Mathematics 106, Winter 2011

Instructor: V. S. Varadarajan
Office hours: MW 10:00-11:30, F 8:00–9:00
Time: MWF 9:00-9:50 MS5200
Discussion: T 9:00-9:50 MS 5127; Th 2:00-2:50 PAB 247
Text book: *The Historical Development of the Calculus*
Author: C. H. Edwards, Jr.

Syllabus: Parts of the twelve chapters of the book. Additional material may be taught as time permits.

Mid term (Closed book): Friday, February 4, 9:00–9:50

Final Exam (Closed book): Tuesday, March 15, 3:00–6:00

Chapter 1 (omit section on Eudoxus p. 12–p. 15)
Chapter 2 (pages 29–62)
Chapter 4 (pages 98–121)
Chapter 5 (pages 122–141)
Chapter 7 (pages 166–188)
Chapter 8 (pages 189–230)
Chapter 9 (pages 231–267)
Chapter 10 (pages 268–281 and 287–292)
Additional material (if time permits)

Homework assignments: All homework assignments for any week will be given by the Friday of that week and will be due on the Friday of the next week. The assignments should be given to the TA directly and recovered from that TA directly. *Late homework will not be graded.*

Office hours: You should try to make use of the office hours to clarify things discussed in the class. *Do not come to the office hours after missing the class, hoping to find out what was done.* The same applies to the office hours of the TA's.

Grades: The grades will be based on a three hour final examination (50%), midterm (25%) and homework (25 %). *There will be one midterm.* If you miss the final examination your score will be 0 for that part. No rescheduling of the finals or the midterm is possible. If you miss the midterm the final will count for 75%.

Virtual office hours, chatlines etc: I will not monitor these. Communications with me can be made personally during the class, office hours, or by email. *Do not use the language you use in texting to your friends. Be professional.*

Discussions: The discussion sections are a vital part of the course. Make an effort to attend them.

Classroom behavior: Use of cell phones, beepers, or pagers during class hours or discussion hours is forbidden. Such activities are a violation of UCLA student Conduct Code.

About the course

It is an almost universal feeling among students, especially among undergraduates, that mathematics is a cold and forbidding subject to learn. One reason for this is that the subject is invariably presented as a succession of theorems, lemmas, and (their) proofs, generally devoid of any motivation. In actual fact, this is not the way mathematics has developed. Ideas were tried, and rejected if they did not lead anywhere. There were false starts and fall backs. By trying to follow the historical development of the subject it is often possible to get a better understanding of the nature, meaning, and the necessity, of the concepts and their inter-relations. From this point of view, one of the best ways to learn a mathematical subject is to study it not as a logical fully formed structure but as a historically evolving organism.

The Calculus is one of the great discoveries in the history of human thought. Its origins go back to ancient times, but the decisive developments came in the sixteenth and seventeenth centuries, with the ideas of Descartes, Fermat, Newton, Leibniz, and ultimately, Euler. Their discoveries were given a firm foundation by Abel, Cauchy, Weierstrass, and their successors. However recent discoveries suggest that mathematicians in the state of Kerala in India had already discovered many parts of calculus in the fourteenth century.

In this course an attempt will be made to make this historical journey with you. That you have a background in calculus is a help, not a hindrance. It will allow you to reflect on how the mathematicians of the past arrived at their discoveries. One way to understand and appreciate their discoveries is to study them with the modern tools you have at your disposal. This will be our *modus operandi*.

There is a course given by Professor David Mumford at the UC Berkeley mathematics department which has some common themes with my course. For information go to

http://www.dam.brown.edu/people/mumford/Math191/PreliminarySyllabus.pdf

Timetable

The following timetable clearly shows the remarkable activity in the sixteenth and seventeenth centuries in mathematics, especially calculus. Each of these contributed in an important way, but Archimedes, Descartes, Fermat, Bernoulli (Johann), Newton, Leibniz, and Euler are the shining figures. After Euler the list includes some of the greatest figures who developed the theory of integration into one of the most far-reaching mathematical tools of today.

Pythagoras(c. 569 B. B.–c. 475 B. C.)
Euclid (c. 325 B.C.–c. 265 B. C.)
Archimedes (287 B.C.–212 B.C.)
Aryabhata (476–550)
Madhavan (c. 1350–1420)
Jyeshtadevan (c. 1500–1610)
Cardano (1501-1576)
Napier (1550–1617)
Kepler (1571-1630)
Descartes (1596–1650)
Cavalieri (1598–1647)
Fermat (1601-1665)
Wallis (1616–1703)
Mercator (1620–1687)
Pascal (1623–1662)
Gregory (1638–1675)
Newton (1643–1727)
Leibniz (1646–1716)
Bernoulli (Jacob) (1654–1705)
Bernoulli (Johann) (1667–1748)
Bernoulli (Daniel) (1700–1782)
Euler (1707–1783)
Lambert (1728–1777)
Lagrange (1736–1813)
Cauchy (1789 –1857)
Riemann (1826-1866)
Dedekind (1831–1916)
Cantor (1845–1918)
Lebesgue (1875–1941)
Wiener (1894–1964)
Siegel (1896–1981)
Robinson (1918–1974)
Feynman (1918–1988)
Langlands (1937–)

Week 1 : Ancient mathematics

This chapter is a summary of what mathematicians were trying to do in ancient times. The Babylonians, Egyptians, the Greeks, and the Indians are the main players in this *prehistory*. The development was due partly to the needs of commerce, and partly out of intellectual curiosity. Areas and volumes were computed for familiar figures. Because the answers had to be given numerically, considerable thought was devoted to number systems and ways of representing them. The Greeks discovered *irrational numbers* by exhibiting pairs of lengths which are not integral multiples of some unit, such as the side and hypotenuse of a right-angled triangle of equal sides ($\sqrt{2} : 1$). This is the same as saying that $\sqrt{2}$ is *not a rational number.* The Indians developed methods for solving equations where the solutions were required to be integers.

The Babylonians (c. 2000 B.C.–1600 B.C.) were quite advanced. They discovered the formula for finding the root of a quadratic equation which is still taught today. They used the *sexagesimal scale* which is the scale of 60 instead of the scale of 10 we use today. Not all contexts use the familiar decimal scale. For example, computer language is based on the *binary* scale. The nautical scale (degrees, minutes, seconds) is sexagesimal.

The babylonians found an approximation to $\sqrt{2}$ in the sexagesimal scale:

$$\sqrt{2} \approx 1; 24; 51; 10 = 1 + \frac{24}{60} + \frac{51}{60^2} + \frac{10}{60^3} \approx 1.414213.$$

The babylonian discoveries were preserved in *cuneiform tablets* of which the most famous is the Plimpton 322, which gives *Pythagorian triplets*, namely triples (a, b, c) where a, b, c are positive integers with

$$c^2 = a^2 + b^2, \quad a, b, c \text{ having no common factors.}$$

It is thus clear that centuries before Pythagoras, people were familiar with the theorem associated with his name, although it is almost certain that there was no *proof* by the standards of Euclid's geometry.

Here are some entries from the Plimpton 322 tablet:

1; 59	2; 49	119	169
56:07	1:20:25	3367	4825
1:16:41	1:50:49	4601	6649
3:31:49	5:09:01	12709	18541

The entries in the first 2 columns are in sexagesimal format, which are then converted to the decimal format in the last 2 columns. The entries in the first column represent the base b while the entries of the second column give the diagonal d. Hence $d^2 - b^2 = \ell^2$ where ℓ is the length. If the computations are correct this should be a positive integer, so that (b, l, d) is a triplet.

How did the babylonians discover such huge triplets? Note that if $b^2 + \ell^2 = d^2$ with b, d having no common factor, then one of b, d is even and the other is odd. Mathematicians coming centuries after the Babylonians discovered that all such triplets can be obtained by giving positive integer values to s, t where

$$b = 2st, \quad \ell = s^2 - t^2, \qquad d = s^2 + t^2 \quad (s > t).$$

It is believed that the babylonians were somehow aware of some such formula. For details see the great book of O. Neugebauer, *The Exact Sciences in Antiquity*, Dover, 1969. See also V. S. Varadarajan, *Algebra in Ancient and Modern Times*, AMS, 1998.

The Mayans used the *vigesimal scale* or the scale with base 20. They also discovered zero and place value. I do not know if the Mayans knew of Pythagorian triplets.

The idea of solving equations where the solutions are required to be integers goes back to the Greek mathematician *Diophantus* (c. 250). Such equations are called *diophantine*. The Indian mathematician *Brahmagupta* (598–665) began a deep study of the equation

$$X^2 - NY^2 = \pm 1$$

where $N > 1$ is a given integer and X, Y are required to be integers. By the twelfth century the Indians had completely solved all such equations.

Areas and volumes of cones and cylinders over plane domains. The ancients used intuitive arguments to show that the volume of a cylinder on a base of area A is Ah where h is the height of the cylinder. For a cone over the base of area A with height h the volume is $(1/3)Ah$.

Imagining a cylinder as being made up of very thin disks on the base area it is easy to see how the ancients might have arrived at the formula base area $\times$ height for the cylinder. The cone volume computation is more subtle and remarkable, because of the factor of $1/3$ in front. See the explanation on page 9 of the book for a triangular base. For a general base use triangular dissection if the base is polygonal, and approximation for general areas. Exercise 11 on page 9 (homework) suggests another way of deriving this. See also pages 19–21 of the book for another proof

The areas of simple figures bounded by straight lines can be computed by dissection into triangles and using the formula for the area of a triangle. For more complicated areas like the circle numerical measurements are worked out by dividing into small squares, rectangles, or triangles; this will not be exact but one can get to any degree of precision by taking the divisions to have arbitrarily small sides. This is the *method of exhaustion*. It was originally developed by Archimedes, and we shall discus it more carefully when we study it in the next chapter. In the questions for the ambitious student I have included an argument that proves that this method of computing areas and volumes (where cubes replace squares) is applicable when the area or volume has a *smooth boundary*.

