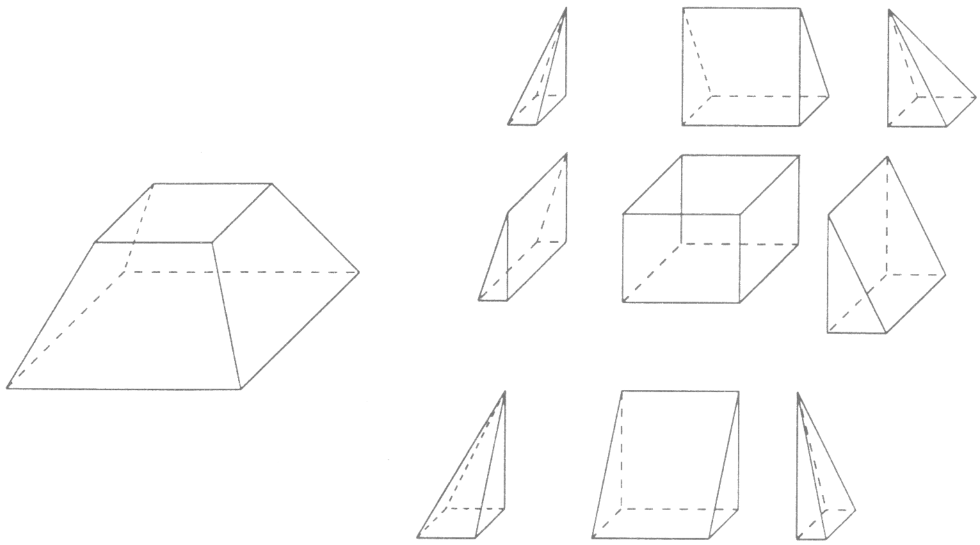


Figure 3.6: The frustum of a square pyramid. The Egyptians knew how to find the volume of this shape.

many Greek writers declaring that the knowledge of geometry first arose among the Egyptians. According to these Greek writers the Egyptians developed geometry to calculate the areas of cultivated plots, whose boundaries were frequently altered by the flooding Nile.

.

We have mentioned the Ahmes papyrus in the previous chapter and discussed the implicit value of π found in the papyrus. Another important source of Egyptian mathematics is the Moscow Papyrus from 1890 BC. Like the Ahmes papyrus it is also a collection of problems. The Moscow papyrus has 25 problems while the Ahmes Papyrus has 87. Virtually all that we know about Egyptian mathematics comes from these two papyri.

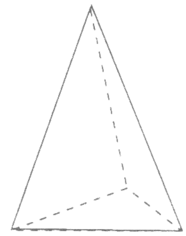


Figure 3.7: A triangular pyramid

In the Moscow Papyrus we find the correct computation of the volume of the frustum of a square pyramid, that is, a pyramid whose top has been cut off. How did the Egyptians obtain this result? Many have suggested that this was done by decomposing the frustum into simpler figures. Figure 3.6 shows how the frustum of a square pyramid could be cut up. Notice that four of the shapes that we obtain are themselves pyramids. Hence the Egyptians must have independently known how to obtain the volume of a pyramid. Let us try and guess how the Egyptians could have found this out.

The cube shown in Figure 3.8 has been cut up into three pyramids. Each of these pyramids has a square base and contains one corner of the cube. In fact, the three pyramids are congruent. That is, to each edge, face and corner in one pyramid, there is a corresponding edge, face and corner in the other two pyramids which are equal to the first. So the volume of each pyramid is $\frac{1}{3}$ the volume of the cube. We could say then that the volume of this pyramid is $\frac{1}{3} \times Areas\ of\ base \times height$. Would this be true for the volumes of all pyramids? We do not still know.

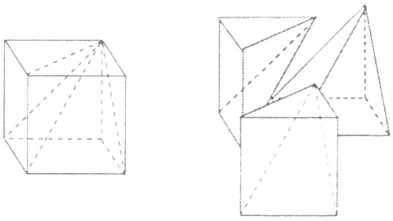


Figure 3.8: A cube cut up into three congruent square pyramids

Let us do a simple experiment. In this case the shape we will consider is a cone, not a pyramid. Take an old postcard or card paper (tetrapack material works best) and cut it into a circle with radius 5.0 cm. Now cut a sector from this circle of 108°. The sector is shown in Figure 3.9. Roll and stick the edges together with gum to obtain a cone as in the figure. This cone has exactly the same height and diameter as a plastic **film roll bottle**. You now have a cone and a cylinder with the same height and circular bases of the same size. Fill the cone to the brim with water or dry sand and pour it into the cylinder. How many times do you have to fill the cylinder before it is full?

It appears from the experiment that the volume of the cone is $\frac{1}{3}$rd that of the cylinder, just as the volume of the pyramid was $\frac{1}{3}$rd that of the cube. We now begin to wonder whether it is true for all pyramids in general that they are have $\frac{1}{3}$rd the volume of the prisms erected on their bases to the same height. Perhaps this was what the Egyptians concluded after conducting some experiments similar to ours. After all, they were right in their calculation of the volume of the truncated pyramid.

We have now two points that suggest the $\frac{1}{3}$rd relationship. The first is cutting up a cube into three identical pyramids. The second is an experimental result. But in order to be sure, we need a mathematical argument preferably a proof. Did the

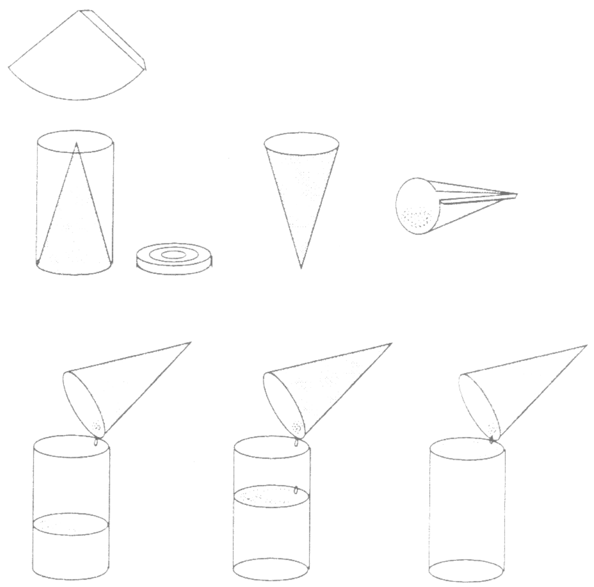


Figure 3.9: Comparing the volume of a cone with a cylinder of the same base area and height

Egyptians have an argument or a proof?

We do not know. But we have Archimedes, the great Greek geometer claiming around 250 BC that Democritus, another great thinker who lived a century and a half before him, was the first to demonstrate the relationship between the volumes of a cone and a cylinder of equal base and height. What was the nature of this demonstration? Archimedes says that Democritus had the idea of cutting up the cone into thin slices each of which could be thought of as a cylinder.

