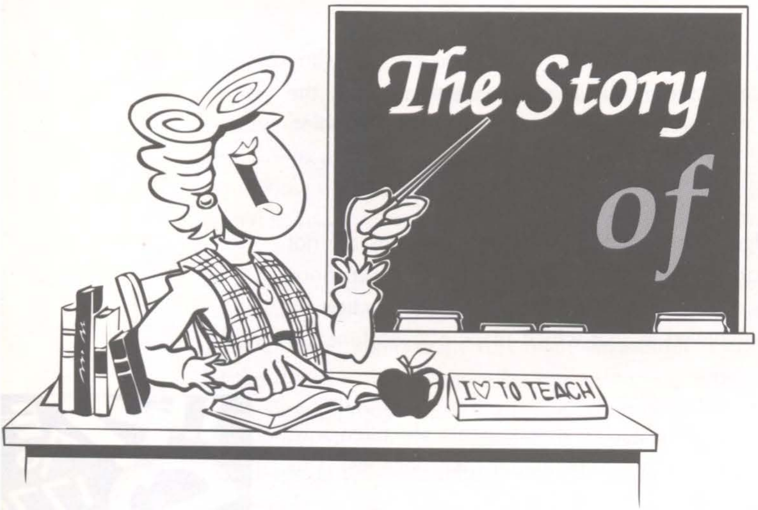

The Story of Calculus I

Dr M K Siu[*]
Translation by Dr P Y H Pang[**]

Most readers of this article have come across the term "calculus", perhaps in secondary school. Even those who have not reached the stage of studying it have probably heard about it. While certain fundamental ideas of calculus were already fermenting in the pioneering works of ancient Eastern and Western mathematicians, as a systematic discipline, calculus has a history of just over three centuries. This article is not about the development of calculus in the last three hundred years, but is the story of its birth.

Nomenclature

The term "calculus" is derived from the Latin word meaning "stone". It refers to *calculation* which, in ancient Europe, was facilitated by stones. Nowadays, the medical term "a calculous man" refers to a patient with kidney stones and not an expert of calculus.

In a way, the terminology is appropriate, as calculus was invented as a powerful tool in mathematical calculations. On the other hand, the name carries the misleading connotation that calculus is just concerned with mechanical computations. This is furthest from the truth. Calculus is one of the greatest achievements, not only in the history of mathematics, but in all of human civilization, with far-reaching consequences in science, technology and philosophy. Today, calculus is an indispensible mathematical tool in many fields, including not only the physical sciences, material sciences and engineering, but also the biological sciences, social sciences and business administration.

Two major historic triumphs of calculus are well-known to all, but nevertheless worth repeating. First, in the mid-17th century, Newton used his newly formulated calculus to expound his theory of mechanics, which explained the movements of the heavenly bodies. This work opened a new chapter in the investigation of the physical universe. Obviously impressed by this important work, the 18th century poet Pope wrote:

Nature and nature's laws
lay hid in night,
God said, "Let Newton be",
and all was light.

Second, in the mid-19th century, Maxwell gave the theory of electromagnetism a mathematical treatment and summarized it in the famous Maxwell (differential) equations. In this work, the existence of electromagnetic waves was predicted, leading to their discovery by the experimental scientist Hertz in twenty years' time. The potential of electromagnetic waves was finally realized when Marconi invented wireless communication thirteen years later.

The Chinese terminology for calculus (weijifen) first appeared in 1859 in a translation of *Analytic Geometry and Calculus* (written by the American mathematician Loomis in 1850) by Li Shanlan and the Englishman Wylie. In the preface, Li wrote "This book deals first with algebra (meaning analytic geometry, known as algebraic geometry at the time) and subsequently differential and then integral calculus, following the order of complexity of the topics, like climbing a staircase. Hence we have given the book the title *Dai Wei Ji Shi Ji* (algebra, differential and integral calculus in ascending steps)." He continued thus "During the time of the Emperor Kangxi, in the West, Leibniz and Newton invented the arts of differential and integral calculus ... which are based on the principle that all planar figures are built up from small to large. Every instantaneous increment in area is called a differential (*wei fen*), and the total area is called an integral (*ji fen*)." This is the origin of the Chinese terminology.