Let me illustrate this method for the disk. It is applicable to balls as well and the treatments are similar and so I confine myself to disks. Let A be the unit disk $\{x^2 + y^2 < 1\}$ with center at the origin. We fix $\varepsilon = 2^{-k} > 0$ and look at the grid in the plane of points $z_{mn} = (m\varepsilon, n\varepsilon)$ where m, n are integers. To each such point we associate the square S_{mn} with the southwest vertex at z_{mn}. Let A^- be the area obtained by forming the union of all S_{mn} such that $S_{mn} \subset A$, and A^+ the area obtained by including also the S_{mn} such that S_{mn} meets the boundary of A. Then clearly

$$A^- \subset A \subset A^+.$$

The method of exhaustion is the statement that

$$\lim_{\varepsilon \to 0} \text{area}(A^-) = \lim_{\varepsilon \to 0} \text{area}(A^+).$$

If this is proved, the common value can be taken as the definition of the are of A. Note that $\varepsilon \to 0$ is the same as saying that $k \to \infty$, i.e., we use finer and finer resolutions. As k increases, it is obvious that A^- increases while A^+ decreases, so that $\lim_{\varepsilon\to 0} \text{area}(A^\pm)$ exist. So all we have to do is to prove that

$$\lim_{\varepsilon\to 0} \text{area}(A^+ \setminus A^-) = 0.$$

Now the boundary is given by the equation $x^2 + y^2 = 1$. We can write $y = f(x)$ with $f(x) = \sqrt{1-x^2}$ or $x = g(y)$ with $g(y) = \sqrt{1-y^2}$. Using either representation we can split the boundary into 4 arcs on each of which either $f'(x)$ or $g'(y)$ is bounded by a constant A. So it is enough to treat the case where the boundary is $y = f(x)$ with $|f'(x)| \le A$. If S_{mn} meets the graph of y, then for some (x, y) on the graph, we must have $|x - m\varepsilon| \le \varepsilon, |f(x) - n\varepsilon| \le \varepsilon$. By the mean value theorem $|f(x) - f(m\varepsilon)| \le A|x - m\varepsilon| \le A\varepsilon$ and so $|n\varepsilon - f(m\varepsilon)| \le (A+1)\varepsilon$. Hence for a given m there are at most $2(A+1)$ values of n. As $|m\varepsilon| \le 1$ we have $|m| \le \varepsilon^{-1}$ and so the number of (m, n) is at most $2(A+1)\varepsilon^{-1}$. So the contribution to $\text{area}(A^+ \setminus A^-)$ from this part of the boundary is at most $2(A+1)\varepsilon$, so that

$$\text{area}(A^+ \setminus A^-) \le 8(A+1)\varepsilon.$$

By dissection and exhaustion, it is clear from the above discussion that if A is an area bounded by smooth curves, then

$$\text{area } (tA) = t^2 \text{ area } (A).$$

By proceeding similarly it follows that for a volume V bounded by smooth surfaces,

$$\text{volume } (tA) = t^3 \text{ volume } (A).$$

Hence we deduce that the area (volume) of a disk (ball) of radius R is

$$c_2 R^2 \quad (c_3 R^3)$$

where c_2 (c_3) is the area (volume) of the unit disk (unit ball). We can determine these by integral calculus (see problem 8 of homework).

It was not until the early days of the twentieth century that the French mathematician *Henri Lebesgue* developed a rigorous theory of areas and volumes in all dimensions. The great modern mathematician *Alexander Grothendieck*, growing up in dire poverty and utter isolation, apparently discovered the entire theory of Lebesgue while he was a high school student. He is the creator of the modern approach to algebraic geometry; his great discoveries were in the 1950's. The discovery that the area of a circle of radius r is πr^2 is a triumph of ancient mathematics, especially *Archimedes*, 2π being defined as the circumference of a circle of radius 1.

Areas are certainly important in real estate transactions which certainly go back to ancient times. Volumes were probably important in measuring out given quantities of medicines in units of various forms-cones, cylinders, etc. For instance, the structure of a nozzle for eye drops is a little intricate and so computation of its volume is tricky.

Commensurability. Irrational numbers. The geometrical constructions of the Greeks led to the discovery that lengths are not always commensurable, i.e., they cannot be integral multiples of the same unit. Thus if we take a right-angled triangle of unit sides, its diagonal is $\sqrt{2}$. There is a famous classical argument that $\sqrt{2}$ is not rational. The Greeks certainly knew that many irrationals arose in natural geometric constructions. Euclid's book X discusses numbers like $\sqrt{a}, \sqrt{\sqrt{a} \pm \sqrt{b}}$, etc. These problems arose because the technology of numbers was lagging behind that of geometry.

Eventually people realized, especially Hilbert in the twentieth century, that number systems and geometry are one and the same. The person who is mainly responsible for a precise construction of the real number system is *Richard Dedekind* (1831–1916).

Numbers: Since in any measurement we have to use a fixed unit (inch, cm, mm, angstrom, etc) it follows that lengths will have to be approximated. So approximations of quantities from geometry in terms of decimal fractions or other formats, became a hot issue. The approximation *par excellence* is for π. Archimedes himself gave the rule (even today not easy to prove)

$$3\frac{10}{71} \ < \pi \ < \frac{22}{7}.$$

A number is called *algebraic* if it satisfies a polynomial equation with integer coefficients. For instance $\sqrt{2}$ is algebraic since it is a root of $X^2-2=0$. Any number obtained using the arithmetical operations on algebraic numbers is algebraic. If a number is not algebraic it is called *transcendental.* There are only countably many algebraic numbers and so, as the real numbers form an *uncountable set*, as was first proved by Georg Cantor (1845–1918), most numbers are transcendental. *Carl Ludwig Siegel* (1896–1981) made deep studies on transcendental numbers. More about this later.

It was eventually left to modern mathematicians of the twentieth century to try to understand the nature of all numbers obtained by solving equations with integer coefficients. This quest is still unfinished and taxes the entire structure of mathematics. The mathematician mainly responsible for this huge project is *Robert Langlands* (1937–). See his lectures on the web site for UCLA Distibguished lecture Series

http://www.math.ucla.edu/dls/2003/langlands042203.pdf

One particular class of numbers that arise from euclidean geometry is the set of those that can be obtained by *euclidean constructions* involving the straightedge (unmarked ruler) and compass. Suppose we call them *euclidean numbers*. Gauss electrified everyone by showing that the side of a regular polygon of 17 sides is a euclidean number. Roughly speaking, if a number is given by a formula involving only square roots and the operations of arithmetic, then it is euclidean. For a regular 17-gon the formula amounts to computing

$$2\cos\frac{2\pi}{17} = \alpha$$

which was shown by Gauss to be

$$\begin{aligned}&\frac{1}{8}\left\{-1+\sqrt{17}+\sqrt{34-2\sqrt{17}}\right\}+\\&\frac{1}{8}\sqrt{68+12\sqrt{17}-16\sqrt{34-2\sqrt{17}}-2(1-\sqrt{17})\left(\sqrt{34-2\sqrt{17}}\right)}.\end{aligned}$$

The side s_{17} of the regular 17-gon is then

$$s_{17} = \sqrt{2-\alpha}.$$

A regular n-gon can be constructed by ruler and compass if and only if n is of the form $2^k p_1 p_2 \dots p_s$ where the p_j are *Fermat primes*, i.e., primes of the form $p = 2^{2^h}+1$. For $h = 0,1,2,3,4$ this gives $p = 3,5,17,257,65537$ which are all primes. No other Fermat prime is known.

The set $\mathbf{E}$ of all euclidean numbers is an infinite field extension of the rational number field $\mathbf{Q}$. All quadratic equations with coefficients in $\mathbf{E}$ can be solved in $\mathbf{E}$. The field $\mathbf{E}$ is deeply related to euclidean geometry (see the remarks of Hermann Weyl on p. 273 of his book *Classical groups.*)

Homework # 1: Due January 14, 2011

1. If $b^2 + \ell^2 = d^2$ where b, ℓ, d are positive integers without a common factor, prove that one of ℓ, b is even and the other is odd.

2. Convert the entries below into the decimal format and find ℓ:

b	d
38;11	59; 01
45	1;15
27; 59	46; 49

3. Convert the numbers above to the Mayan scale of base 20.

4. Exercise 3 (a) (b) on p. 2.

5. Exercise 11 on p. 9.

6. Exercise 12 on p. 11.

7. Give an argument to show that $\sqrt[3]{2}$ is not rational.

8. By using integral calculus show that

$$\text{area of the unit circle} = \pi, \quad \text{area of the unit ball} = \frac{4}{3}\pi.$$

9. Use trigonometry to derive the formula

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta.$$

Hence show that $y = 2\cos 20^\circ$ satisfies

$$y^3 - 3y - 1 = 0.$$

This is the basis of the proof that there is no straight-edge and compass method for trisection of the angle 60°.

10. We consider the equation

$$a^2 - 2b^2 = 1$$

where a, b are positive integer unknowns. Prove that if $(a_1, b_1), (a_2, b_2)$ are two solutions, so is (a, b) where

$$a = a_1 a_2 + 2b_1 b_2, \quad b = a_1 b_2 + a_2 b_1$$

These formulae are due to Brahmagupta. Hence, starting from the solution $(3, 2)$ calculate the first 4 of an infinite sequence of solutions.

For the ambitious student

1. Try to define formally a euclidean number, namely, one that can be obtained by taking repeated square roots and using arithmetical operations. Use this definition to show that $\cos 20°$ is not a euclidean number. (*Hint*: If x is a euclidean number, it belongs to a field extension F of the rational number field $\mathbf{Q}$ of degree 2^k. If y lies inside F, y must generate a field whose degree over $\mathbf{Q}$ must also be a power of 2. But y generates a field of degree 3 over $\mathbf{Q}$!!)

2. Prove the representation of pythagorean triplets given in the notes above.

3. Prove that $\sqrt{2} + \sqrt{3}$ is algebraic by constructing an equation with integer coefficients satisfied by it.

4. Extend the argument in the discussion above to the area of a region bounded by a collection of smooth curves.

Week 2 : Archimedes (287 B. C.–212 B. C)

The circle and the parabola

The circle. Archimedes is universally regarded as the greatest figure of ancient science. Many of his manuscripts have been lost because of their antiquity and also because of the destruction foolishly carried out by the armies of warring factions. However what is left is enough to stamp him as a genius, an innovator, and a man equally skilled in mathematics for its own sake as well as its applications, including armed warfare.

It is believed that a large part of his work has been lost. The works that are available include

1. Measurement of a Circle
2. Quadrature of the Parabola
3. On the Sphere and the Cylinder
4. On Spirals
5. On Conoids and Spheroids
6. The Method

The work of Archimedes in pure mathematics consisted in developing techniques for computing areas and volumes of geometrical figures. I shall now discuss these ideas. For any curve C denote its length by $\ell(C)$; for any domain D denote its area by $a(D)$; and for any region R in three dimensional space let $v(R)$ be its volume. It is implicitly assumed that these are well defined for all the curves, domains, and regions that one comes across. For curves length is understood as follows. Let P be a polygonal line with vertices on the curve. Then $\ell(P)$ is well defined. The number $\ell(C)$ is defined as the limit of the numbers $\ell(P)$ as the polygon approaches the curve in the limit. If the curve is *closed* and *convex* like the circle, or the ellipse, and there is no difficulty in accepting this idea of length. We call the polygons P *inscribed*. One can also take polygons whose sides touch the curve, called *circumscribed* polygons. Already the duality in geometry between points and lines is exhibited in these ideas. The length $\ell(C)$ is also understood as the limit of $\ell(P)$ for circumscribed polygons P which approach the curve in the limit.