This is a very powerful idea, and if Democritus is truly the originator of this idea, then we are indebted to him. But it is quite possible that the Egyptians too used an argument similar to this to arrive at the $\frac{1}{3}$rd proportion of the pyramid to the cube or cuboid. How would they have argued ? We will try and reconstruct an argument which allows us to derive the volume of the cone and all pyramids including slanting ones. The basic idea in the argument, its creative kernel, is the idea of cutting up a solid figure into thin slices. And the aim of the argument will be to establish the following two propositions.

Proposition 3.1 Two prisms of the same height have volumes proportional to the area of their bases.

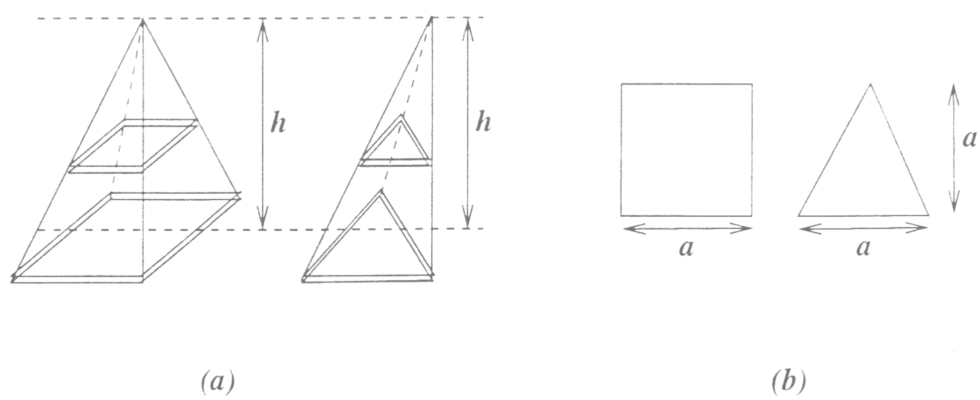


Figure 3.10: Pyramids of the same height but with different bases

Proposition 3.2 Two prisms with the same area of base have volumes that are proportional to their heights.

Cutting them up into slices

In Figure 3.10 (a), two pyramids are shown. One of these is a square pyramid and so has a square base of side a. The height of the pyramid is h. The other is a pyramid with the same height, but with a triangular base. To simplify matters the triangle at the base of the triangular pyramid has both base and height equal to a. The square and the triangular bases are shown in Figure 3.10 (b). It is clear from the figure that the area of the base of the second pyramid is half the area of the base of the first pyramid.

Now let us divide both the pyramids into an equal number of equally thin slices. We could consider each slice to be a prism. This is of course, only an approximation but we could get as accurate as we wanted by making the slices very thin. The volume of each slice is the area of the slice in cross-section multiplied by the thickness.

Take a slice from the first pyramid at some height h_k and a corresponding slice from the second pyramid. Both the slices will be the kth slice starting from the bottom, where k is some number. Compare the shapes of the two slices in cross-section. The shapes are still similar to those shown in Figure 3.10 (b). Both the slices are smaller than the bases of their pyramids, but their cross-sections are in the same ratio to the bases. This can be easily proved by considering each of the faces of the pyramid and applying the theorem about parallel lines in a triangle. The base and height of the triangular cross-section are equal and can be named a_k. The side of the square cross-section is also a_k. The area of the triangular cross-section is half the area of the square cross-section. Both the slices have the same thickness. Hence the volume of the triangular slice is half the volume of the square slice.

The same argument holds true for all the slices. Each triangular slice has half the volume of each corresponding square slice. So we can conclude that the volume of the second pyramid in half the volume of the first pyramid. Notice too that the argument can be made perfectly general although we took simple shapes and sizes to illustrate the point. In general the volumes of corresponding slices will be the same proportion as the area of the bases. Hence we conclude that two pyramids with the same heights have volumes proportional to the area of their bases. The argument clearly holds for pyramids with any polygonal base. Moreover nothing in the argument prevents it from being applied to cones, where the slices will be cylinders instead of prisms. Another nice result is that the argument applies equally well to slanting pyramids and cones. As long as their heights are the same, the thickness, the area and the number of slices remain the same. So if two pyramids or cones are of equal height, then their volumes are proportional to the areas of the base, even if the pyramids and cones are slanting. We see therefore that Proposition 3.1 is true.

What about Proposition 3.2 above, which speaks of the volumes of pyramids with identical bases but with different heights. Figure 3.11 shows two pyramids with identical bases but with different heights. Let us say that the second pyramid is 1.5 times as high as the first. Now cut the two pyramids again into an equal number of thin slices each of which is approximately a prism. Again we can get as accurate as we wish by making the slices very thin. But this time the thickness of the slices in the two pyramids are different although in each pyramid all the slices are equally thick. Since the total number of slices in both pyramids is the same, each slice in the second pyramid would be 1.5 times as thick as each slice in the first pyramid. Further slices which are half way to the top in both pyramids would have the same cross-sectional area. Similarly slices which are at 0.9 times the height in both pyramids would have the same cross-sectional area. In other words, each slice in the first pyramid would have the same area in cross-section as the corresponding slice in the second pyramid. But the slice in the second pyramid would be 1.5 times as thick as the slice in the first pyramid. Since this is true of each pair of corresponding slices, the total volume of the second pyramid would be 1.5 times

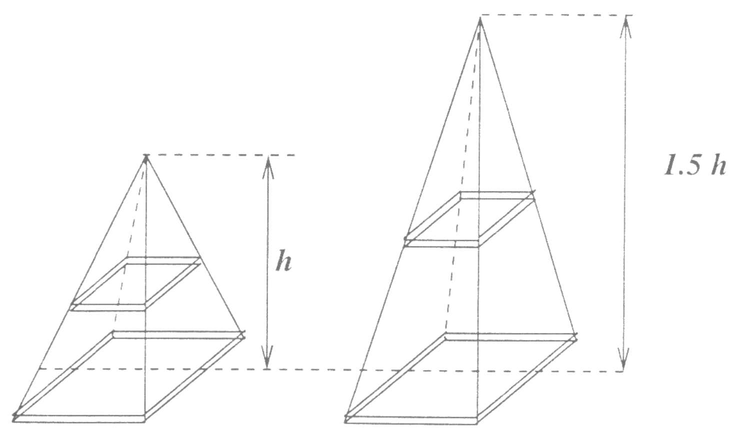


Figure 3.11: Pyramids of the same base but with different heights

the volume of the first pyramid.

The argument would hold even if the ratio of the heights of the two pyramids was not 1.5. Whatever be the ratio, we cut an equal number of slices in the two pyramids. For a pair of corresponding slices, the cross-sections would be the same but the thicknesses would be different. The ratio of their thicknesses is the same as the ratio of the heights of the pyramid. It is easy to see that the argument would then be perfectly general. So we have also demonstrated the second proposition – the volumes of two pyramids with bases having the same area are proportional to their heights.

The idea of taking thin slices of the two pyramids (or any solid body in general), comparing corresponding slices and then comparing the volumes is better known as the Cavalieri principle. Bonaventura Cavalieri was an Italian physicist and mathematician who lived in the Seventeenth Century AD. But like the Pythagoras theorem and the Pascal's triangle, the Cavalieri principle was known and used long before the time of the mathematician after whom it is named.

