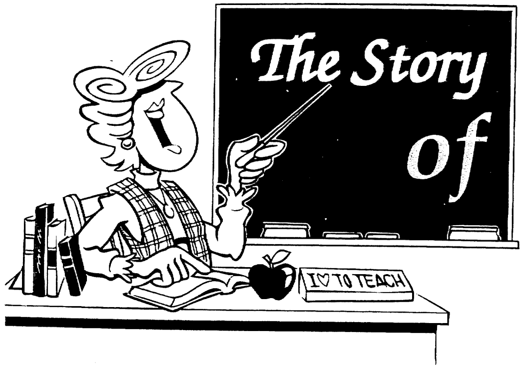
Mathematical Medley Vol. 23, No. 1 (1996), 28-32




# 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

$$V_B = 2/3\, V_C.$$

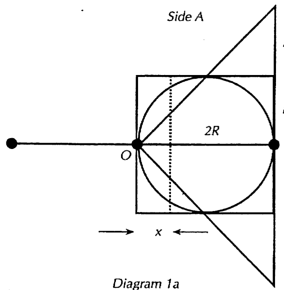


Diagram 1a

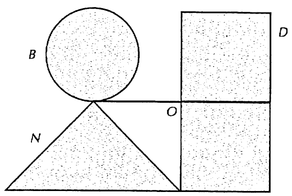


Diagram 1b

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.

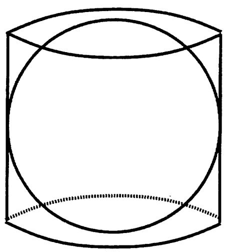


Diagram 2

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.

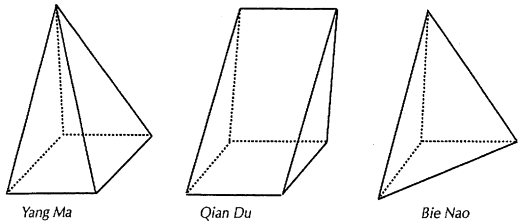


Diagram 3

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 $\pi$). 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

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

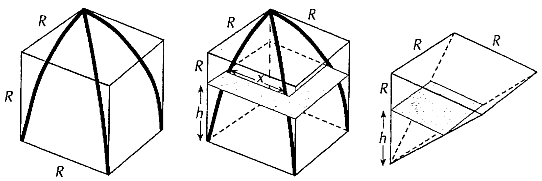



Diagram 4

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.

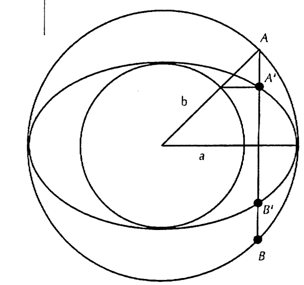



Diagram 5

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 $\pi ab$, using the knowledge that the area of the circle was $\pi a^2$. M

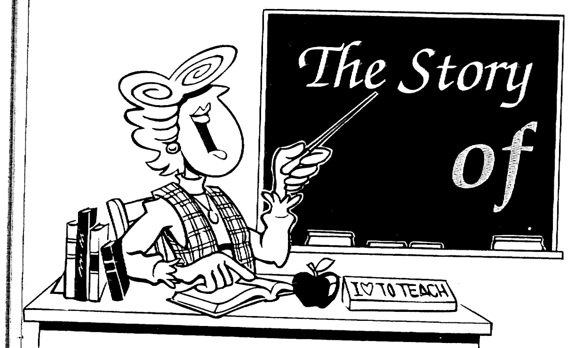
*Editor's Note: This is a translation of an article (in Chinese) by Dr M K Siu, which forms Chapter 2 of his book One, Two, Three and Beyond, published by Guangdong Jiaoyu Chubanshe, 1990. The Singapore Mathematical Society wishes to thank Dr Siu for allowing this translation to be published in the Mathematical 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.*



Mathematical Medley Vol. 23; No. 2 (1996), 51–55.