In the nineteenth century these notions were subjected to critical analysis. Mathematicians like *Guiseppe Peano* (1858–1932) realized that not all curves have length in the above sense, as the above-mentioned limits *may not exist.* Indeed, Peano constructed a curve which filled the entire unit square!! Such curves are called *space filling* or *Peano curves.* Curves for which the definition of length given above works are called *rectifiable*, and the process of determination of the lengths of rectifiable curves is called *rectification.* The Peano curves are of course not retifiable. Once people understood calculus, they were able to show that if a curve C is *parametrized* by a map

$$t \longmapsto p(t) = (x(t), y(t)) \qquad (a \le t \le b),$$

then the curve is rectifiable as soon as the functions $x(t), y(t)$ are *differentiable with continuous derivatives*; in this case the length is computed by the remarkable formula

$$\ell(C) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\ dt$$

where $f'(t) = df/dt$ for any function f. The intuitive basis for this formula is that the total length is understood as the sum of the infinitesimal lengths, namely,

$$\sum\{\text{distance between the points } p(t+\Delta t) \text{ and } p(t)\}$$

which is

$$\sum ||p(t+\Delta t) - p(t)||$$

and

$$||p(t+\Delta t) - p(t)|| \approx \sqrt{x'(t)^2 + y'(t)^2}\ \Delta t.$$

For surface areas there is an analogous method. The surface is supposed to be given by an equation

$$z = f(x,y) \quad ((x,y) \in D)$$

with a function f which is continuously differentiable, and the surface area is

$$\int_D \sqrt{1 + ||\nabla f||^2}\ dxdy$$

where ∇f is the gradient of f so that

$$||\nabla f||^2 = (\partial f/\partial x)^2 + (\partial f/\partial y)^2.$$

Notice that these definition could not be made unless the curve or surface is given in terms of *differentiable* functions. Therefore non-rectifiable curves must be described by functions that are *continuous but not differentiable*. Such functions were discovered only in the nineteenth century.

I give these modern definitions, which were first treated decisively by Newton and Leibniz, to underscore the gigantic nature of Archimedes's achievement, almost two millennia before the discovery of the calculus. In this sense it is natural to regard him as the forerunner of people like Newton, Leibniz, and Euler, and finally Riemann and Lebesgue.

Definition and calculation of π. For any polygon P, any polygonal domain D, and region R, it is clear that

$$\ell(tP) = t\ell(P), \qquad a(tD) = t^2 a(D) \qquad v(tR) = t^3 v(R) \qquad (t > 0)$$

which shows how length, area, and volume behave under *dilations*. Let C_t, D_t, B_t denote the circle, disk, and ball of radius t. Then writing

$$c_1 = \ell(C_1) \qquad c_2 = a(D_1) \qquad c_3 = v(B_1)$$

we see that

$$\ell(C_t) = c_1 t^2 \qquad a(D_t) = c_2 t^2 \qquad v(B_t) = c_3 t^3.$$

Archimedes's stunning achievement was to show first of all that if we define π by

$$\pi = \frac{1}{2} c_1$$

then

$$c_1 = 2\pi \qquad c_2 = \pi \qquad c_3 = \frac{4}{3}\pi.$$

Moreover if S_t is the spherical surface of radius t, then

$$a(S_t) = 4\pi t^2$$

where we understand by $a(S)$ the surface are of any surface S in three-space. *In other words, all constants of proportionality are determined in terms of a single one, namely π.* Thus we have the four formulae

$$\ell(C_t) = 2\pi t \qquad a(D_t) = \pi t^2 \qquad v(B_t) = \frac{4}{3}\pi t^3 \qquad a(S_t) = 4\pi t^2.$$

Since the number π was introduced indirectly as the ratio of the circumference of the unit circle to its diameter, it became essential to compute it numerically. Archimedes found the remarkable approximations

$$3\frac{10}{71} < \pi < 3\frac{1}{7} \qquad (1).$$

Ever since his work, the problem of finding as much of the decimal expansion of π as possible has attracted mathematicians. By using very refined mathematics people can now compute billions of digits in its decimal expansion, and have algorithms for computing the entire expansion.

Some history on the nature of π. The nature of π has intrigued mathematicians from the ancient era itself. The Indian mathematician and astronomer *Aryabhata* (476–550) discovered the approximation 3.1416 for π. The Indian (Kerala) mathematicians of the 14^{th} century, Jyeshtadevan and Nilakanthan already wondered if π is irrational. *Johann Heinrich Lambert* (1728–1777), Swiss mathematician, physicist, and astronomer, proved in 1761 that π is irrational. More generally he proved that if $x \neq 0$ is rational, then $\tan x$ is irrational. Thus, as $\tan \pi/4 = 1$, $\pi/4$, thus π itself, has to be irrational. To prove irrationality or transcendentality of a specific number is always difficult. *Ferdinand von Lindemann* (1852–1939) proved in 1882 that π is a transcendental numbe. The number e is also transcendental, as was proved in 1873 by *Charles Hermite* (1822-1901). *Joseph Liouville* (1809–1882) found a general method of constructing transcendental numbers. *Georg Cantor* (1845–1918), creator of modern set theory, proved that almost all numbers are transcendental. The irrationality of e and e^2 was proved by Euler. For an English translation of Euler's paper in Latin, see *An essay on continued fractions*, Myra F. Wyman and Bostwich F. Wyman, Mathematical Systems Theory, **18**(1985) 295–328.

The method of Archimedes for determining π consists in computing the lengths and areas of the inscribed and circumscribed *regular* polygons of the unit circle. Let $P_n^{\pm}$ be the circumscribed and inscribed regular polygons and let

$$C_n^{\pm} = \ell(P_n^{\pm}) \qquad A_n^{\pm} = a(P_n^{\pm}).$$

Let C be the length (=circumference) of the entire circle. Its existence as well as its value will follow from the discussion below.

Let θ_n be one half of the angle subtended at the origin of a side of the regular n-gone P_n^- inscribed in the unit circle. Then

$$C_n^- = 2n \sin\theta_n \qquad C_n^+ = 2n \tan\theta_n.$$

Observe now that

$$\sin\theta < \theta < \tan\theta$$

and so

$$C_n^- < C < C_n^+.$$

Now, as $n \to \infty$ we have

$$\theta_n \to 0$$

so that

$$\frac{C_n^+}{C_n^-} = \frac{1}{\cos\theta_n} \to 1.$$

On the other hand, as $\theta_n \to 0$, we have

$$\frac{\sin\theta_n}{\theta_n} \to 1$$

so that

$$C_n \to \gamma$$

where we denote by γ the total angle at 0. Hence

$$\gamma = 2\pi$$

and

$$C_n^- \to 2\pi, \qquad C_n^+ \to 2\pi.$$

For the areas we use the formula $(1/2)ab\sin\theta$ for the area of a triangle with sides a, b and the included angle θ. Hence

$$A_n^- = 2n(1/2)\sin 2\theta_n = (1/2)C_n^- \to \pi \qquad A_n^+ = 2n(1/2)\tan 2\theta_n = (1/2)C_n^+ \to \pi.$$

We have used in a non-trivial manner the theory of the trigonometric functions in this argument. An algebraic proof can be given as follows. Let us take a triangle with one vertex at the origin (center of the unit circle) and the other two on the circle. Let $2s$ be the length of the side opposite the origin. Let 2θ be the angle at the origin. Then the area Δ of the triangle is

$$\Delta = s\sqrt{1-s^2}.$$

This algebraic formula shows clearly how the side of the polygon determines its area. If s_n is the half-length of the side of C_n^- we get from this

$$\frac{A_n^-}{C_n^-} = \frac{1}{2}\sqrt{1-s_n^2} \to \frac{1}{2}$$

since $s_n \to 0$. This shows that $c_2 = \pi$.

For his numerical calculation Archimedes starts with a regular k-gon for a small value of k and repeatedly doubles it to go to regular n-gons where $n = k, 2k, 4k, 8k, 16k$ etc. To do this we need a recursive formula for relating the regular n-gon to the regular $2n$-gon. Since the angle is halved, we need formulae relating the trigonomeric functions of an angle α with those of the angle 2α. The formulae are:

$$\sin^2\alpha = \frac{\sin^2 2\alpha}{2+\sqrt{4-4\sin^2 2\alpha}}, \qquad \tan\alpha = \frac{\tan 2\alpha}{1+\sqrt{1+\tan^2 2\alpha}} \tag{2}$$

Let t_n denote the half length of a regular circumscribed n-gon and s_n the half length of a regular inscribed n-gon. Then

$$t_n = \tan\theta_n, \quad t_{2n} = \tan\frac{\theta_n}{2}, \quad s_n = \sin\theta_n, \quad s_{2n} = \sin\frac{\theta_n}{2}$$

so that

$$t_{2n} = \frac{t_n}{1+\sqrt{1+t_n^2}}, \qquad s_{2n}^2 = \frac{s_n^2}{2+\sqrt{4-4s_n^2}} \qquad (2a).$$

The formula (4) on page 34 differs from the above formula relating s_{2n} to s_n, but remember that the book defines s_n as the full side of the inscribed regular n-gon whereas for us it is the half side. The formula relating t_{2n} to t_n remains the same because the book defines t_n as half the side of the circumscribed regular n-gon. The basic inequality is

$$ns_n < \pi < nt_n.$$

Let us start with $n = 6$. Then $s_6 = 1/2, t_6 = 1/\sqrt{3}$ and so we have

$$3 < \pi < 2\sqrt{3} \Rightarrow 3 < \pi < 3.47$$

which is poor but serves as a starting point for computing with regular polygons of 12, 24, 48, 96 sides. For $n = 12$ we have

$$s_{12} = \frac{\sqrt{2-\sqrt{3}}}{2} > 0.2588, \qquad t_{12} = 2-\sqrt{3} < 0.2680$$

giving

$$3.1056 < \pi < 3.2160$$

which still does not fix the first decimal. This suggests that the convergence to π by the method of doubling the regular polygons is very slow. Starting with a regular decagon one can get by this method, going up to regular polygons of 640 sides the approximation

$$\pi \approx 3.1416$$

rounded off to 4 decimals. For getting (1) Archimedes used the 96-gon, and at each step, used precise rational approximations for the square roots. By going to 96 we get a good approximation to 2 decimals 3.14, also the approximation (1). However this verification of (1) is cheating; Archimedes used deep rational approximations to derive (1) using the 96-gon. It is easy to write a program for this and with current computers we can carry this method much farther and with geater ease.

We also used trigonometry to establish (2a). Actually the formulae (2a) can be obtained from geometrical considerations (see pages 33-34 of the book).

The circle was regarded as the perfect curve in ancient times. people made an attempt to describe every curve in terms of the circle. For example, the cycloid, which formed the basis of Ptolemic astronomy, was obtained as the locus of a point, fixed on a circle, as the circle rolls with uniform velocity on a line.