Euclid's proofs

Propositions 3.1 and 3.2 do not still tell us how to calculate the volume of a pyramid. This requires another proposition, which is one of the propositions found in Euclid's *Elements*[Book XII, Proposition 7]. The proof of this proposition as we find it in the *Elements* is remarkable in its beauty and simplicity.

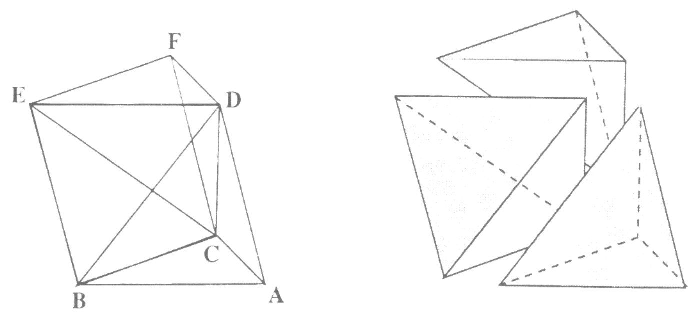


Figure 3.12: Cutting up a prism into 3 equal pyramids

Proposition 3.3 Any triangular prism can be cut into three pyramids of equal volume

The triangular prism shown in Figure 3.12 is cut into three prisms by the lines joining the vertices in the figure. ABED is a parallelogram and hence the triangles ABD and EBD are congruent. So the two pyramids EBDC and ABDC have congruent bases. Further their bases lie on the same plane and they have a common vertex C. So the height of the vertex C from the base of both pyramids is the same. From Propositions 3.1 and 3.2 it follows that the volumes of the two pyramids are equal.

We have looked at two pyramids whose bases lie on the face ABED of the prism. Now consider another face of the prism BEFC which is also a parallelogram. Triangles BEC and FEC are therefore congruent. But these two triangles are the bases of the pyramids BECD and FECD, which share a common vertex at D. The bases of these two pyramids BECD and FECD lie in the same base and hence the common vertex D is at the height from the base in both pyramids. It follows that these pyramids have the same volume.

Now notice that the pyramid BECD is the same pyramid that we named EBDC in the earlier paragraph. So we have

$$\begin{aligned}\text{volume of pyramid EBDC} &= \text{volume of pyramid ABDC}\\ \text{volume of pyramid FECD} &= \text{volume of pyramid EBDC}\end{aligned}$$

Hence the volumes of the three pyramids in the prism are equal. Thus the volume of a pyramid with a triangular base is $\frac{1}{3}$rd the volume of a prism of equal height erected on its base.

Just as any prism with a polygonal base can be cut up into prisms with triangular bases, any pyramid with a polygonal base can be cut up into pyramids with triangular bases. Since the volume of each pyramid with a triangular base will be $\frac{1}{3}$rd of the volume of the corresponding prism, we have a general result – the volume of any pyramid with a polygonal base is $\frac{1}{3}$rd the volume of a prism with equal base and height.

We can think of the cone as a pyramid whose base is a polygon with an infinite number of sides. So we would expect the volume of a cone to be $\frac{1}{3}$rd the volume of a cylinder with the same height and base. Let us then write down the formula for the volume of a cone.

Volume of a cone = $\frac{1}{3}$ × volume of a cylinder with the same base and height

The important steps in this derivation were Propositions 3.1, 3.2 and 3.3. All of these including the expression for the volume of a cone are derived in Book XII of Euclid's *Elements*. Euclid uses the method of exhaustion repeatedly in this book. In the previous chapter, we have seen how this method is applied to show that the ratio of the area of a circle to the square of the diameter is the same for all circles. In demonstrating Propositions 3.1 and 3.2 above we have used something similar to a method of exhaustion. The Cavalieri principle which involves the cutting up of the pyramid into a number of thin slices (in the limit on infinite number of slices) is a version of the method of exhaustion. However, Euclid uses a different approach which is also an application of the method of exhaustion. Although this approach is not general and is applicable specifically to the triangular pyramid, it is interesting and intuitively appealing. We also find echoes of this approach in Chinese work on the volume of solid bodies. So let us take a brief look at this approach.

Recall Euclid's proof that the area of a circle is proportional to the square on its diameter. We looked at this proof in detail in the previous chapter. In the course of this proof we needed to find a figure which in the limit would exhaust the area of the circle. Moreover, this figure had to be one whose area was known. Such a figure was the polygon. The polygon is composed of triangles and hence its area can be determined. As we increase the number of sides, the polygon exhausts the circle.

The proposition that we now wish to demonstrate is Proposition 3.1 above, which states that pyramids with equal heights have volumes proportional to the areas of

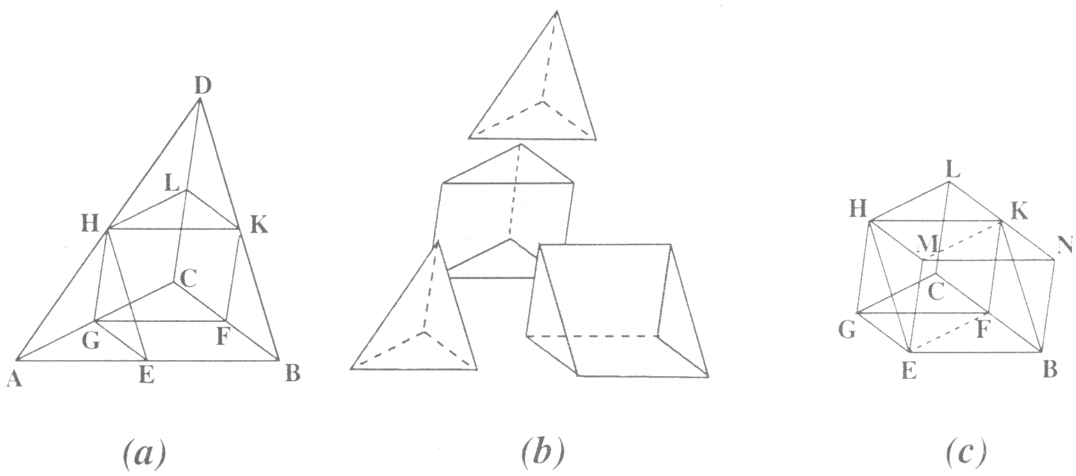


Figure 3.13: A pyramid cut into two equal pyramids and two prisms of equal volume

their bases. Euclid first proves it for pyramids with triangular bases. The approach that Euclid uses is the method of exhaustion. In the case of the triangular pyramid, just as in the case of the circle, we need to find a solid whose volume is known and which can exhaust the volume of the pyramid. It turns out that such a solid is the triangular prism. The triangular pyramid is exhaused by suitable triangular prisms. This is however a little more complicated than exhausting the the circle by inscribed polygons.