Volume Calculations in Ancient Greece

Calculations of areas and volumes have appeared in ancient Eastern and Western mathematical literature going back thousands of years, but it was not until the 4th and 3rd centuries BC when such formulae were given mathematical proofs in Greece. The foremost contributor in this enterprise was Archimedes, who lived in the 3rd century BC. His deductive reasoning is considered rigorous even by today's standards. However, his greatness lies in his ingenious usage of intuition and conjectures, and application of other fields to establish deep results in mathematics.

To establish his area formulae, he used the method of "exhaustion", which was based on the work of the Greek mathematician Eudoxus who lived one century before him. Eudoxus observed that if a magnitude was reduced by at least half, and the remainder reduced by at least half, and so on, the remainder could be made arbitrarily small after sufficiently many steps.

The central idea of the method of "exhaustion" can be illustrated by the proof of the following simple assertion which is the second theorem in Volume 12 of Euclid's *Elements*: The ratio of the areas of two circles is equal to the ratio of the squares of their radii. In other words, if a circle has radius d and area a, and a second circle has radius D and area A, then

$$a / A = d^2 / D^2. \qquad (*)$$

Effectively, this observation gives the formula for the area of a circle. The proof given in *Elements* goes as follows: Suppose the equality (*) does not hold. Then either the left hand side (LHS) is larger than the right hand side (RHS) or vice versa. If the LHS is larger, that is, if $a / A > d^2 / D^2$, pick an $a_1 < a$ such that $a_1 / A = d^2 / D^2$, and denote $a - a_1$ by e. Consider successive regular N-gons inscribed in the circles, with N doubling at each step; for example, start with $N = 3$, i.e., equilateral triangles, followed by regular hexagons ($N = 6$), etc. Denote by $p(N)$ and $P(N)$ the areas of these regular N-gons, inscribed in the circles of radii d and D respectively. Then it is easy to see that $p(N) / P(N) = d^2 / D^2$, and hence $p(N) / P(N) = a_1 / A$. Each time such an inscribed polygon doubles in the number of sides, the difference between its area and that of the circumscribing circle reduces by more than half (an exercise left to the readers). Therefore, by Eudoxos' Principle, after a certain number of steps, a value N is arrived at such that $a - p(N) < e$, that is, $p(N) > a_1$ and hence $P(N) > A$. This statement that the inscribed polygon has a larger area than the circumscribing circle is of course absurd, showing that the original assumption that LHS is larger is flawed. Similarly we can rule out RHS being larger.

In fact, the concepts of the infinitesimal and limit in calculus were already alluded to in Eudoxos' Principle, albeit disguised in a language involving finitely many steps, thus hiding the essentiality of the infinite. The fact that calculus was not born earlier was due in a large part to this attitude of ancient mathematicians that the infinite ought to be avoided. Nevertheless, the rigour demonstrated by mathematicians over 2000 years ago is well worth our respect.

Using this technique, Archimedes established many area and volume formulae. However, one mystery remained. One could prove that a formula was the right one by this method if the formula was indeed the right one, but how was one to come up with the right formula in the first place? Was Archimedes divinely inspired? This mystery was finally unveiled in 1906 by Heiberg, a German scholar who specialized in ancient Greek mathematics. He found a parchment in a monastery in Constantinople with prayers from the 13th century. However underneath the prayers, some other writing could barely be discerned. Through Heiberg's extreme care and persistance, it was finally revealed that it was a 10th century copy of a missing manuscript of Archimedes. It was a letter to the mathematician Eratosthenes explaining how he discovered the area and volume formulae. This precious document is now known as "The Method".

Archimedes explained that he first put the geometric object on one side of a hypothetical balance. The geometric object was viewed as being made up of infinitesimal cross sections, which he would move, one by one, from one side of the balance to the other until the two sides balanced. The area or volume formulae were then calculated by the Principle of Moment of Force.