# The Story of Calculus II[1]


**Dr. M. K. Siu[2]**
**Translation by Dr. P. Y. H. Pang[3]**


*In the last episode, we mentioned the work of the father and son mathematician team Zu Chongzhi and Zu Geng during the North-South Period (AD 420—589) and how one of their results (now known as Cavalieri's Principle) was rediscovered in Europe about one thousand years later and was used by Kepler to calculate the area of an ellipse. So why was Kepler interested in the area of an ellipse?*

## 4. The Rise of Mathematics in Western Europe

Two millenia prior to Kepler, the ancient Greeks had discovered ellipses. However, their interest in them was purely as an exercise in geometry; it was Kepler who brought the ellipse into the realm of practical science. Curiosity in the mysteries of the universe led Kepler to Tycho Brahe, the leading astronomer of his time, to be his assistant. Upon the sudden demise of Tycho Brahe two years later, Kepler inherited his position, and with it the richest and most elaborate depository of astronomical data in Europe at the time, and Kepler was determined to find the laws that governed the universe from this wealth of data. Kepler kept to the highest standards of empiricism, rejecting one theory after another only for minute discrepancies with the data. Finally in the year 1609, he published two celebrated Laws:

(I) The planets follow elliptic orbits around the sun, with the sun as a focus;

(II) the straight line segment joining the sun and the planet sweeps out equal areas in equal time intervals.

Ten years later, he published the Third Law:

(III) The square of the period of the planet's revolution around the sun is proportional to the cube of the semi-major axis of the elliptic orbit.

While the Polish astronomer Copernicus had made the ground-breaking proposal in 1543 that planets revolved around the sun, he was unable to break completely from ancient Greek influences and believed that they did so in circular orbits with constant speed. Kepler's discoveries not only turned the world to the investigation of non-circular and non-uniform motions, they contributed directly to the development of calculus. It was in order to understand Kepler's Three Laws better that led Newton to his theory of mechanics, for which calculus was a prerequisite.

[1] The first part of this article appeared in the last issue of the Medley.

[2] Reader, the University of Hong Kong

[3] Senior Lecturer, National University of Singapore

Contemporarily, the Italian physicist Galileo Galilei also played a key role in the development of calculus. Basically, he initiated the "mathematization" of science. He believed that scientific investigation must start with natural phenomena. From the intricacies of natural phenomena one extracts the basic concepts of matter and motion. By using mathematics, new results can be deduced from these basic concepts, and the new results must then be verified by laboratory experimentation. Thus mathematics is an inherent partner of science. To quote a famous saying Galileo made in 1610: "The mysteries of the universe are all written in this great book in front of us, and in order to understand it, we must first learn the language in which it is written ... this language is mathematics."

As a result of the Renaissance of the 15th century, Europe in the 17th century was imbued with the keen desire to understand nature. This desire stimulated the thinking and widened the horizon of people. At the same time, advances in industry and commerce brought new and pressing scientific and technological problems. Specifically, these problems involved variable quantities which could not be tackled with the "static" mathematics of the time. Naturally, these "dynamical" challenges attracted most of the first-rate minds of the day.

### 5. Differentiation and Integration

Let us first consider an important concept, namely, rate of change. For simplicity of exposition, let us take the example of the rate of change of distance versus time – in other words, speed. First, we must distinguish between two types of speed: average speed and instantaneous speed. Average speed is a simple concept: it is the quotient of the distance covered in a certain time interval divided by the length of that time interval. But what is often of more interest is instantaneous speed. For example, when a bullet hits a target, the average speed of the bullet does not tell much, it is the instantaneous speed at impact which carries the force! How can we understand instantaneous speed? Suppose a bullet is moving forward, and between the time instant $T$ and $T + 0.5$ sec it has travelled 207.5 m. Then the average speed for this time period would be 415 m/s. If we shorten the time of observation to between $T$ and $T + 0.1$, and suppose the bullet has travelled 41.9 m in this time interval, then the average speed would be 419 m/s. Repeating this now for the time interval between $T$ and $T + 0.01$, we may observe that the distance covered becomes 4.199 m, giving an average speed of 419.9 m/s. Keep on shortening the time interval this way, the average speed will tend towards a certain number. This number should be the instantaneous speed of the bullet at time $T$. You may argue: "If an instant is a time interval of length zero, how can the bullet change position in an instant and how can there be any (instantaneous) speed?" Indeed, this is an archaic argument first put forward by the ancient Greeks in the 5th century BC. Let us not get bogged down by its debate – anyone who has run into an object and got bruised would not doubt the evidence of instantaneous speed.