Figure 3.13 shows how a pyramid can be decomposed into two prisms and two smaller pyramids. The triangle HKL is halfway to the top of the pyramid ABCD. So H, L and K are midpoints of AD, CD and BD respectively. G,E and F are midpoints of the sides of the triangular base AC, BC and AB respectively. It is clear that $\triangle$HKL $\cong\triangle$GFC $\cong\triangle$AEG. Further all these triangles are similar to $\triangle$ABC.

The figure shows that the original pyramid ABCD is now broken up into two pyramids HKLD and AEGH and two prisms. The first prism is GCF-HLK and the second prism is HGE-KFB. It is easy to see that the two small pyramids are congruent. What may perhaps be surprising is that the volumes of the two prisms are also equal. Euclid actually proves this separately as the last proposition of Book XI [Prop 39]. To obtain the proof we need to complete the second prism HGE-KFB into a parallelogram-prism, which is the prism HMNK-GEBF in Figure 3.13 (c). We will call the new prism the **completed prism**. Clearly the volume of the prism

HGE-KFB is half the volume of the completed prism. What we have to show is that the volume of the prism HLK- GCF is also half of the completed prism.

In Figure 3.13 (c), the two dotted lines MK and EF have been drawn. This gives us a new prism GEF-HMK which we will now show is congruent to the prism GCF-HLK. Since G, E and F are midpoints of the sides of $\triangle$ABC, $\triangle$GCF $\cong\triangle$GEF. In the completed prism the parallelograms GEBF and HMNK are congruent. Hence $\triangle$GEF $\cong\triangle$HMK. So the volume of the prism GEF-HMK is half the volume of the completed prism. But $\triangle$HLK $\cong\triangle$GCF $\cong\triangle$HMK. Hence the volumes of the prisms GCF-HLK and GEF-HMK are both equal. Hence the volume of the prism GCF-HLK is equal to half the volume of the completed prism. So we see that the two prisms GCF-HLK and HGE-KFB in the original pyramid have equal volumes.

Each of these prisms has thrice the volume of each of the smaller pyramids. So the smaller pyramids together make up $\frac{1}{4}$th the total volume of the original pyramid. We can cut up these smaller pyramids further into smaller prisms and pyramids. The remaining pyramids can be cut further. We can continue this till all the volume is exhausted by prisms and the volume of the pyramids left over is as small as we please. Now the original pyramid is exhausted by prisms whose volumes are known.

Now if we have two pyramids of equal height but different bases, then we can exhaust both pyramids by prisms as we have seen above. The volumes of two prisms of the same height are proportional to the area of their bases. Hence the combined volume of the prism pieces for each pyramid will be proportional to the areas of the bases of the two pyramids. Hence the volume of the two pyramids is proportional to the area of their bases. We will skip the full rigorous statement of the later steps of the proof. Once we have proved Proposition 3.1 we can prove Proposition 3.3 , that the volume of a triangular pyramid is $\frac{1}{3}$rd the volume of the prism erected on its base, as we did earlier.

We have seen how ancient mathematicians found the volume of a cone by first finding the volume of a triangular pyramid and extending it to a pyramid with a polygonal base. We have looked at Euclid's proof since it successfully implements what intuition would suggest, namely, breaking up the pyramid into known solids. We find a very similar approach taken by the Chinese mathematician Liu Hui in the 3rd century A.D in order to obtain the volume of a square pyramid. On the basis of the propositions we have discussed Euclid derives the volume of all pyramids and cones. It is impossible to discuss all these proofs in this booklet. Many of the later propostions also involve the method of exhaustion. The full proofs are found in

Book XII of Euclid's *Elements*. One cannot but be struck by the rigour and beauty of these proofs. In each of them the method of exhaustion is used in a consistent manner.

In the next chapter, we will move on to a problem which is more difficult that the problem of finding the volume of a pyramid or a cone. The cone is a relatively simpler solid in comparison with the sphere, which is a truly curved surface. We now turn to an account of how the ancient mathematicians tackled the sphere.

4 The Volume of the Sphere

A Sphere is a truly curved body and the geometry of the sphere is more difficult than that of the cone. As we would expect, it took longer for mathematicians to find the volume and surface area of a sphere than to find these for a cone. Archimedes, who lived nearly a century after Euclid in 250 BC, was the first geometer who determined the volume and surface area of a sphere. Much later, two Chinese mathematicians, a father and son team, Zu Zongshi and Zu Xuan independently found the volume of a sphere. Zu Zongshi was determined to remove the "scar on mathematics" – the scar was the lack of an expression for the volume of a sphere.

The problems of finding the volume of a sphere and the surface area of a sphere are closely bound up with each other. We find a simple argument in *Siddhanta Shiromani* written by the great Indian mathematician Bhaskara, which derives the volume of a sphere from the surface area of a sphere. Bhaskara considers a tiny circle drawn on the surface of a sphere as in Figure 4.1. The centre of the sphere is connected to all the points on the circumference of this small circle. By this construction we obtain a cone whose base is actually a curved surface which is part of the surface of the sphere. The slant height of the cone is r, the radius of the sphere. By making the circle drawn on the sphere very small we can approximately take the base of the cone to be a flat circle. We can also take its height to be r. The volume of the cone then is $\frac{1}{3} \times area\ of\ the\ base \times r$.

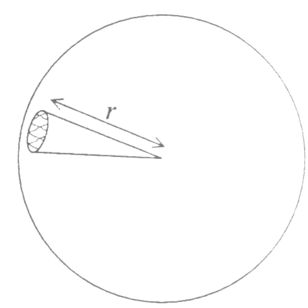


Figure 4.1: Bhaskara's derivation of the relation between surface area and volume of a sphere

We can think of the sphere as composed of a large number (actually an infinite number) of such cones. All the cones are of height r. Hence the volume of the sphere is the combined volume of the cones. So we can write

Volume of the sphere = $\frac{1}{3}$ × Combined area of the base of all cones × r.

But the combined area of the bases of the cones is the surface area of the sphere. We will assume that we know the surface area of a sphere to be $4\pi r^2$. Hence we have

Volume of the sphere = $\frac{1}{3} \times 4\pi r^2 \times r = \frac{1}{3} \times 4\pi r^3$

This argument also allows us to work backwards. If we know the volume of the sphere we would be able to find its surface area. But how do we find the first one of these two – volume and surface of the sphere? One method would be to cut the surface of the sphere into thin strips which are like the circles marked by the latitude lines on the globe. One could then cut open the strips and find their surface areas. This is the approach that is followed in Indian mathematical texts from about the 10th Century onwards. The approach requires that we know how to find the sum of the first n squares.

We do not know if the Greeks knew how to sum a series of squares. But we find a truly marvelous determination of the volume of the sphere by Archimedes in 250 BC. The proof is also a proof of Archimedes' sheer ingenuity. Archimedes himself was so proud of this achievement, that he wanted a diagram of the sphere inscribed in a cylinder to be drawn on his tombstone.

How did Archimedes find the volume of a sphere?