To illustrate this, let us take a round ball of radius R, a circular cylinder of radius R and height $2R$, and a circular cone of base radius $2R$ and height $2R$, and put them on side A of a balance as shown in diagram 1a. Measuring from the fulcrum O of the balance, remove a cross section of a very small thickness L at a distance x (see diagram 1a). Since L is very small, we may suppose that the volumes of the cross sections of the ball, cylinder and cone are respectively $\pi x(2R - x)L$, $\pi R^2 L$, and $\pi x^2 L$ (why?). Take the cross sections from the ball and the cone and move them to side B of the balance, at a distance $2R$ from O. From the theory of mechanics, we can calculate the torque from side B to be $2R[\pi x(2R - x)L + \pi x^2 L] = x(\pi(2R)^2 L)$. Thus, if we double the radius of the cylinder on side A, the two sides will balance. By moving the ball and the cone to the other side of the balance cross section by cross section, they will eventually balance the cylinder with the radius doubled on the other side (see diagram 1b).

Thus, denoting the volumes of the ball, the original cylinder, the enlarged (radius doubled) cylinder, and the cone by V_B, V_C, V_D and V_N respectively, we obtain the formula that

$$2R \times (V_B + V_N) = R \times V_D.$$

Since

$$V_N = 1/3 V_D,$$

it follows that

$$V_B = 1/6 V_D.$$

Furthermore, as $V_D = 4V_C$, we conclude that

$$V_B = 2/3 V_C.$$

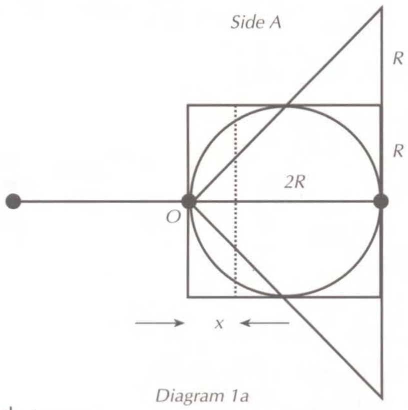

Diagram 1a

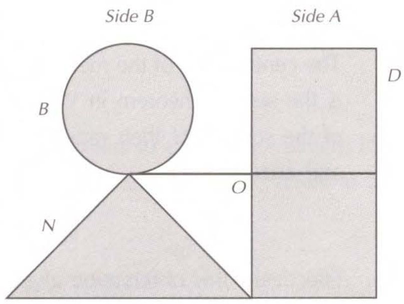

Diagram 1b

In addition, Archimedes determined that the surface area of the sphere is two-thirds that of the (original) cylinder. He was so proud of this formula that he decided that the geometric figure depicting the ball B contained in the cylinder C (see diagram 2) was to be his epitaph.


Diagram 2

Archimedes' death was a tragedy to the world. There was a legend that when the Roman general Marcellus took the city Syracuse in south Italy in which Archimedes lived, the soldiers found an old man in a room drawing pictures in sand. When the soldiers ruined the drawings, the old man lost his temper and was subsequently killed by the soldiers. The old man was Archimedes. Another legend had it that Marcellus was much impressed by Archimedes' genius and wanted to meet him. When the soldiers went to fetch Archimedes, he was completely absorbed in his problem-solving and refused to go, thereupon the soldiers lost their cool and had Archimedes killed.

Alas, after two thousand years, Archimedes is still being remembered, but who still remembers Marcellus? To finish the legend, when he learned of Archimedes' death, Marcellus was so remorseful he erected a monument on which Archimedes' beloved geometric shapes were engraved. As time went by, people forgot about this, until it was rediscovered in 1965 when the site was acquired for hotel development.

Volume Calculations in Ancient China

To continue with our story, but now moving to the Far East, the earliest complete and systematic mathematical treatise in China is *Jiu Zhang Suanshu* (Nine Chapters on the Mathematical Art), which contained results in mathematics up to the Han dynasty (206 BC - AD 220), including, without proof, many area and volume formulae.

During the time of the Three Kingdoms (AD 220 – 265), Liu Hui provided explanatory notes for *Jiu Zhang Suanshu*. The fifth chapter, entitled "Shang Gong" (discussing work), dealt with engineering mathematics, but in fact contained mainly volume calculations. Among the problems discussed was one on the volume of a "yang ma", which was the building term at the time meaning the pyramid with rectangular base. Liu Hui wrote: "Bisecting a cube along the diagonal yields two 'qian du'. Each 'qian du' can further be divided along a diagonal into a 'yang ma' and a 'bie nao', at the constant ratio of 2 : 1 in volume." "Qian du" is the term for the wall around a moat, here meaning a prism whose cross sections are right-angled triangles. "Bie nao" is the term for a particular tortoise bone, here meaning a pyramid with right-angled triangular base (see diagram 3). Thus, according to Liu Hui, the volume of a "yang ma" is twice that of a "bie nao", whereas two "yang ma" and two "bie nao" together make one cube. Hence, the volume of a "bie nao" is one-sixth that of a cube, and the volume of a "yang ma" is one-third that of a cube, giving the formula that the volume of a "yang ma" is equal to one-third the product of the height and the base area.