The following principle was observed by Galileo: When an object falls freely from rest, the distance fallen is directly proportional to the square of time. When a quantity varies with another quantity, this relationship is called a function in mathematics. Let's denote by $S$ the distance fallen, and by $T$ the time in which the fall takes place, then we say that $S$ is a function of $T$. If we measure $S$ in metres and $T$ in seconds, then Galileo's principle can be expressed as $S = 5T^2$ (5 is an approximate value). We may ask what is the object's instantaneous speed after 5 seconds. (Incidentally, the average speed for the first 5 seconds is 25 m/s as the object falls 125 m.) The instantaneous speed can be worked out as follows: We observe that when $T = 5$, $S = 125$; and when

$$T = 5 + h,\ S = 5(5 + h)^2 = 125 + 50h + 5h^2. \qquad (1)$$

Thus the average speed in this time interval of length $h$ is $50 + 5h$ m/s. As $h$ gets smaller and smaller, the average speed tends to 50 m/s. Thus, the instantaneous speed of the free-falling object five seconds after its release is 50 m/s. Your attention is drawn to the following very important fact: It seems that the instantaneous speed is obtained by putting $h = 0$, in (1); however, this does not concur at all with the reasoning we have just given. Indeed, if we put the length of the time interval $h = 0$, then the difference $S$ would also be 0, in which case when we calculate the average speed we would get 0/0, which of course is meaningless.

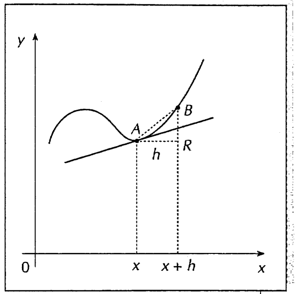


Diagram 1

Our consideration of the instantaneous speed above in fact contains the essential concepts of calculus. In a more general setting, if $y$ is a function of $x$, i.e., $y$ is a quantity which varies as $x$ varies, then the instantaneous rate of change of $y$ with respect to $x$ is called the *derivative* of $y$ and the method for its computation is called *differentiation*. As another example, consider $y$ being atmospheric pressure and $x$ being height, then the derivative of $y$ with respect to $x$ measures the instantaneous rate of change of atmospheric pressure with respect to height (note that here the word "instantaneous" does not refer to a time instant but a height level). The usual mathematical notation for this derivative is $\frac{dy}{dx}$, and geometrically it can be described as follows: Consider the graph of the function $y$ with respect to $x$ in the usual $x$ - $y$ coordinate system. Note that the average rate of change between $x$ and $x + h$ is the slope of the secant joining the points $A$ and $B$ as indicated in Diagram 1.

As we shorten $h$ repeatedly, the secant approaches the tangent $R$ to the graph at $x$ (see Diagram 1). Thus, the instantaneous rate of change is the slope of the tangent.

Computing the slope of a tangent was an important area of research in the 17th century. Around 1630, the French mathematician Fermat proposed a systematic method for their computation, which was basically the method we described in the last paragraph. In this way, Fermat computed the slopes of the tangents to the curves $y = x^2$, $y = x^3$, $y = x^4$, etc., and found that at the point $x = a$, they were $2a$, $3a^2$, $4a^3$, etc., respectively. At the same time, Fermat made important contributions to the computation of the area under a curve. He considered a lot of very narrow rectangles under the curve (see Diagram 2). The sum of their areas would then give an estimate of the area under the curve from below. As these rectangles became narrower and narrower (and thus more and more of them would be needed), the sum of their areas would tend to be the actual area under the curve. Using this idea, Fermat computed the areas under the graphs $y = x^2$, $y = x^3$, $y = x^4$, etc., and found that between $x = 0$ and $x = a$, they were $a^3/3$, $a^4/4$, $a^5/5$, etc., respectively. This is in fact the fundamental concept of integral calculus. At this point, the groundwork for differential and integral calculus had been laid.