You have probably heard of Archimedes. He was the Greek who jumped out of his bathtub and ran through the streets shouting "Eureka". Or so the story goes. Many of these stories, like the story of Newton and the apple, were first born in the imagination of some later biographer. But we do know, that Archimedes was the one who discovered the principle of the lever. He probably even put it to use by constructing terrible war engines that hurled huge rocks at invading Romans. For this story however, we have to rely on what was written by Plutarch, a Greek and Roman biographer. The principle of the lever or the principle of balance (which is also a kind of lever) was a powerful metaphor for Archimedes. It played a role in his development of the theory of floating bodies. Many centuries later, when Galileo read Archimedes' writings in the 16th Century, he too was impressed by the power of this metaphor. The idea of the 'balance' helped Galileo to develop his theory of freely falling bodies.

This is not something that you would expect. Archimedes actually used the principle of the balance in deriving the relation for the volume of a sphere. Let us see how he was able to do this.

Consider a sphere, a cone and cylinder of the same height d as in Figure 4.2. The cone and the cylinder have identical circular bases. The diameter of the base for

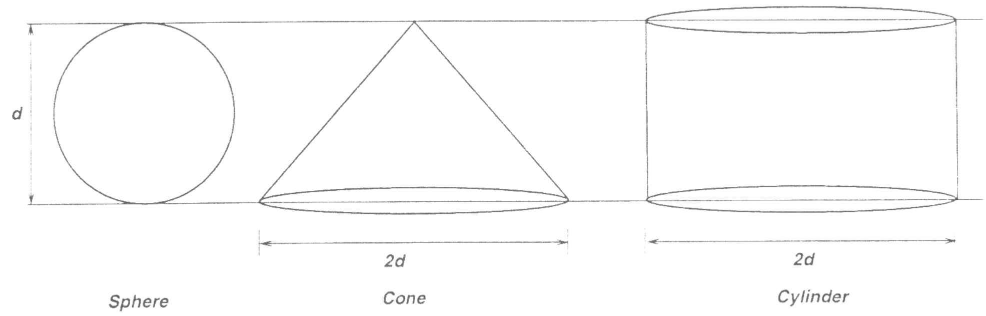


Figure 4.2: The three solids compared by Archimedes

both the cone and the cylinder is twice the height $2d$. In other words, the radius of the base and the height of the cone and the cylinder are equal to d. Hence if the cone is viewed in profile, the angles at the base are each equal to 45°.

Archimedes thought of a wonderful imaginary arrangement of the three solids. He imagined that the cone and the sphere were placed inside the cylinder. Although they were solid, be imagined that they would pass through one another like ghosts. The arrangement that he imagined is shown in Figure 4.3. The base of the cone lies exactly on the base of the cylinder and sphere is placed exactly in the middle of the cylinder just touching the top and bottom surfaces.

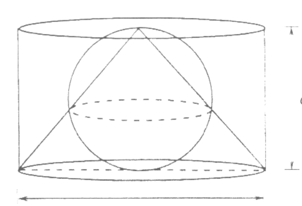


Figure 4.3: Archimedes' arrangement of the three solids

Now, following Archimedes, we will conduct an imaginary or a thought experiment. We cut the combined arrangement in Figure 4.3 into thin slices. This means that the sphere, the cone and the cylinder are all cut into slices of equal thickness. Let us number the slices 1 to n starting from the top. Let all the slices have a thickness t. Take the kth slice of the three objects. Figure 4.4 shows the geometry of the three kth slices. The three slices are all viewed from the edge in Figure 4.4 and hence they all coincide with the line MN. Let the radius of the slice of the cone be a_k and the radius of the slice of the sphere be c_k. The radius of the slice of the cylinder is d. Notice that as k increases from 1 to n, the radius of the kth slice of the cone increases, the radius of the kth slice of the sphere first increases and then

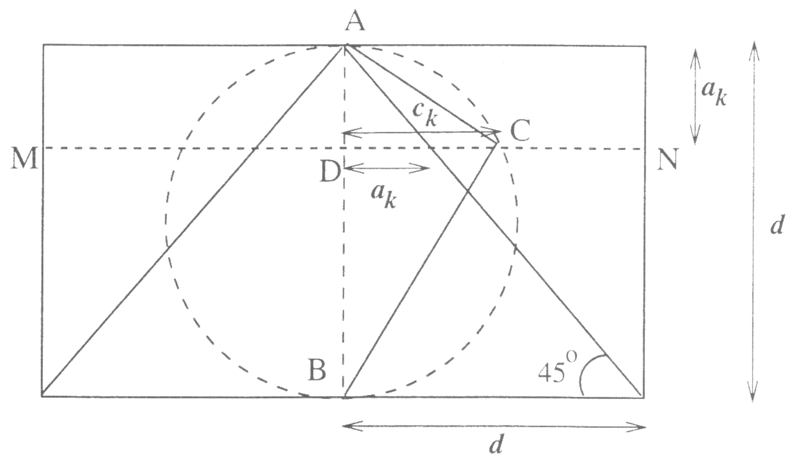


Figure 4.4: The geometry of the slices of the three solids

decreases. However all the slices of the cylinder have the same radius d.

Now look at △ABC in Figure 4.4 . It is a right angle triangle because ∠ACB is in a semi-circle. CD is perpendicular to AB. So △ADC and △CDB too are right angled triangles. The three triangles ABC, ACD and BCD are similar because they are all right angle triangles and have the same angles at the other two corners.

$$\triangle ABC \sim \triangle ACD \sim \triangle BCD$$

$$\frac{CD}{AD} = \frac{BD}{CD}$$

$$CD^2 = BD \times AD$$

CD is the radius of the slice of the sphere c_k, BD is the radius of the cone slice a_k. ∠BCD is 45°. So BD = CD = a_k. Therefore we have

$$C_k{}^2 = a_k \times (d - a_k) = da_k - a_k{}^2$$

$$\text{or } c_k{}^2 + a_k{}^2 = da_k$$

This is the equation that we get from the geometry of the kth slice. How does the balance come into picture? Here we get a glimpse of Archimedes' genius who

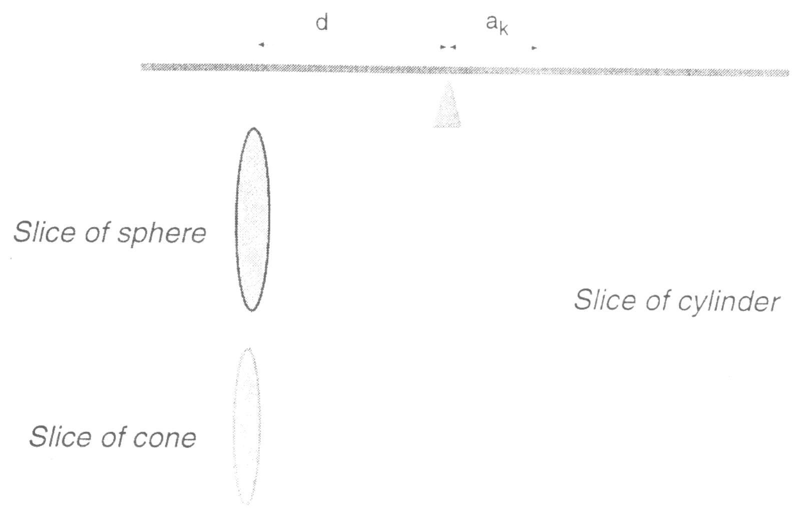


Figure 4.5: Balancing the kth slices of the three solids

interprets this equation in terms of a balance. First let us multiply both sides of the equation by πtd. Recall that t is the thickness of each slice.