The parabola. The *parabola* is one of the conic sections, namely a curve obtained by a plane section of a right circular cone. It is not like the usual curves one encounters because it is not enclosed in a finite part of the plane, like the ellipse. It is obtained if the plane cutting the cone, is parallel to a generator of the cone (see fig. 5, p. 36). It is more like the hyperbola. If we take the ellipse

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1$$

and translate so that the point $(-a, 0)$ is moved to $(0, 0)$, we get

$$\frac{(X-a)^2}{a^2} + \frac{Y^2}{b^2} = 1$$

or

$$Y^2 - 2\frac{b^2}{a}X + \frac{b^2}{a^2}X^2 = 0$$

If we now let $a, b \to \infty$ but in such a way that $2b^2/a$ goes to a constant $k > 0$ we get, since $b^2/a^2 = (2b^2/a)(2/a) \to 0$, the curve

$$Y^2 = kX \qquad (k > 0)$$

which is the equation of a parabola. To get the parabola

$$Y^2 = kX \qquad (k < 0)$$

we do the same limiting process with the hyperbola

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1.$$

There are also other forms of equations that describe a parabola, which change its orientations.

The parabola is important in many ways. It has an *axis* of symmetry with the property that the curve is unchanged under reflection in the axis. For the curve

$$Y^2 = kX \qquad (k > 0)$$

the parabola lies entirely to the positive side of the Y-axis which is the axis of symmetry of the curve. It has a *focus*, a point on its axis, with the following property: any line parallel to the axis and hitting the curve, will get reflected into a line through the focus. This property is used heavily in applications, in the construction of parabolic mirrors or antennae. Archimedes is believed to have used the focussing property of parabolic mirrors to burn ships of invading armadas.

The trajectories of missiles are parabolic. To see this we need Newton's laws of motion. Suppose we fire a missile like a rocket, having a mass m, with an initial angular velocity (a, b). So the initial speed is $\sqrt{a^2 + b^2}$ in the direction (a, b). We assume $a > 0, b > 0$. Once fired the missile is not accelerated and falls under the earth's gravitation. If we work in the xz-plane the equation

$$\text{Force } = \text{mass } \times \text{acceleration}$$

leads to the equations

$$\frac{d^2x}{dt^2} = 0, \qquad \frac{d^2z}{dt^2} = -mg$$

where g is the constant acceleration due to gravity. The initial conditions are

$$x(0) = z(0) = 0, \qquad \frac{dx}{dt}(0) = a, \ \frac{dz}{dt}(0) = b.$$

The solution is

$$x(t) = at, \qquad z(t) = -\frac{1}{2}mgt^2 + bt \quad (t \geq 0).$$

Eliminating t we get

$$z = -Ax^2 + Bx \qquad \left(A = \frac{mg}{2a^2}, \quad B = \frac{b}{a}\right).$$

This is a parabola with a maximum at

$$\left(\frac{B}{2A}, \frac{B^2}{4A}\right).$$

If we introduce new coordinates

$$X = x - \frac{B}{2A}, \quad Y = z - \frac{B^2}{4A}$$

we get the equation

$$Y = -AX^2$$

which is a parabola going downwards. Changing the sign of Y and interchanging X and Y we get it in the earlier form

$$Y^2 = kX \qquad \left(k = \frac{1}{A}\right).$$

Quadrature of the parabola. Quadrature means computing the area. Since the parabola extends to infinity, one cannot speak of the area it encloses, and so the next best thing is to take a line meeting the parabola and try to compute the area of the finite part cut off by the line. This is what Archimedes did in a beautiful calculation. Let A, B be points on the parabola and S be the short part of the curve cut off by the segment AB (see figure 8 in p. 37 of the book). By the *vertex* of S we mean the point P on the parabola farthest from the line AB. It is also the point at which the tangent to the parabola is parallel to the line AB. Then Archimedes proved the magnificent result

$$a(S) = \frac{4}{3}a(\Delta APB) \tag{3.}$$

The proof of this formula is a remarkable illustration of the method of exhaustion developed by Archimedes. The line AB gives rise to lines AP, PB and two vertices P_1 above the segment PB and P_2 above AP. Archimedes proves that

$$a(\Delta AP_2P) + a(\Delta PP_1B) = \frac{1}{4}a(\Delta APB) \tag{4.}$$

We can then continue finding vertices above AP_2, P_2P, PP_1, P_1B, and so on, to get triangles that have areas always (1/4) of the previous set. Clearly this procedure fills the area S completely in the limit, or as we say, *exhausts the area* S. Thus

$$a(S) = a(\Delta APB)\left(1 + \frac{1}{4} + \frac{1}{4^2} + \dots\right) = \frac{4}{3}a(\Delta APB).$$

Obviously (4) is the key step. The proof depends on properties of the parabola which were familiar to Archimedes from earlier work of his predecessors like Apollonius (who wrote a famous treatise on conic sections). In the text book, (4) is proved by using the properties of the parabola listed without proofs on p. 37 therein. I shall indicate a proof of (4) *analytically*, leaving some steps for you to verify in the homework problems.

The first step is to show that the two definitions of the vertex are the same. Next one has to show that a line from the vertex P parallel to the axis meets AB at its mid point M. So the line from P_1 parallel to the axis meets PB at its mid point Y. Hence M_1, the point of MB where P_1Y meets MB, bisects MB. The next step is to prove

$$YM_1 = 2P_1Y \qquad (5).$$

Suppose that we have proved (5). Then

$$\begin{aligned} a(\Delta PP_1B) &= 2a(\Delta PP_1Y) \\ &= a(\Delta PM_1Y) \\ &= \frac{1}{2}a(PM_1B) \\ &= \frac{1}{4}a(\Delta PMB) \\ &= \frac{1}{8}a(\Delta APB). \end{aligned}$$

Similarly we have

$$a(\Delta AP_2P) = \frac{1}{8}a(\Delta APB).$$

Hence we get (4). In this derivation we have repeatedly used the principle that if two triangles have a common vertex and bases on a given line, the ratio of their areas is the same as the ration of their bases.

It remains to prove (5). This is done by computing the coordinates of P_1, Y, M_1 in terms of the coordinates of A and B.

Points of $Y^2 = X$ are precisely those of the form (t^2, t) for some real number t. Only the origin corresponds to $t = 0$; all other points have $t \neq 0$. Let $A = (a^2, a), B = (b^2, b), a \neq b$. The equation to AB is

$$(b^2 - a^2)(Y - a) = (b - a)(X - a^2) \quad \text{or } (b + a)(Y - a) - (X - a^2) = 0.$$

Now the square of the distance of a point (t^2, t) not on the line given by an equation $uX + vY + w = 0$, to that line, is

$$\frac{(ut^2 + vt + w)^2}{(u^2 + v^2)}$$

which is maximized if t satisfies

$$(ut^2 + vt + w)(2ut + v) = 0$$

or

$$2ut + v = 0$$

(since the first factor cannot be 0). For the line AB, for which $u = -1, v = a + b$, this means that $t = (a + b)/2$. Hence the coordinates of P are given by

$$P = (c^2, c) \quad c = \frac{a+b}{2} \qquad (6).$$

It is now easy to show that P is also the unique point on the parabola with the property that the tangent to the parabola at P is parallel to the line AB; I shall leave this as a homework problem.

The formula (6) is very powerful. It follows from (6) that the line from P parallel to the axis meets AB at the point

$$M = \left(\frac{a^2 + b^2}{2}, \frac{a+b}{2}\right) \qquad (7)$$

showing that M bisects AB. Since we know the coordinates of P we can apply (6) to get coordinates of P_1. Thus

$$P_1 = (c_1^2, c_1) \quad c_1 = \frac{c+b}{2} \qquad (8).$$

The point Y bisects PB and so its coordinates are

$$Y = \left(\frac{c^2 + b^2}{2}, \frac{c+b}{2}\right) \qquad (9).$$

Since M_1 bisects MB its coordinates can be found from those of M, B. From the knowledge of the coordinates of P_1, Y, M_1 we can verify (5).

Homework # 2: Due January 21, 2011

1. Prove that $\sin\theta < \theta < \tan\theta$ for $0 < \theta < \pi/2$ by integrating the inequality $\cos\theta < 1 < \sec^2\theta$.

2. Prove Heron's formula that the area of a triangle with sides a, b, c is

$$\sqrt{s(s-a)(s-b)(s-c)} \qquad s = \frac{a+b+c}{2} \text{ is the semiperimeter.}$$

3. Derive the formulae (2) on page 13 of these notes.

4. Prove that

$$\sin 18^\circ = \frac{\sqrt{5}-1}{4}.$$

5. Ex. 6 on p. 35.

6. Exercise 7 on p. 35.

The next set of problems aim to prove (5) of the text above.

7. Show that
$$Y^2 = 4kX \qquad (k > 0)$$
is the locus of a point which moves in such a way that its distance from $(k, 0)$ is equal to its distance from the line $X = -k$.

8. Deduce from (6) of these notes that P is the unique point on the parabola such that the tangent to the parabola at P is parallel to AB, and establish (7).

9. Prove that
$$M_1 = \left(\frac{a^2 + 3b^2}{4}, \frac{a + 3b}{4}\right)$$
and deduce from this the formula (5).

For the ambitious student

1. Compute directly, using coordinates, the areas of the domain S and the triangle APB to verify the theorem of Archimedes that
$$a(S) = \frac{4}{3}a(\Delta APB).$$

2. Find the focus of the parabola $Y^2 = kX(k > 0)$ and verify the focussing property.

3. There is a formula due to the Indian mathematician *Brahmagupta* (c. 597–680) that gives the area Q of a quadrilateral inscribed in a circle of sides a, b, c, d. It is
$$Q = \sqrt{(s-a)(s-b)(s-c)(s-d)} \qquad s = \frac{a+b+c+d}{2}.$$
Give a proof of this.

4. Show that e is irrational using the infinite series for e. (*Hint*: If $e = p/q$, with $p, q \geq 1$, show that $q!e$ is an integer and equals an integer θ where
$$\theta = \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \ldots$$
and
$$0 < \theta < \frac{1}{q+1} + \frac{1}{(q=1)^2} + \ldots = \frac{1}{q} < 1,$$
a contradiction.)

5. Weierstrass constructed the first examples of continuous functions which are nowhere differentiable. One such is
$$f(x) = \sum_{n=1}^{\infty} a^n \cos(b^n \pi x)$$
where $0 < a < 1$, $b > 0$ is a positive odd integer, and $ab > 1 + \frac{3\pi}{2}$. Prove this.

Week 3 : Archimedes (continued)

Volumes, surface areas, and the spiral

Surface areas and volumes of spheres. Archimedes conceived of the idea of defining the surface area of the sphere and other surfaces like the paraboloids and developed what may be regarded as an axiomatic method for *convex* surfaces. He obtained several remarkable formulae connecting surface areas with volumes, in generalization of what he did for lengths and areas in two dimensions. For example, he proved that if in three dimensional space S_t is the sphere of radius t and B_t is the ball of radius r, then writing $a(\cdot)$ for the surface area,

$$a(S_t) = 4\pi t^2, \qquad v(B_t) = \frac{4\pi}{3}\pi t^3.$$

He also showed that the surface area of a sphere is the same as the surface area of an enclosing cylinder (minus the top and bottom areas)

The formulae show that

$$v(B_t) = \frac{1}{3}a(S_t).$$

If this can be established, it is then sufficient to prove only one of the two previous ones. Archimedes's proof of this is very simple. He thinks of the ball as the union of pyramids with vertex at the center and base a small square whoese vertices are on the sphere, very much analogous to the approximation of the circle through regular inscribed polygons. The volume and surface area of the pyramids are related as above, which gives the above formula by summing and going to the limit. We have already discussed as a problem the computation of the volume by using integral calculus. We now discuss the surface area in the same fashion.