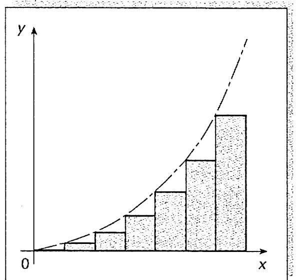


Diagram 2

Judging from Fermat's calculations, it is hard to believe that he was not already aware of the relationship between differential and integral calculus. However, in the mathematical literature, the first to note the connection seems to be the English mathematician Barrow, who was Newton's teacher at Cambridge University. This was later clarified by Newton and the German mathematician Leibniz, and is known today as the *Fundamental Theorem of Calculus*. Very roughly, this theorem says that if we first integrate a function and then differentiate the integral, we recover the original function. To be more precise, let $f$ be a continuous function on the closed interval $[a, b]$, and let

$$F(x) = \int_a^x f(t)\,dt,$$

then $f$ is the derivative of $F$. In fact, today, most people compute an integral by using this theorem – first, find the indefinite integral (also called the primitive, which is a function whose derivative is the given one) for which there is a host of methods, then evaluate it at the upper and lower limits and take the difference. For example, for $f(x) = x$, an indefinite integral is $x^2/2$, and thus

$$\int_a^b x\,dx = \frac{b^2}{2} - \frac{a^2}{2}$$

However, we must bear in mind that this is a *technique* for computing an (definite) integral and is not its *definition*. An (definite) integral is, by definition, the limit of an infinite sum (see explanations given in the last paragraph). It must be pointed out that the integration technique using the Fundamental Theorem of Calculus is in practice rarely effective, and in some cases even impossible as was demonstrated by mathematicians in the first half of the 19th century, except for certain simple problems or textbook exercises! By formalizing these concepts, differential and integral calculus turn out to be much more versatile than just computing speed or area.

## 6. The Work of Newton and Leibniz

When we mention calculus, the names of the Englishman Newton and the German Leibniz immediately come to mind. These two outstanding scholars of the 17th century are accredited with the invention of calculus by most textbooks. It is most unfortunate that their concurrent but independent founding of their own schools of calculus led to an ugly plagiarism scandal, which resulted in a century of no exchange between British and continental mathematicians. Because of this, British mathematics did not advance for a century! Mathematicians must heed this historical lesson. The most regrettable aspect of this dispute was that it was actually instigated by an "outsider", the Swiss mathematician de Duillier, who held a grudge against Leibniz. As it intensified, the dispute became a display of nationalistic bigotry, and the protaganists could not escape involvement. In fact, earlier, Leibniz had very high esteem for Newton, and was quoted as saying that Newton's contribution to mathematics had been the better half of the sum of all before him.

This dispute of precedence is absolutely pointless as numerous scholars had made contributions to the developement of calculus in the 17th century. It is fair to say that by the mid-17th century, all the groundwork had been prepared and the time was ripe for calculus to make its grand entrance; Newton and Leibniz compiled the knowledge and put on the final touches. This assessment does not in any way lessen the important contributions by these two great mathematicians. We must bear in mind what era they lived in. Some people have called the 17th century "the age of geniuses". The "geniuses" came about because they exceeded others in observational powers, knowledge, diligence and commitment. Thus they were able to make full use of the opportunities of their times to make the maximum contributions. Newton had said: "If I have seen further than [others], it is because I have stood on the shoulders of giants." His secret of success was that he never stopped thinking. It was usual for him to work 18–19 hour days and he had spent many sleepless nights in the laboratory.

Calculus without Newton and Leibniz is like the Trojan War without the wooden horse. In view of the fact that the groundwork had been laid before, what then really were the contributions of these two great men?

First, they put calculus into a systematic framework. According to Newton, his work on calculus was done during the period 1665–1666; however, distaste for publication led to the delay of its dissemination. His major publications include *De analysi per aequationes numero terminorum infinitas* (completed in 1669 but published in 1711), *Methodus fluxionum et serierum infinitorum* (completed in 1671 and published in 1742), *De quadratura curvarum* (completed in 1676 and published in 1693). His magnum opus, though, is his treatise *Philosophiae naturalis principia mathematica*, in which he expounded his theories of calculus as well as mechanics. This treatise was completed later than the three mentioned above but was the first in print (1687), thanks to his good friend and fellow mathematician Halley, who dug into his own pockets to see its publication through.