$$\pi td \times {c_k}^2 + \pi td \times {a_k}^2 = \pi td \times da_k$$

Rearranging the terms we have

$$\pi {c_k}^2 t \times d + \pi {a_k}^2 t \times d = \pi d^2 t \times a_k \tag{4.1}$$

Now look at the picture of the balance in Figure 4.5. Can you interpret the meaning of the equation that we have first written in terms of the balance? You may need to assume that the cone, the cylinder and the sphere are made of the same material with uniform density. This is, of course, a reasonable assumption to make. The balance shows the kth slices of the cone and the sphere hung on one side and the kth slice of the cylinder hung on the other side. Note the distance of each slice from the fulcrum or the pivot of the balance. Will the slices balance each other? See if the equation written above helps you answer this question.

We have checked one set of the slices for balance. What would happen to the next slice, say, the $(k+1)$th set of slices? Following the pattern for the kth slice.

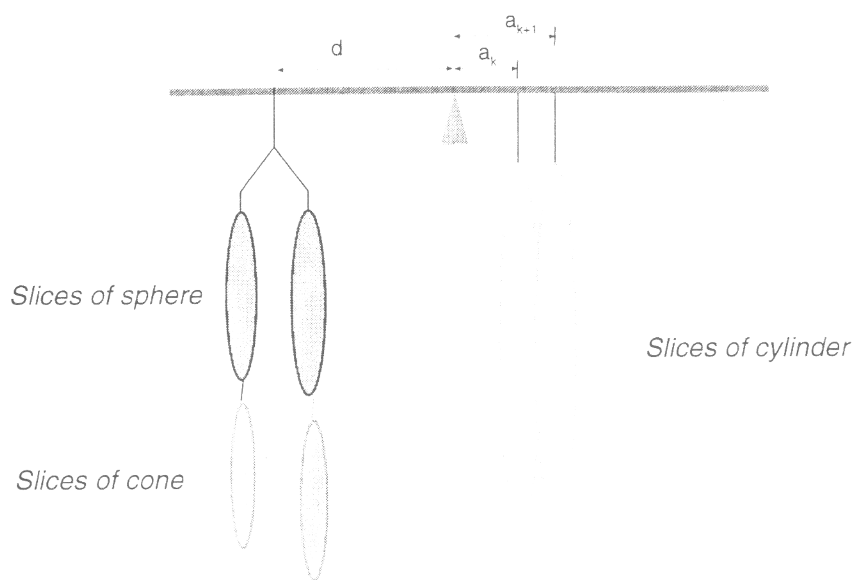


Figure 4.6: Balancing the $(k + 1)$th set of slices

we could say that the radius of the $(k + 1)$th slice of the cone and the sphere are respectively a_{k+1}, c_{k+1}. The radius of the $(k + 1)$th slice of the cylinder, of course, remains d. Now we can write down a "balance" equation for this set of slices following the pattern of Equation (4.1).

(4.2) $$\pi {c_{k+1}}^2 t \times d + \pi {a_{k+1}}^2 t \times d = \pi d^2 t \times a_{k+1}$$

Now let us hang these slices at the appropriate places on the balance as in Figure 4.6 . Which of the slices are hung at the same point as the kth slice? Which slices have moved?

We see that in order to balance the slices, the slices of the cone and the sphere are hung at the same point, a distance d, from the fulcrum. The slice of the cylinder is hung at a distance a_{k+1}. For the next set of slices, that is, the $(k + 2)$th set, the slice of the cylinder would be at a distance a_{k+2}. Therefore when we have finished hanging all the slices up on the balance, we would have

- All the slices of the cone, hung at a distance d on the left

- All the slices of the sphere hung at a distance d on the left
- Every ith slice of the cylinder hung at a distance a_i on the right

This is equivalent to hanging the entire cone and the entire sphere on the left at a distance d from the fulcrum, and hanging the entire cylinder at a distance $\frac{d}{2}$ on the right. The arrangement is shown in Figure 4.7 . Convince yourself that this arrangement is indeed right.

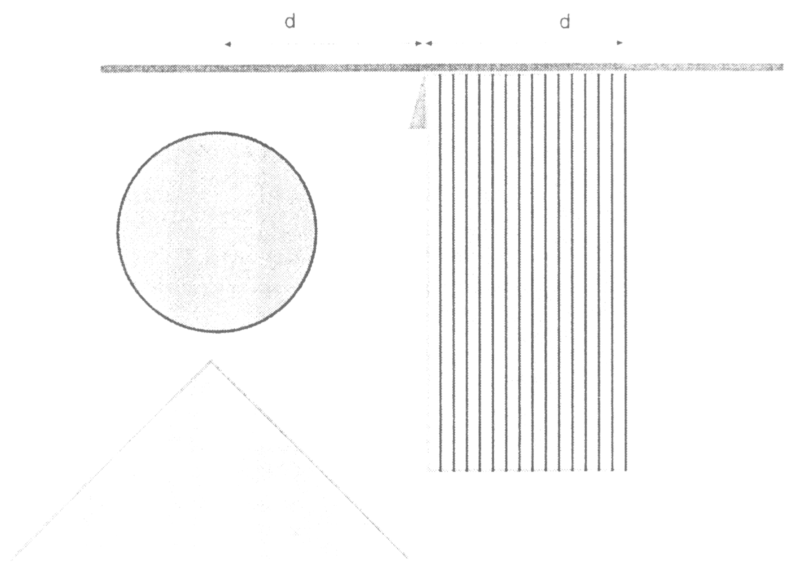


Figure 4.7: Balancing the three solids

From the arrangement on the balance we can conclude, just as Archimedes, did that

$$volume\ of\ the\ sphere\ \times\ d\ +\ volume\ of\ the\ cone\ \times\ d$$
$$=\ volume\ of\ the\ cylinder\ \times\ \frac{d}{2}$$

or

$$volume\ of\ the\ sphere\ +\ volume\ of\ the\ cone$$
$$=\ \frac{1}{2}\ \times\ volume\ of\ the\ cylinder$$

We have discussed the history of the relation for the volume of the cone in the previous chapter. Archimedes knew the volume of the cone since Democritus had determined it more than a century and a half earlier. Eudoxus had proved that the volume of a cone was $\frac{1}{3}$rd that of a cylinder of equal length and base about a century earlier. So all that Archimedes had to do was to substitute those values in the relation and obtain the volume of the sphere. Let us substitute expressions for the volumes of the cylinder and the cone.

$$volume\ of\ the\ sphere + \frac{1}{3} \times \pi d^3 = \frac{1}{2} \times \pi d^3$$

or

$$volume\ of\ the\ sphere = (\frac{1}{2} - \frac{1}{3}) \times \pi d^3 = \frac{1}{6} \times \pi d^3$$

Putting $d = 2r$, we have

$$volume\ of\ the\ sphere = \frac{1}{6} \times \pi \times 8r^3 = \frac{4}{3}\pi r^3$$

Archimedes describes this derivation of the volume of a sphere in a book called *Method*. A manuscript of this book, incidentally, was found only in 1899 in Jerusalem by the Greek palaeographer Papadopulos Cerameus. It was found on a parchment written in the 10th Century AD. The parchment on which Archimedes' book was written had been washed out in the 12th or 13th century and some prayers had been copied on to the parchment. Fortunately the washed out writing did not disappear entirely and could be read and decoded in the 20th Century. To the great excitement of historians of mathematics, the parchment contained large portions of six books by Archimedes. Of these four books were already available, but two books, which were known to exist, had no known copies. Recently, about a year ago, the parchment was auctioned at the famous auctioneer Christie's for two million dollars.