The sphere is the union of the upper and lower hemispheres. The upper hemisphere is given by the equation

$$z = +\sqrt{t^2 - x^2 - y^2}.$$

Hence

$$a(S_t) = 2\int\int_{x^2+y^2\le t^2} f\,dxdy$$

where

$$f(x,y) = \left(1 + (\partial z/\partial x)^2 + (\partial z/\partial y)^2\right)^{1/2}.$$

This can be done by using polar coordinates. See homework problem 1.

What is the heuristic derivation of the formula for the surface area of the graph of a function $z = g(x,y)$ defined over a domain D? Following Archimedes we erect rectangular prisms over small rectangles in the domain D. The surface area above a small square, approximated by the slanted square (like a roof) above the base rectangle, has the value

$$\left(1 + (\partial g/\partial x)^2 + (\partial g/\partial y)^2\right)^{1/2} \Delta x \Delta y$$

where Δx and Δy are the sides of the small rectangle centered at (x, y). In the limit the sum of these areas becomes the double integral. Clearly our modern understanding of surface areas depends on both differential and integral calculus in two variables.

The Archimedean spiral. The equation for the spiral in polar coordinates is

$$r = a\theta \qquad (a > 0 \text{ is a constant})$$

which arises as the locus of a point moving with uniform speed on a rod that is revolving with uniform angular velocity about one of its endpoints which is fixed, the moving point starting from that end point. It can be parametrized by

$$x = at\cos t, \quad y = at\sin t \quad (t \geq 0).$$

Other spirals are obtained by

$$x = f(t)\cos t, \quad y = f(t)\sin t \quad (t > 0)$$

for diverse choices of the function $f(t)$, leading to the equation

$$r = f(\theta).$$

A special case is the *logarithmic spiral*, or *spira mirabilis*

$$r = be^{\theta}.$$

This appears in nature as the spiral of the beautiful *nautilus shell.* As $\theta \to 0$ the logarithmic spiral approximates the archimedean spiral

$$r = b + b\theta.$$

Archimedes made a deep study of his spiral and discovered many beautiful properties of it. The first is the formula for the part of the spiral starting from the origin up to the first turn. This area is

$$A_1 = \frac{1}{2}\int_0^{2\pi} r^2 d\theta = \frac{\pi}{3}(2\pi a)^2 \qquad (1).$$

Archimedes did this of course by his approximation procedures, during which he derived formulae for the sum of squares of numbers in an arithmetic progression. The above formula can be generalized to give the area between any two radial cutouts of the spiral (see (36) on page 59 of the book and the homework problem 3). The area A_n bounded by the n^{th} turn $(2\pi(n-1) \leq \theta \leq 2\pi n)$ and the portion of the polar axis joining its end points, and the area $R_n = A_n - A_{n-1}$ of the ring between the $(n-1)^{\text{th}}$ and n^{th} turns between successive turns of the spiral are given by

$$A_n = \frac{4\pi}{3}(\pi a)^2(3n^2 - 3n + 1), \qquad R_n = 6(n-1)A_1 \qquad (2).$$

Other curves and surfaces. Archimedes also evaluated the area of the ellipse and the volumes and surface areas of paraboloids and ellipsoids of revolution.

Concluding remarks on Archimedes and ancient mathematics. One can say with great conviction that Archimeds took the first steps towards a theory of integration. The derivations of his formulae by modern integral calculus highlight the fact that his methods of approximation and geometric intuition form the foundations of the modern theory of integration. It was not until Riemann's work that a formal theory of integration became available, via the theory of Riemann sums and the view of the integral as the limit of Riemann sums.

One feature of his work that we have not touched is the so-called *archimedean axiom* for the real numbers. This states that if $a > 0, b > 0$ are two numbers, then there is a positive integer n such that $na > b$. This means that you can always measure a length b, no matter how large it is, by using the length a as a unit. In modern physics however, there are *absolute limits to measurement*. In fact, inside regions of linear dimension smaller than 10^{-33} cm, no measurements are possible, let alone comparisons of length. So in such regions (called *Planckian*), the geometry of space-time may not be archimedean. One has to use number systems other than real numbers to construct mathematical models of such space-times. The p-adic numbers provide such number systems!! They play a vital role in modern number theory. Non-archimedean number systems have also come to the fore in *non-standard analysis.*

Homework # 3 : Due January 28, 2011

1. Show that for

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

we have

$$f(x, y) = \left(1 + (\partial z/\partial x)^2 + (\partial z/\partial y)^2\right)^{1/2} = \frac{t}{z}.$$

Hence show that

$$\int\int_{x^2+y^2 \le t^2} dxdy = 4\pi t$$

using polar coordinates.

2. Ex. 17, p. 57

3. Prove (36), p. 59, by using the formula for area in polar coordinates.

4. Ex. 18, p. 62.

5. Write the equation of the archimedean spiral in Cartesian coordinates.

6. Find the area of the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

by dilating the x-coordinate axis.

For the ambitious student

1. Ex. 16, p. 56.

2. How do the p-adic numbers violate the archimedean axiom?

Week 4 : Precalculus

Beginnings of integral calculus. Greek mathematics declined and disappeared while Europe entered the dark ages. During this time it was kept alive in India and the Arab world, and the discoveries made there eventually made their way to Europe via the so-called *silk route*. It was not until the sixteenth century that mathematics started flourishing again in Europe.

Cavalieri. *Bonaventura Cavalieri* (1598–1647) was an influential figure who developed infinitesimal techniques for computing areas and volumes. He wrote two books outlining his ideas, *Geometria indivisibilibus* and *Exercitationes geometricae sex*. One of his main results is that if two solids have the property that the areas of their horizontal cross sections are in the same constant ratio, their volumes are also in the same ratio. He also developed ingenious methods to calculate the volume of a single solid but they were not convincing.

Wallis. In the period before Newton and Leibniz, people like *John Wallis* (1616-1703) started the beginnings of *integral calculus*, by trying to develop general methods of computing the area under various curves, in the tradition of Archimedes. Typically they started with a function $y = f(x)$, say for x between 0 and 1, and tried to compute the area between $x = 0$ and $x = 1$ as a limit

$$\lim_{n\to\infty} \frac{f(1/n) + f(2/n) + \ldots + f(1)}{n}$$

In modern notation this would mean that

$$\int_0^1 f(x)dx = \lim_{n\to\infty} \frac{f(1/n) + f(2/n) + \ldots + f(1)}{n}.$$

This would not become clear until Riemann who invented the Riemann sums and showed that the integral is the limit of Riemann sums. Wallis himself tried to compute this limit when

$$f(x) = x^{p/q}$$

and attempted to show that

$$\lim_{n\to\infty} \frac{1^{p/q} + 2^{p/q} + \ldots + n^{p/q}}{n^{1+p/q}} = \frac{q}{p+q}.$$

He was not successful but such work slowly paved the way to an understanding of the integral as a limit of sums, not just for a few isolated functions, but for a very large class of functions.

Fermat. In the town hall of Toulouse, France, there is a bust of *Pierre de Fermat*(1601-1665), native son, with the brief legend: *Discoverer of Calculus*. This is pardonable pride but it also has a substantial element of truth. It is not correct to expect that such a great discovery as the Calculus came out of one man's imagination, mature and fully grown. Many people (before Newton and Leibniz) made discoveries that slowly developed and suggested that a new idea will be decisive. Newton and Leibniz were fortunate enough to come upon such a new idea.

Fermat made contributions to many parts of mathematics and science. He was the first number theorist and had the great insight to conjecture what is now called *Fermat's last theorem* which was only proved a few years ago, by Andrew Wiles and Richard Taylor, using almost all resources of contemporary mathematics. Fermat discovered the principle that *light travels along paths for which the travel time is least*, from which almost all results of optics follow; he himself derived *Snell's law of refraction* by this argument. The method of calculating the path with least time duration is an example of the *calculus of variations*, first developed by Euler and then by Lagrange, almost a century later, and is absolutely fundamental in all of physics. He developed methods of calculating the tangents to various curves given by diverse equations, as well as their areas. With *Blaise Pascal* (1623–1662) Fermat created a new discipline, *The theory of probability*, although *Gerolamo Cardano* (1501-1576) had anticipated many aspects of it in an earlier work.

Fermat was able to complete the calculation of

$$\int_0^a x^{p/q}dx = a^{p+q/q}\frac{q}{p+q}.$$

His method was to chose a subdivision of the interval $(0, a)$ not by equally spaced points but by points in a *geometric progression*, namely $ar^n(n = 0, 1, 2,)$. Notice that this is an *infinite* sequence of points. The approximation of the area by rectangles erected at these division points gives (see page 116 of the book)

$$A(r) = \sum_{n=0}^{\infty}(ar^n)^{p/q}(ar^n - ar^{n+1}) = a^{p+q/q}\frac{1 + t + t^2 + \ldots + t^{q-1}}{1 + t + t^2 + \ldots + t^{p+q-1}} \qquad (t = r^{1/q})$$

and

$$\int_0^a x^{p/q}dx = \lim_{t\to 1} A(r) = a^{\frac{p+q}{q}}\frac{q}{p+q} = \frac{a^{\frac{p}{q}+1}}{\frac{p}{q}+1}.$$

For us this is a special case of the general result

$$\int_0^\alpha x^\alpha dx = \frac{a^{\alpha+1}}{\alpha+1} \quad (\alpha \text{ any real number } > -1).$$

This can be obtained by a limit argument, taking rational approximations to α. Fermat restricted himself to rational exponents p/q because no one knew how to define powers like $x^{\sqrt{2}}$. It was only after Dedekind estblished the foundations of real analysis that such irrational powers were brought into the fold of analysis. Finally, the condition $\alpha > -1$ is needed for convergence; for $\alpha = -1$ the integral diverges as $\log x$ becomes infinite as $x \to 0+$.

Descartes. *Rene Descartes*(1596–1650) is universally regarded as one of the greatest philosophers of western thought. He was also a remarkable mathematician, the discoverer of analytic geometry, and the developer of infinitesimal methods of calculation of tangents to curves given by algebraic equations. He was therefore a true forerunner of Newton and Leibniz. He is the author of one of the most famous remarks in philosophy : *cogito ergo sum* meaning *I think, therefore I am.*

His greatest mathematical discovery is that *coordinates* can be assigned to points on the plane in such a way that a straight line is given by a *linear* equation, a conic section is given by a *quadratic* equation, and so on. This allowed the powerful methods of algebra and analysis to bear upon geometrical problems, and, as a consequence, thoroughly revolutionized geometry. It also led to the creation of *new geometries*, like *algebraic geometry*, *non-euclidean geometry*, *projective geometry* and *riemannian geometry*; riemannian geometry is the basis for Einstein's theory of gravitation.