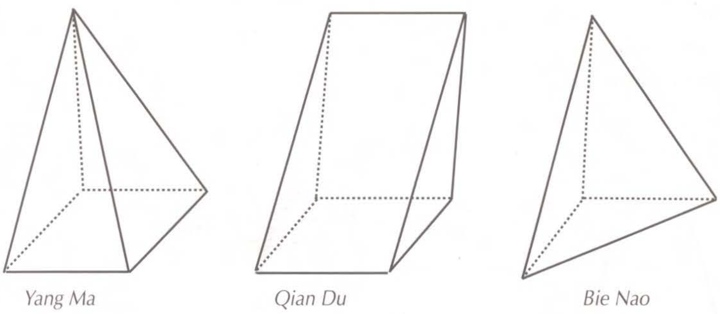

Diagram 3

More than one thousand and six hundred years after Liu Hui, the German mathematician Hilbert raised the following famous problem: Is it possible to re-arrange a subdivision of a polyhedron to obtain another polyhedron of the same volume? The answer is that this is not always possible. The solution of this problem has a profound mathematical meaning, which, put simply, is that calculus is necessary for volume calculations.

So was calculus used in Liu Hui's derivation? The answer was in fact yes, but probably Liu Hui himself was not aware of it, as it was quite beyond the level of mathematical sophistication of his time. As a matter of fact, after the passage quoted above, Liu Hui's explanations continue, and they can be summarized as follows: Divide the "yang ma" into two smaller "yang ma" and four "qian du", and divide the "bie nao" into two smaller "bie nao" and two "qian du". Put aside all the smaller "yang ma" and "bie nao", then the remaining parts of the original "yang ma" and "bie nao" are in the ratio of 2 : 1 in volume. Repeat this process to each of the smaller "yang ma" and "bie nao". Upon iteration of this process, the "yang ma" and "bie nao" will get smaller and smaller. As they become infinitesimal, they become "formless and thus negligible". This explains why the volumes of the original "yang ma" and "bie nao" are in 2 : 1 ratio. Of course, by today's standards of rigour, "becoming formless and thus negligible" is not quite acceptable, but there is no doubt that Liu Hui has captured the basic idea of the infinitesimal.

In the last section, we saw how Archimedes treated geometric objects as being made up of infinitesimal cross sections. This point of view was also often employed by Liu Hui in the form of the following principle: If the cross sections at the same height of two objects have a constant ratio in area, then the volumes of these two objects are also in the same ratio. An interesting example occurred when he corrected a mistake in the original text in chapter four, entitled "Shao Guang" (short width). The original text stated that the ratio between the volumes of a round ball and its circumscribing cube was $\pi^2 : 4^2$ (in the original text, 3 was used as an approximation to π). It had been known that the ratio between the area of a circle and its circumscribing square was $\pi : 4$, and therefore the volumes of a circular cylinder and its circumscribing cube were in the same ratio. The above mistake stemmed from the misconception that the ratio between the volumes of a round ball and its circumscribing circular cylinder was also $\pi : 4$.

This error was pointed out by Liu Hui. In fact, he stated that $\pi : 4$ was the ratio between the volumes of the round ball and the object which he called "mou he fang gai". Visualize the round ball as being made up of cross sections of circles, from a point at the north pole increasing in size to the equator and then decreasing back to a point at the south pole. Each circle has a circumscribing square. The object formed by these circumscribing squares is a "mou he fang gai". It is also the intersection between two identical circular cylinders which are placed such that their axes intersect perpendicularly. Thus, if one knew the volume of the "mou he fang gai", the volume of the round ball could be inferred. However, Liu Hui failed in his attempt to evaluate the volume of the "mou he fang gai", and he wrote: "Due to the extremely intricate interaction between the circle and the square, I fail to obtain the answer. I leave this problem to a more able person. It is not right for me to make irresponsible comments." His frankness, humility and integrity are indeed remarkable.