Archimedes called the method we have described above a 'mechanical method' since it was based on principle of mechanisms like the lever or balance, and did not consider it a 'proof'. He thought that it was a useful method to determine the result at first and made it easier to search for a proof. We find the actual 'proof' of the relation between the volumes of a sphere and a cylinder set out in the book 'on the sphere and cylinder'. We shall not discuss this proof here. Instead we will turn to another remarkable and successful attempt at finding the volume of the sphere in another corner of the world, nearly six centuries later.

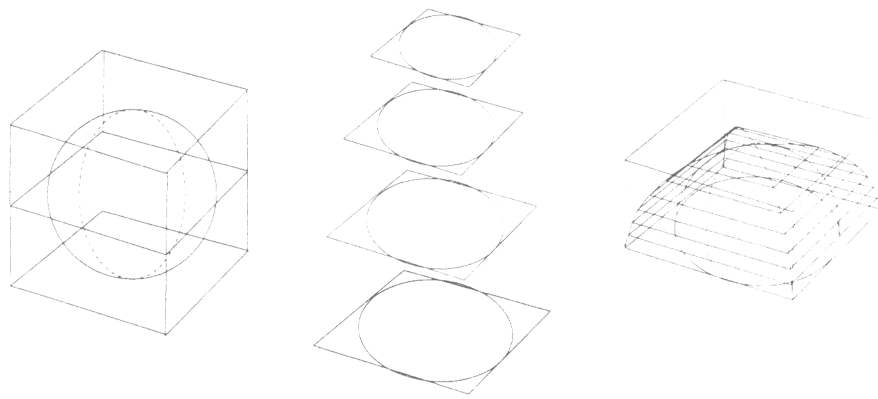


Figure 4.8: The umbrella solid

The Chinese find the volume of the sphere using an umbrella solid.

The Chinese arrived at the expression for the volume of the sphere through an entirely different and independent method. The pioneering work in Chinese mathematics on several geometric solids was done by the great mathematician of the 3rd century AD, Liu Hui. Hui wrote a commentary on the classic *Nine chapters on the mathematical Art*, the oldest extant Chinese work in mathematics, stemming from about 200 BC or earlier. Liu Hui made many important contributions to Chinese mathematics. One of the results that he established is already familiar to us – the volume of a pyramid is $\frac{1}{3}$rd the volume of a prism with the same base and height.

Liu Hui took the first steps towards a Chinese derivation of the volume of a sphere by considering a peculiar solid, the double vault or the double 'umbrella' solid. The umbrella solid is something like a mosquito net in the shape of a square umbrella used to protect sleeping infants from mosquitoes. We describe how such a solid is constructed briefly.

Imagine that a sphere is enclosed in a cube so that the sphere exactly touches the cube as in Figure 4.8(a). Consider only the upper half of the sphere and the cube. Divide the height of the hemisphere into an equal number of parts. Now consider the cross-sections of the sphere at each division of the height. Of course, each of these cross-sections is a circle. Imagine that each of the circular cross-sections are inscribed in a square. Figure 4.8(b) shows a few such cross-sections. If every circular cross-section of the hemisphere is inscribed in a square, all these squares form the cross-section of a new solid – the umbrella solid as in Figure 4.8(c). If we divide the height into a large number of equal divisions (an infinite number) the square cross-sections form a smooth umbrella solid. The double umbrella solid is made up of two umbrella solids joined at the bases.

Liu Hui obtained the relation between the volume of the double umbrella solid and a sphere. The argument was simple and used the Cavalieri principle that we encountered when we discussed the volume of a cone. Compare the umbrella solid with the hemisphere as in Figure 4.8(c). The diameter of the hemisphere is equal to the side of the umbrella solid. Cut both of these into an equal number thin slices of equal width. At each height we will then have a slice of the hemisphere, which has the shape of a square and a slice of the umbrella solid which has the shape of a square. The diameter of the circular slice is equal to the side of the square slice as can be seen in Figure 4.8 (b). The areas of the circular slices hence are $\frac{\pi}{4}$ times the area of the square slice. Since the slices are of the same thickness we can conclude that

$$volume\ of\ the\ hemisphere = \frac{\pi}{4} \times volume\ of\ the\ umbrella\ solid.$$

Liu Hui did not know how to determine the volume of the umbrella solid. He died in 280 AD without proceeding further. Two centuries later, a father and son team of mathematicians, Zu Zongshi and Zu Xuan solved the problem. They found the volume of the umbrella solid and then the volume of the sphere. Once they had the right idea, it turned out to be simple.

We considered the umbrella solid above while explaining Liu Hui's discovery. Now consider $\frac{1}{4}$th part of the umbrella, (this is $\frac{1}{8}$th of the original double umbrella.) This is shown in Figure 4.9 (a). The base of this $\frac{1}{4}$th umbrella solid or quarter umbrella is a square of side r. (r is also the radius of the original sphere in Figure 4.8.) We can imagine the quarter umbrella to be enclosed in a cube of side r.

Again we think of this quarter umbrella and the cube in which it is enclosed as made up of slices. Consider the slices at a height h from the base. The slice of the cube is a square of side r. So its area is r^2. The slice of the quarter umbrella is also a square whose side is smaller, say x. The area of this slice is x^2. The difference between the area of the cube slice and the quarter umbrella slice is an l-shaped area which is shaded in Figure 4.9 (a). This l-shape has an area equal to $(r^2 - x^2)$

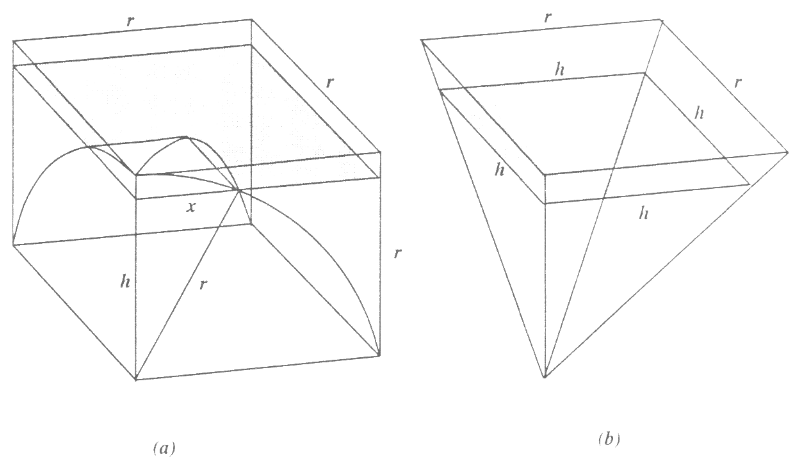


Figure 4.9: Finding the volume of the quarter umbrella solid

Applying the Pythagoras theorem to the right angled triangle which can be seen in Figure 4.9(a), we have

$$r^2 = x^2 + h^2$$

or

$$h^2 = r^2 - x^2$$

So the area of the L-shaped slice is h^2 where h is the height from the base. The total volume of these L-shaped slices is the difference in the volume between the cube and the quarter umbrella.