Beginnings of differential calculus. The basic question became clear slowly: If we are given a function $f(x)$, what is the tangent line to the graph of $y = f(x)$ at the point $(a, f(a))$? Descartes and Fermat tried to answer this question and succeeded in doing so for some functions. Descartes worked with polynomial functions while Fermat made no such restriction. For example, the Archimedean spiral in Cartesian coordinates is not an algebraic function.

If f is a polynomial, Descartes found a method of writing the *normal* to the curve at the point in question; the tangent is then the line in the orthogonal direction.

Dynamical interpretation of tangent lines. The idea is that the *instantaneous velocity* vector of a point moving according to the law

$$p(t) = (x(t), y(t))$$

at the time t_0 is

$$p'(t_0) = (x'(t_0), y'(t_0)).$$

The tangent to the curve described by $p(t)$ at $t = t_0$ has the slope

$$\frac{y'(t_0)}{x'(t_0)}.$$

For specific curves this method was applied successfully by *Gilles Personne de Roberval* (1602–1675).

The Kerala school, starting with Madhavan. *Madhavan* (c. 1350–1425) was the creator of a great school of mathematics and astronomy in Kerala, a srate in deep southwest of India. Current research has established that he and his disciples had discovered both differential and integral calculus, as well as the fact that differentiation and integration are inverse processes, at least for the circular arcs which are described by $\sqrt{1-x^2}$. They obtained the infinite series expansions of $\arctan x, \sin x, \cos x$ as well as constructed deep approximations to π. More about this later.

Homework # 4 : Due February 4, 2011

1. Ex. 6, p. 105.

2. Ex. 7, p. 105.

3. Ex. 17, p. 112.

4. Use the previous exercise to show that

$$\lim_{n\to\infty} \frac{1^k + 2^k + \ldots + n^k}{n^{k+1}} = \frac{1}{k+1}$$

   for any integer $k \geq 1$, a derivation attributed to Pascal in the book, although Madhavan knew of this long before Pascal.

5. Ex. 20. p. 117.

6. Ex. 4, p. 127

7. The *cycloid* is the trajectory of a point P on a circle that was initially at the origin when the circle starts rolling on a line with uniform speed. Prove that its parametric representation is

$$x(t) = a(t - \sin t), \quad y(t) = a(t - \cos t)$$

   where a is the radius of the circle. Hence compute the velocity vector at any t.

8. Compute the velocity vectors for the parabola $y^2 = 4px$ and ellipse $x^2/a^2 + y^2/b^2 = 1$.

9. Find a parametrization of the Archimedean spiral and hence calculate the velocity vector.

For the ambitious student

1. Show that in any medium light travels by straight lines as a consequence of Fermat's principle of least time.

2. If light travels from a point A in one medium to a point B in another medium, with speeds c_1 and c_2 respectively, show that the path is a broken line, the break occurring at the surface of separation of the two media. The exact break leads to Snell's law of refraction.

3. Find out about Pascal's mystic hexagon.

Week 5 : Infinite methods

Infinite series and products. Already in the seventeenth century, geometers found the Cartesian restriction to algebraic functions quite constraining. Thus, after the introduction of the logarithm by *John Napier* (1550–1617), it was found that the logarithm was *not* a polynomial function: in fact

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots + (-1)^{n-1}\frac{x^n}{n} + \ldots,$$

a series discovered by *Nicolas Mercator* ((1620–1687). He did this by computing

$$\int_0^x \frac{1}{1+t}dt,$$

not, as we do now, by integrating the geometric series

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \ldots + (-1)^n t^n + \ldots,$$

but by actually computing the integral as a limit of sums (see pages 162–163 of the book).

Wallis discovered an *infinite product* for π. It is

$$\frac{\pi}{2} == \prod_{k=1}^{\infty}\left(1+\frac{1}{2k-1}\right)\left(1-\frac{1}{2k+1}\right) = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdot\frac{8}{7}\cdot\frac{8}{9}\cdots.$$

This follows simply from Euler's infinite product for $\sin x/x$ by taking $x = \frac{\pi}{2}$. An infinite series for π is that attributed to Leibniz:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \ldots + (-1)^n\frac{1}{2n+1} + \ldots.$$

The simplest way to prove this is to integrate

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - \ldots$$

to get the series

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^n\frac{x^{2n+1}}{2n+1} + \ldots$$

giving the series for $\pi/4$ at $x = 1$. But this series is useless for numerical calculation as it converges very very slowly. For getting an approximation up to 2 decimals one needs to sum hundreds of terms. The simplest way to get a better series is to observe that since $\tan(\pi/6) = 1/\sqrt{3}$, one can get

$$\frac{\pi}{6} = \sqrt{\frac{1}{3}}\left\{1 - \frac{1}{(3)3} + \frac{1}{(3^2)5} - \frac{1}{(3^3)7}\cdots\right\}$$

which is rapidly convergent since the terms are *geometric*. There are even better ones (see homework problems). Leibniz himself regrouped the terms of the series to obtain series converging a bit faster (see p. 249 (21) of the book and problem # 7 in homework # 7.) The Madhavan school of Kerala mathematicians got the approximations

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \ldots + (1-)^{\frac{n-1}{2}}\frac{1}{n} + (1-)^{\frac{n+1}{2}}\frac{n^2+4}{2n(n^2+5)}.$$

For $n = 19$ this gives a value of π differing only by 1 in the 8^{th} decimal place, and for $n = 49$, it is accurate to 11 decimal places. These and other approximations of the kerala school are obtained as the convergents to the continued fraction

$$\frac{1}{2} \cdot \frac{1}{n+} \frac{1^2}{n^+} \frac{2^2}{n+} \frac{3^2}{n+} \cdots.$$

Similar continued fractions were discovered in Europe only in the 17^{th} century. Another famous rational approximation to π is 355/113.

The binomial series In the above infinite series the derivation and convergence are easily dealt with. This was not the case with the binomial series

$$(1+x)^a = 1 + \binom{a}{1}x + \binom{a}{2}x^2 + \ldots + \binom{a}{n}x^n + \ldots$$

If a is a positive integer this series terminates but otherwise is infinite. The series was first established by Newton for fractional exponents a. Its convergence needs $|x| < 1$. The case when $|x| = 1$ was not understood fully until Gauss studied this and more general series (the *hypergeometric series*).

Homework # 5 : Due February 11, 2011

1. Prove that the condition $|x| < 1$ is necessary for the convergence of the series for $\log(1+x)$.

2. Writing $a(u) = \arctan u$, and using the formula

$$\tan(A+B) = \frac{\tan A + \tan B}{1 - \tan A \tan B}$$

deduce that

$$a(u) + a(v) = a\left(\frac{u+v}{1-uv}\right)$$
$$a(u) + a(v) + a(w) = a\left(\frac{u+v+w-uvw}{1-uv-vw-wu}\right)$$

3. Use the previous exercise to verify (see H. Eves, *An introduction to the history of mathematics*, Holt, Rinehart, and Winston, 1976. pp. 96–102)
   (i) $\arctan\frac{1}{2} + \arctan\frac{1}{3} = \frac{\pi}{4}$
   (ii) $\arctan\frac{1}{2} + \arctan\frac{1}{5} + \arctan\frac{1}{8} = \frac{\pi}{4}$
   (iii) $4\arctan\frac{1}{5} - \arctan\frac{1}{239} = \frac{\pi}{4}$

(iv) $4 \arctan \frac{1}{5} - \arctan \frac{1}{70} + \arctan \frac{1}{99} = \frac{\pi}{4}$
(Hint : For (iii) and (iv), show first that $3 \arctan \frac{1}{5} = \arctan \frac{37}{55}$)

4. Ex. 2, p. 170

5. Ex. 3, p. 170

6. Ex. 4, p. 170

7. The *cissoid* is the curve
$$y = x^{3/2}(1-x)^{-1/2} \quad (0 < x < 1).$$
Prove that y has one derivative at $x = 0$ but not two, and goes to ∞ as $x \to 1 - 0$, with the line $x = 1$ as an asymptote. Show also that
$$\int_0^1 y dx$$
is convergent and its value is
$$\int_0^1 x^{3/2}(1-x)^{-1/2} = \frac{3\pi}{8}.$$
(*Hint*: Use the substitution $x = \sin^2 \theta$. Such integrals were first systematically studied by Euler, so they are called *Eulerian integrals*).

8. Show from the binomial series that
$$(1-x^2)^{-1/2} = 1 + \sum_{n=1}^{\infty} \frac{(2n)!}{2^{2n}(n!)^2} x^{2n}.$$

For the ambitious student

1. Show by explicit squaring of the infinite series that
$$\left(1-x^2)^{-1/2}\right)^2 = 1 + x^2 + x^4 + \ldots + x^{2n} + \ldots.$$

2. Define Euler's *Gamma function* $\Gamma(a)(a > 0)$ by
$$\Gamma(a) = \int_0^{\infty} e^{-t} t^{a-1} dt.$$
Prove that
$$\Gamma(a = 1) = a\Gamma(a), \quad \Gamma(n+1) = n!(n \text{ an integer } \geq 1), \qquad \Gamma(1/2) = \sqrt{\pi}.$$

(*Hint*: For the last formula, due to Euler, first take $t = u^2$ and write

$$\Gamma(1/2) = 2\int_0^\infty e^{-u^2} du$$

so that

$$(\Gamma(1/2))^2 = 2\int_{u>0.v>0} e^{-u^2-v^2} dudv$$

and use polar coordinates.)

3. Prove that

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad (a > 0, b > 0).$$

Hence derive the area of the cissoid.

Week 6 : Newton

Newton. *Isaac Newton* (1642–1727) is regarded as one of the greatest scientific figures of all time. His discovery of the Calculus, of the universal law of gravitation, of the corpuscular theory of light with new insights in both theory and experiment, and a host of other achievements, large and small, are the reasons behind the veneration he is accorded. It was not until Einstein came up with his epoch-making discoveries that a scientific figure was allowed to be mentioned on equal terms with Newton. This is not the place to go into areas of darkness in his persona cast by so much light. We shall confine ourselves to calculus.

He called his method the calculus of *fluxions*. It is basically what we teach students to day. He denoted the rate of change of a quantity f as $\dot{f}$ and called it the *fluxion* of f. His idea was that the quantity f evolved as expressed in terms of another variable, for example time, and the fluxion was the rate of change of this flux or flow. His method yielded that for a curve of equation

$$f(x, y) = 0$$

where f is a polynomial, we have

$$\dot{x}\frac{\partial f}{\partial x} + \dot{y}\frac{\partial f}{\partial y} = 0$$

giving

$$\frac{\dot{y}}{\dot{x}} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

He discovered the following as part of his calculus of fluxions.