After about two hundred years, during the North-South Period (AD 420 – 589), the father and son mathematician team Zu Chongzhi and Zu Geng solved this problem in the following ingenious way: Divide the "mou he fang gai" into eight equal parts and put one part in a cube with side R (where R is the radius of the round ball). Now the volume of the space between the cube and the 1/8 "mou he fang gai" inside can be computed. Note that if we take a cross section of this space at a distance h from the bottom of the cube, its area is given by the Pythagoras Theorem to be $R^2 - x^2 = h^2$ (see diagram 4). Now take an upside down pyramid of height R with a square base (now on top) with side R. Its cross section at a distance of h from the bottom (the tip of the pyramid) also has an area of h^2. Therefore, the space between the cube and 1/8 of the "mou he fang gai" has the same volume as the pyramid. Now the pyramid is 1/3 of a cube with side R, therefore, the volume of the pyramid is $(1/3)R^3$, from which we deduce that the area of the "mou he fang gai" is $(2/3)D^3$ where D is the diameter of the round ball. Finally, we conclude that the volume of the round ball is

$$\pi / 4 \times 2/3D^3 = 1/6\pi D^3.$$

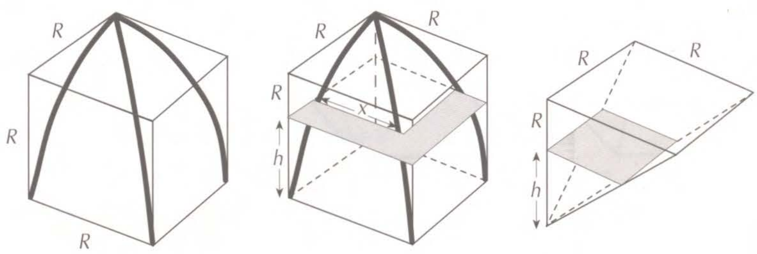

Diagram 4

It is a pity that their book *Zhui Shu* has been lost since the North Song dynasty (AD 960 – 1126) and very little is known about it. Our account of the above calculation is based on explanatory notes on *Jiu Zhang Suanshu* written by the Tang dynasty (AD 618 – 907) mathematician Li Chunfeng.

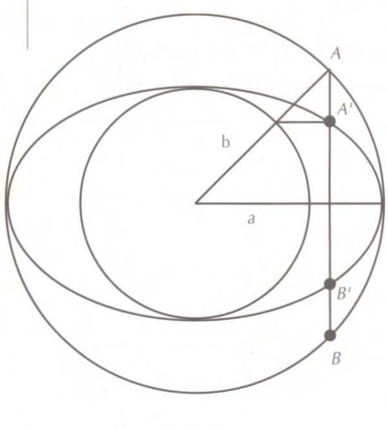

Diagram 5

The arguments by the Zu's were based on the following principle: Given two objects whose cross sections at the same height have the same area, then the two objects have the same volume. This is known in the West as Cavalieri's Principle, after the Italian mathematician who published it in 1635 and used it to establish many volume formulae. In fact, even in the West, this principle had been commonly used before the time of Cavalieri. For example, in the early 17th century, the German astronomer Kepler used it to find the area of an ellipse. He first put an ellipse of semi-major axis a and semi-minor axis b in a circle of radius a (see diagram 5). Then he observed that the ratio between the lengths of $A'B'$ and AB is $b : a$. As AB formed the cross sections of the circle, and $A'B'$ formed the cross sections of the ellipse, he concluded that the area of the ellipse was πab, using the knowledge that the area of the circle was πa^2.

Editor's Note: This is a translation of an article (in Chinese) by Dr M K Siu, which forms Chapter 2 of his book One, Two, Three and Beyond, published by Guangdong Jiaoyu Chubanshe, 1990. The Singapore Mathematical Society wishes to thank Dr Siu for allowing this translation to be published in the Mathematical Medley. Due to the length of the article, it will be published in two parts; the second part will appear in the next issue of the Medley.

**Dr M K Siu is Reader of Mathematics at the University of Hong Kong and is the author of numerous popular mathematics books.*

***Dr P Y H Pang is Senior Lecturer in Mathematics at the National University of Singapore.*

The translator would like to express his thanks to Professor Lam Lay Yong for her helpful advice.