On the other hand, Leibniz "discovered" calculus independently in 1676, but published his findings earlier than Newton. His first paper on differentiation appeared in a German journal in 1684, followed in 1686 by a paper on integration. He had originally used the terms *Calculus Differentialis* and *Calculus Summatorius*, which the Swiss mathematician James Bernoulli later changed to *Calculus Integralis*. This is the origin of the terminology in English. Before Newton and Leibniz, calculus consisted of a collection of "clever tricks" proposed by many different people. It was Newton and Leibniz who organized them into a systematic, widely-applicable discipline with standardized notation. Before that, its use had been extemporaneous – every problem required a different "trick" and the hard work of many first-rate minds. We must thank Newton and Leibniz if only for this.

Their (especially Newton's) second main contribution was that they brought "infinity" into the realm of mathematics. Newton observed that the infinite sum could be viewed as an extension of ordinary algebraic operations. As an example, his celebrated Binomial Theorem was an extension of the binomial theorem in algebra. For a natural number $N$, it had been known that

$$(A+B)^N = \binom{N}{0} A^N + \binom{N}{1} A^{N-1} B + \dots + \binom{N}{N-1} AB^{N-1} + \binom{N}{N} B^N, \qquad (2)$$

where $\binom{N}{r}$ denotes the binomial coefficient which is the number of ways one can choose $r$ objects out of $N$ different ones. For a number $N$ which is not a natural number, Newton showed that

$$(A+B)^N = A^N + NA^{N-1}B + \frac{N(N-1)}{1\cdot 2} A^{N-2}B^2 + \frac{N(N-1)(N-2)}{1\cdot 2\cdot 3} A^{N-3}B^3 + \dots \qquad (3)$$

Of course, today we can see that the formula (3) is an obvious generalization of (2) by writing out the binomial coefficients. (Note that the illusion that the infinite sum (3) is merely a straightforward extension of the finite sum (2) is fallacious.) However, this is far from the way of thinking during Newton's time; he derived the formula (3) while considering an integration problem. On the other hand, the discovery of this formula allowed him to compute the derivative of $x^n$, and it was through this sort of connection that he finally came upon the Fundamental Theorem of Calculus. The incorporation of "infinity" into mathematics broke the Hellenic tradition of avoiding infinite processes and began a new chapter for mathematics.

In the subsequent two centuries, many advances in calculus had been made. Theorems upon theorems had been discovered, leading to the solution of one practical problem upon another. But still, the logical foundation of calculus was pitifully deficient. The anology is a barely-equipped commando team going deep into enemy territories, taking one stronghold after another, while its supplies and back-up are nowhere in sight. Most people find this phenomenon shocking as they believe that mathematics is a science in which logic and rigour reign supreme. The 18th century English mathematician de Morgan said: "The driving force of mathematical discoveries is not deduction but imagination." It was by the sheer courage of 18th and 19th century mathematicians that calculus became a profound and versatile body of knowledge. While there had been minor mistakes, it verged upon a miracle that, lacking a solid foundation, there had not been any major breakdown in more than two hundred years and that it had not gone astray. In addition to the extraordinary intelligence and foresight of the master mathematicians, this was due to the fact that, at the time, mathematics developed hand in hand with physics and astronomy. Mathematical discoveries were "empirically verified" by scientific experiments, and at the same time, problems originating in science prompted the development of mathematical theories. This partnership with science provided the strong dose of confidence that mathematics needed.

In 1734, the English Bishop Berkeley wrote *The Analyst* in which he attacked calculus for its lack of logical foundation. This tract was subtitled "Or a Discourse Addressed to an Infidel Mathematician [presumably Halley]. Wherein It Is Examined Whether the Object, Principles, and Inferences of the Modern Analysis are More Distinctly Conceived, or More Evidently Deduced, than Religious Mysteries and Points of Faith." From this, Bishop Berkeley's motivation for attacking calculus was clear. However, from the mathematical point of view, he did raise numerous issues which had to be addressed. As more and more such questions badly needed answers, mathematicians in the mid-19th century realized the imperiousness of building a rigorous foundation for calculus. Thus the need to give calculus a new lease on life gave birth to a vast subject in mathematics – mathematical analysis. M²