1. The formula for the tangent line to an algebraic curve $f(x, y) = 0$ at any point of it
2. Fundamental Theorem of Calculus
3. Chain rule for differentiation and integration by substitution
4. Infinite series expansions and inversion or reversion of power series
5. Finding approximations to roots of equations

The scientific works. He was a student of Trinity College and became seriously interested in mathematics and the natural sciences from 1664. It was during the years 1664-1666, when Cambridge university was shut down because of the Great Plague, that Newton, in isolation in his country home in Lincolnshire, made his greatest discoveries. However he did not publish his results immediately and many of them were published decades after his original discoveries. His scientific papers and correspondence have been collected and published.

Calculus. He collected together all his work during the years 1664–1666 in *The October 1666 Tract on Fluxions*. When Mercator's work *Logarithmotechnia* on the power series for $\log(1 + x)$ appeared in 1668, Newton, alarmed that his priority to those and even more general results would not be recognized, wrote a brief account of his results called *De Analysi per Aequationes Numero Terminorum Infinitas*, and had it published in 1669. His entire work on the calculus was published in 1671 in *De Methodis Serierum et Fluxionum*.

Principia. This is an abbreviation of *Philosophiae Naturalis Principia Mathematica*, his magnum opus. It is written in a very opaque style that tends to mystify rather than illuminate, with all propositions proved as in Euclidean geometry. It is characteristic of his personality: secretive, fearful of losing his priority, and quarrelsome nature. It is nevertheless a monument to human discovery.

In the *Principia* published in 1687, he outlined his theory of mechanics of bodies under the influence of forces. He discovered his famous *laws of motion*, made the hypothesis of gravitational attraction between any two bodies and formulated his *law of universal gravitation.* As a consequence of his theory he worked out the planetary orbits and found them to be ellipses, and verified the three laws of planetary orbits first discovered empirically by *Johannes Kepler* (1571-1630). His theory put to rest all doubts about the heliocentric nature of the solar system and silenced all religious discussions about these matters, at least in the northern european countries including England. The law of universal gravitation states that there is a universal constant $G > 0$ such that any two bodies of masses m_1, m_2 and a distance $r > 0$ apart attract each other by a force of magnitude

$$\frac{m_1 m_2 G}{r^2}$$

and the direction of the force is along the line joining the two bodies (viewed as point masses).

De Quadratura Curvarum. This work, his last on calculus, was written during 1691–1693 and appeared as a mathematical appendix to his treatise *Opticks*. In it he worked the integrals of the form

$$\int x^\theta (e + f x^\eta) dx.$$

Opticks. Published in 1704, this contains all his discoveries on the nature of light, made during the *biennium mirabilissimum*, 1664-1666. He showed that white light can be split into a *spectrum* of colored lights and then recombined into white light again. He developed a *corpuscular theory* of light to explain his experiments

Newton today. If we examine what has happened today to the theories discovered by Newton, we see plainly the difference between mathematics and physics. In mathematics, once a theory is discovered, it is *permanent*. We still teach calculus according to Newton-Leibniz-Euler. Of course some of their arguments needed modifications to satisfy modern demands of rigor, but their theories remained basically intact. Subsequent theories are generalizations of earlier theories, such as Riemann's integration, and its generalization, Lebesgue's. As another instance, modern algebraic geometry consists in working over number systems far different from the real numbers, and for this, a generalization of calculus and differentials had to be invented, which would be compatible with the Newton-Leibniz theory over the reals (and not supplant it).

In physics however, new technology allows us a look into deeper layers of reality and it sometimes happens that the earlier theories are not adequate in explaining the new experiments, and entirely new approaches are needed. For instance, in the nineteenth century, *Augustin-Jean Fresnel* (1788–1827), discovered *wave properties* of light like diffraction and established a *wave theory of light* which overthrew Newton's particle theory. Then in the twentieth century Einstein showed that the *photo-electric effect* can be explained only on the assumption that light consists of particles (photons). This created a contradiction with wave theory which was not resolved until quantum mechanics showed that all particles possess both wave and particle properties. This *wave-particle duality* was established mathematically by Dirac who constructed the modern theory of radiation to explain transitions among the energy levels of an atom. The modern theory of light bears little resemblance to the theories of Newton or Fresnel.

The same fate overtook Newton's theory of gravitation. *Ole Christensen Rømer* (1644–1710), a Danish astronomer made the first quantitative measurements of the speed of light in 1676, and discovered that *light travels with a finite velocity.* This immediately dealt a fatal blow to the Newton's theory of gravitation where it is assumed that the forces act *instantaneously*. Since light signals are the speediest way to send information, it became clear that any theory of gravitation must incorporate this fact. The theory of electromagnetic waves developed by *James Clark Maxwell* (1831–1879) is a *field theory* in which forces are propagated with finite velocities. In a poll conducted among 100 greatest living physicists the following rating of physicists of all time was obtained:

1. Albert Einstein
2. Isaac Newton
3. James Clerk Maxwell
4. Niels Bohr
5. Werner Heisenberg
6. Galileo Galilei
7. Richard Feynman
8. Paul Dirac
8. Erwin Schrdinger
10. Ernest Rutherford

One should not take this ordering too seriously as it reflects to some extent the views and prejudices of contemporary physicists, but the first three places are not in dispute.

Nevertheless it should be noticed that the earlier theories are still valid to a great degree of accuracy in the regimes where they were first developed. Thus Newton's gravitation theory is still adequate for all rocket flights inside or near the solar system. What characterizes the newer theories is not just that they are more accurate, but they arise out of a radically new view point on space, time, and the physical events that take place in space-time. Here is the view of Brian Greene, from Columbia University, New York, USA:

> Einstein's special and general theories of relativity completely overturned previous conceptions of a universal, immutable space and time, and replaced them with a startling new framework in which space and time are fluid and malleable.

According to Newton, space and time are absolute, and events take place with such a space-time as background. In Einstein's theory, this is not so: space and time vary with the observer. In particular, time flows differently for different observers. The theories of motions of particles and gravitation arising from these foundations that Einstein created remain the most perfect of physical theories so far.

Detailed analysis of Newton's calculus contributions. We take one by one the various aspects of Newton's calculus work mentioned earlier.

1. The formula for the tangent line to an algebraic curve $f(x, y) = 0$ at any point of it. The main result is that if f is a polynomial and x, y are functions of t, then the fluxions $\dot{x}, \dot{y}$ arer related by

$$\dot{x}\frac{\partial f}{\partial x} + \dot{y}\frac{\partial f}{\partial y} = 0.$$

The tangent line to the graph of $f = 0$ at (a, b) has the slope

$$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} = -\frac{\partial f/\partial x}{\partial f/\partial y}.$$

His proof was to write (al la Fermat and Descartes) $x + \dot{x}o$ for x and $y + \dot{y}o$ for y and neglect second and higher powers of o, leading to the formula above (see the derivation in pages 193–194 of the book).

2. The fundamental theorem of calculus. This is the fundamental result because it connects, in a remarkable manner, differentiation with the much older theory of integration. The basic result is the following. If $f(x)$ is a function and $A(t)$ is the area of the graph of $y = f(x)$ between $x = 0$ and $x = t$, i.e.,

$$A(t) = \int_0^t f(x)dx,$$

then

$$\frac{dA}{dx} = f(x) = y.$$

This formula contains the quadrature results no powers of x obtained earlier. Its real meaning is that *Integration is the process inverse to differentiation.*

3. Chain rule for differentiation and integration by substitution. The chain rule is the following. If $y = f(g(x))$, then, writing $z = g(x)$,

$$\frac{\dot{y}}{\dot{x}} = \frac{\dot{y}/\dot{z}}{\dot{x}/\dot{z}}.$$

In current notation,

$$\frac{dy}{dx} = \frac{df(z)}{dz}\frac{dz}{dx} = f'(g(x))g'(x).$$

Newton applied this to several examples. Since integration is the inverse of differentiation, the chain rule implies a formula for integration:

$$\int_0^x f'(g(t))g'(t)dt = f(g(x)) - f(g(0)).$$

Here it should be noted that Newton regarded the integral as an *indefinite integral*, with an arbitrary lower limit, which he often assumed to be zero.

4. Infinite series expansions and inversion or reversion of power series. Newton was the first to realize the importance of power series expansions for theoretical as well as empirical purposes.

First he discovered the binomial series for arbitrary, not just positive integral exponents. One knew, from Pascal's law, that if α is a positive integer,

$$(1+x)^\alpha = 1 + \binom{\alpha}{1}x + \binom{\alpha}{2}x^2 + \ldots + x^n = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \ldots + x^n.$$

The coefficients are found by applying Pascal's law of triangles. Newton discovered the series by assuming that the Pascal law remains valid no matter what the exponent is, except that one will find that the coefficients never vanish, the coefficient of x^r being

$$\frac{\alpha(\alpha-1)\alpha-2)\ldots(\alpha-r+1)}{r!}$$

which will never be 0 if α is not a positive integer; if α *is* a positive integer, the expression vanishes for $r \geq n+1$. Newton's proof of this was not complete and involved very complicated calculations supplemented by various assumptions which were not justified, but strengthened by various special cases. To prove it from our point of view write

$$y_\alpha = (1+x)^\alpha = 1 + a_1(\alpha)x + a_2(\alpha)x^2 + \ldots + a_n(\alpha)x^n + \ldots.$$

Since $dy_\alpha/dx = \alpha y_{\alpha-1}$ we have

$$(n+1)a_{n+1}(\alpha) = \alpha a_n(\alpha - 1) \quad (n = 1, 2, \ldots), \quad a_1(\alpha) = \alpha.$$

These relations are *recursive* and allow us to calculate the $a_n(\alpha)$ *recursively*, i.e., one after the other. It is straightforward to check that

$$a_r(\alpha) = \binom{\alpha}{r} = \frac{\alpha(\alpha-1)\alpha-2)\ldots(\alpha-r+1)}{r!} \qquad (1).$$

The proof is *still not complete*; one must still show that the series derived actually sums to $(1+x)^\alpha$. This requires additional ideas.

Newton derived many series expansions. For instance he found

$$\arcsin x = x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \frac{5}{112}x^7 + \ldots.$$

This can be obtained by noting that

$$\arcsin x = \int_0^x \frac{dt}{\sqrt{1-t^2}}$$

and then using the expansion

$$\frac{1}{\sqrt{1-t^2}} = \sum_{n=0}^{\infty} \frac{(2n)!}{(n!)^2 2^{2n}} x^{2n} \qquad (2)$$

obtained by the binomial series and integrating tem by term. The result is

$$\arcsin x = \sum_{n=0}^{\infty} \frac{(2n)!}{(n!)^2 2^{2n}} \frac{x^{2n+1}}{2n+1} \qquad (3).$$

Notice that this gives not only the first few terms written above but also the *general term*.

If a function y has an expansion

$$y = a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots$$

one can ask if the inverse function can be expressed as a series and if so what that series is? At the *formal evel* this is asking for a power series expansion for x as

$$x = b_1 y + b_2 y^2 + \ldots + b_n y^n + \ldots$$

such that is this is substituted for x in the series for y we get just y. We assume that the series have no constant terms; this assumption implies that if we want to compute b_n, we need only use the terms up to y^n in the series for y. Newton used this method to start from $y = \arcsin x$ in its series form, to get the series for $x = \sin y$, as

$$x = \sin y = y - \frac{y^3}{3!} + \frac{y^5}{5!} - \ldots + (-1)^n \frac{y^{2n+1}}{(2n+1)!} + \ldots \tag{4}$$

To get the general term by this process seems impossible and Newton did not get it; it appears he seemed certain of what the general term is.