Now consider a square pyramid whose height is r and whose base is a square of side r. This pyramid is shown upside down in Figure 4.9 (b). Any slice of the pyramid at a height h from its vertex has an area h^2. So, using the Cavalieri principle, we can say that the volume of this pyramid is equal to the total volume of the L-shaped slices in Figure 4.10 (a). Since Liu Hui had already derived the volume of a pyramid two centuries earlier, Zu Zongshi and Zu Xuan could use this result. They found that

$$combined\ volume\ of\ the\ L-shaped\ slices$$
$$= volume\ of\ the\ square\ pyramid\ of\ base\ \ r^2\ \ and\ height\ \ r$$

$$= \frac{1}{3} \times r^2 \times r \quad = \quad \frac{1}{3} \times r^3$$

But the combined volume of the L-shaped slices is equal to the difference in the volume of the cube of side r and the quarter umbrella. Hence

$$volume\ of\ cube\ of\ side\ r - volume\ of\ quarter\ umbrella = \frac{1}{3}r^3$$

$$r^3 - volume\ of\ quarter\ umbrella = \frac{1}{3}r^3$$

or

$$volume\ of\ quarter\ umbrella = \frac{2}{3}r^3$$

Hence

$$volume\ of\ umbrella\ solid = 4 \times \frac{2}{3}r^3 = \frac{8}{3}r^3$$

But we know that

$$volume\ of\ the\ hemisphere = \frac{\pi}{4} \times volume\ of\ the\ umbrella\ solid$$

Hence

$$\begin{aligned} volume\ of\ the\ hemisphere &= \frac{\pi}{4} \times \frac{8}{3}r^3 \\ &= \frac{2}{3}\pi r^3 \end{aligned}$$

Hence,

$$volume\ of\ the\ sphere = \frac{4}{3}\pi r^3$$

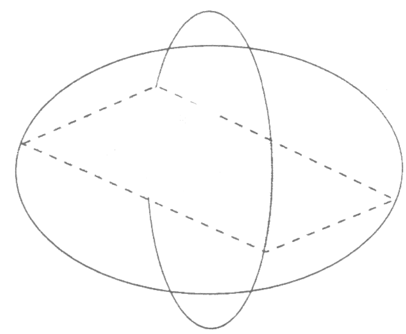


Figure 4.10: The double umbrella solid as it appears in Archimedes

What we have here therefore is a derivation which is much simpler than Archimedes derivation of the volume of a sphere. A remarkable fact is that Archimedes in fact computed the volume of the umbrella solid. In his book *Method* in which we find the use of the metaphor of the balance, we find him setting out the volume of the double umbrella solid as $\frac{2}{3}$ of the whole cube in which it is inscribed. The double umbrella solid is described interestingly as formed by the intersection of two equal cylinders perpendicular to each other as in Figure 4.10. Archimedes, of course, did not use the relation for the volume of the double umbrella solid in his derivation of the volume of a sphere.

We have explored a few of the problems relating to solid bodies that were tackled by ancient mathematicians. As we go over their arguments in detail, we come to understand the brilliant leaps that their intellects took. Today we have so much more powerful techniques through the use of calculus, that the problems themselves seem simple and easy. But clearly to those who do not have access to these tools the problems are not trivial. Moreover, even if the problems are simple, the ideas that were brought to bear on these problems are not trivial even today. Exploring the paths that were taken by the old masters lets us feel the excitement of mathematics as it is being created.

References

Books

Bhanu Murthy, T.S. *A Modern Introduction to Ancient Indian Mathematics*. New Delhi, Wiley Eastern Limited, 1992

Boyer, Carl B. and Merzbach, Utac. *A History of Mathematics*. New York, John Wiley and Sons, 1989.

Clawson, Calvin C. *The Mathematical Traveler: Exploring the Grand History of Numbers*. New York and London, Plenum Press, 1994.

Dehaene, Stanislas. *The Number Sense*. London, Penguin Press, 1998.

Dunham, William. *The Mathematical Universe*. New York, John Wiley and Sons, 1994.

Dutta, B.V. and Singh, A.N. *History of Hindu Mathematics, A Source Book Parts-I and II*. Bombay, Asia Publishing House, 1962.

Flegg, Graham. *Numbers Their History and Meaning*. London, Penguin Books, 1983.

Heath, Thomas L. *The Thirteen Books of Euclid's Elements Volume-III*. Cambridge, The University Press, 1926.

Heath, T.L. *The Works of Archimedes*. New York, Dover, (Original 1897, 1912).

Heilbron, J.L. *Geometry Civilized : History, Culture, and Technique*. Oxford, Clarendon Press, 1998.

Joseph, George Gheverghese. *The Crest of the Peacock, Non- European Roots of Mathematics*. New Delhi, Affiliated East-West Press, 1995.

Martzloff, Jean-Claude. *A History of Chinese Mathematics* Berlin, Springer-Verlag, 1997.

Nelson, Joseph and Williams. *Multicultural Mathematics*. Oxford, Oxford University Press, 1993.

Neugebauer, O. *The Exact Sciences in Antiquity*. Providence, Brown University press, 1957.

Sarasvati Amma, T.A. *Geometry in Ancient and Medieval India.* Delhi, Motilal Banarsidass, 1979.

Smith, D.E. *History of Mathematics, Volume-II.* New York, Dover Publications, 1958.

Swetz, Frank J. *From Five Fingers to Infinity: A Journey through the History of Mathematics.* Chicago, Open Court, 1995.

Wilson, A.M. *The Infinite in the Finite.* Oxford, Oxford University Press, 1995.

ARTICLES

Swetz, Frank J. 'The Volume of a Sphere: A Chinese Derivation', *Mathematics Teacher*, Vol. 88, 1995, pp. 142-145.

Eagle, Ruth. 'A Typical Slice', *Mathematics in School*, Sept., 1998, pp. 37-39.

June 1999

Homi Bhabha Centre for Science Education
Tata Institute of Fundamental Research
V. N. Purav Marg, Mankhurd
Mumbai – 400 088