5. Finding approximations to roots of equations. It is known that very few equations can be solved by explicit formulae. It is more practical to look for approximate solutions. The most famous method is Newton's. Let the equation to be solved be $f(x) = 0$. If x_n is a given approximation to a root, then we write $x_* = x_n + p$ and find that

$$0 = f(x_*) = f(x_n + p) = f(x_n) + pf'(x_n) + \ldots .$$

Hence neglecting terms of order p^2 and higher we get

$$p \approx -\frac{f(x_n)}{f'(x_n)}$$

giving

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

If we start with a first approximation that is close, this process will converge usually.

The same idea can be applied to solve equations

$$f(x, y) = 0$$

to get y as a power series in x. Newton illustrates this with

$$y^3 + a^2 y + axy - x^3 - 2a^3 = 0.$$

Homework # 6 : Due February 18, 2011

1. Ex. 2, p. 194.
2. Ex. 4, p. 198
3. Ex. 9. p. 203
4. Ex. 10, p. 206
5. Ex. 11, p. 206

6. Prove the formula (1) for the coefficients of $(1+x)^\alpha$.

7. Prove (2) and (3).

8. Show that

$$\tan y = y + \frac{y^3}{3} + \frac{2}{15}y^5 + \dots$$

by inverting the series for $y = \arctan x$. Why is it that only odd powers of y will appear in the series for $\tan y$?

For the ambitious student

1. Prove, using Taylor's series with Lagrange's remainder, that the binomial series for $(1+x)^\alpha$ actually converges to $(1+x)^\alpha$ for $|x| < 1$.

2. Find the power series for $\sin x, \cos x$ by Taylor series with Lagrange's remainder.

Week 7 : Leibniz

Leibniz. *Gottfried Wilhelm Leibniz* (1646–1716) Leibniz is one of the greatest philosopher-scientists of the seventeenth century. He was a diplomat, historian, archivist, and at the same time did some of the greatest mathematics of his time. He is, as modern research has now shown without any question, the independent co-discoverer of Calculus along with Newton. In the eighteenth century, Newton and his sycophants launched a tainted investigation of Leibniz's discoveries and issued a report, almost certainly written by Newton himself, with the conclusion that Leibniz plagiarized Newton. This absurd conclusion is now rejected by historians. Leibniz himself, in his famous book *Historia et origo calculi differentialis* gives a masterly exposition of his discoveries beginning with these eloquent words (see A. Weil, Collected Papers, Vol III, pp. 366–378):

> It is most useful that the true origins of memorable inventions be known, especially of those which were conceived not by accident but by an effort of meditation. The use of this is not merely that history may give everyone his due and others be spurred by the expectation of similar praise, but also that the art of discovery may be promoted and its method become known through brilliant examples. One of the noblest inventions of our time has been a new kind of mathematical analysis, known as the differential calculus; but, while its substance has been adequately explained, its source and original motivation have not been made public. It is almost forty years now that its author invented it..... . ..

The difference with Newton is striking. Leibniz's was an open personality unlike Newton's, and he had a great mastery of the art of exposition, as would be obvious from his career as a lawyer, logician, diplomat, and archivist. Newton wrote two letters to Leibniz and Leibniz replied to these. The difference between the two men can be seen in a striking manner in this correspondence.

So far as the calculus was concerned, Leibniz obtained all the fundamental results and did everything that Newton did by methods which were based on completely different ideas, and without any knowledge of Newton's methods, and in many cases, even of his results. Indeed, Leibniz published his work in 1684 before Newton's *Principia* which appeared in 1687. His basic idea of *differentials*, obscure at that time, has been understood by now. The differentials emerge in two different modern incarnations: in the concept of *exterior differential forms*, discovered by *Elie Cartan* (1869–1951), which forms the basis of all topology and geometry, and in *non-standard analysis*, a new subject created by *Abraham Robinson* (1918–1974). The Leibniz theory of differentials treated expressions like dx, dy as objects with definite rules of operation; the derivative dy/dx was a result of these operations.

The most convincing defense for Leibniz, if one is even needed, is that today we use his ideas and notations universally. It is also true that English mathematics, by the insistence of English mathematicians of Newton's priority, and their consequent ignoring of continental mathematics and mathematicians, suffered terribly. England became a backwater in mathematics, and most of the progress in mathematics came from european sources, especially France and Germany.

The characteristic triangle. This refers to the right-angled triangle associated with a function $y = f(x)$, of sides dx, dy and hypotenuse ds (see fig. 3, p. 241 in the book). The normal to the curve at (x, y) meets the x-axis at some point and n is the length of the segment from (x, y) to this point. Then, as a simple consequence of similarity of triangles one gets, a la Leibniz,

$$\frac{ds}{n} = \frac{dx}{y} \Rightarrow yds = ndx \qquad (1).$$

Summing the infinitesimals according to his scheme, Leibniz gets

$$\int y ds = \int n dx \qquad (2).$$

Proceeding further, let t be the length of the segment of the tangent line to the curve at (x, y) between the point where this tangent line meets the x-axis and the point where it meets a fixed vertical line (see fig. 4 on page 242 of the book). Let a be the length of the segment of this vertical line from its base on the x-axis to the point where it meets the tangent line. **In the book the vertical line is mistakenly denoted by a.** Then, by a similar triangle argument, one gets

$$\frac{ds}{t} = \frac{ds}{a} \Rightarrow a ds = t dy \qquad (3)$$

leading to

$$\int a ds = \int t dy \qquad (4).$$

All the results of quadrature and arc length computations can be obtained from these results, as Leibniz himself demonstrated.

Transformations and transmutations. Unlike Newton (at least in calculus) Leibniz was focused on *general methods and concepts* and not on *specific calculations*. He made the great discovery that the symbol $f(x)dx$ remains *invariant* if we substitute $x = g(t)$:

$$f(x)dx = f(g)dg = f(g(t)g'(t)dt \qquad (5)$$

which gives, on integration, the change of variables formula:

$$\int f(x)dx = \int f(g(t))g'(t)dt.$$

If there is more than one variable, say $x_1, x_2, \ldots x_n$ the Leibniz differentials $dx_1, dx_2, \ldots, dx_n$ can be combined to form a *Grassmann algebra* with the rules

$$dx_i dx_i = 0, \qquad dx_i dx_j = -dx_j dx_i.$$

The result of the invariance of $dx_1 \ldots dx_n$ is then the change of variables formula for n-dimensional integrals!!

Leibniz described the advantages of his transmutation theorem in his letter to Newton (see p. 245 of the book). The formula (19) on page 247 of the book is the *transmutation theorem.* He applied it, among other things, to get the quadrature of the circle and his famous series

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$$

and the less known

$$\begin{aligned} \frac{\pi}{8} &= \left(\frac{1}{2} - \frac{1}{6}\right) + \left(\frac{1}{10} - \frac{1}{14}\right) + \left(\frac{1}{18} - \frac{1}{22}\right) + \ldots = \frac{1}{3} + \frac{1}{35} + \frac{1}{99} + \ldots \\ \frac{\pi}{8} &= \frac{1}{1.3} + \frac{1}{1.3} + \frac{1}{1.3} + \ldots. \end{aligned} \qquad (6)$$

The transmutation theorem also led to the quadrature of the cycloid (see pp. 250–251 of the book). See pages 246–247 for the derivation of the transmutation theorem.

In the modern view the difference between the Newtonian and Leibnizian points of view can be summarized as follows. Newton worked with *vector fields* while Leibniz worked with *differential forms*. The reason why the Leibnizian point of view is more fundamental is that differential forms behave naturally under a change of context while vector fields do not*.

In addition to all of these ideas, Leibniz was responsible for our modern understanding that we integrate differentials and not functions, and that it is the differentials, not functions, that transform naturally under changes of variables. These are among his greatest achievements and are the reason why his impact has been so great and so permanent.

Homework # 7 : Due February 25, 2011

1. Ex. 9, p. 242

2. Ex. 10, p. 243 (Recall interpretation of formula (14) of the book on p. 243

3 . Obtain the formula

$$ds = \sqrt{\left(1 + \frac{dy}{dx}^2\right)}\, dx$$

for the infinitesimal arc length from (14) of p. 243.

4. Ex. 12, page 249.

5. Ex. 13, p. 251.

6. Ex. 18, p. 261.

7. Verify that the general terms of the two series for $\pi/8$ are respectively

$$\left(\frac{1}{8k+2} - \frac{1}{8k+6}\right), \qquad \frac{1}{(4k+1)(4k+3)}$$

and obtain (6) by grouping the terms suitably in the Leibniz series for $\pi/4$.

For the ambitious student

1. Let A be the set of polynomials in two variables t, u over a field F and D a linear map $A \longrightarrow A$ such that

$$D(fg) = Df{\cdot}g + f{\cdot}Dg.$$

* I owe this remark to Professor J. A. C. Kolk.

Prove that $D1 = 0$ and that

$$D^n(fg) = \sum_{r=0}^{n} \binom{n}{r} D^r f \cdot D^{n-r} g.$$

If $Dt = \alpha, Du = \beta$, then prove that

$$D = \alpha \frac{\partial}{\partial t} + \beta \frac{\partial}{\partial u}.$$

2. If $x_1, \ldots, x_n$ are n variables and $y_i = f_i(x_1, \ldots, x_n)(1 \le i \le n)$, show that

$$dy_1 \ldots dy_n = J dx_1 \ldots dx_n$$

where J is the *Jacobian* $\det(\partial f_i / \partial x_j)_{1 \le i,j \le n}$. The differentials dx_i are *anti-commutative*, i.e.,

$$dx_i^2 = 0, \quad dx_i dx_j = -dx_j dx_i \quad (i \ne j).$$

Week 8 : Euler

Euler. *Leonhard Euler* (1707–1783) inaugurated the modern era of the calculus. He revolutionized it completely, introducing new ideas, discovering and proving thousands of new results. By combining calculus with infinite processes like infinite series and products, he created a new subject, *mathematical analysis.* He wrote a monumental two-volume treatise *Introductio in analysin infinitorum* on analysis which has served as the blue prints for *all* works on analysis and calculus ever since. Even today the *Introductio* is better than any contemporary text on calculus and analysis.

He was the greatest mathematician and natural philosopher of the 18^{th} century and one of the greatest of all time. He worked on all branches of mathematics, both pure and applied, known in his time. He created new branches of mathematics like combinatorial topology, graph theory, and the calculus of variations. He was the founder of modern differential and integral calculus as we know them to day, and his books introduced algebra and calculus and their applications to enormous numbers of students. It could be said without exaggeration that he did to analysis what Euclid did to geometry, except that Euler himself created a huge part of what went into his books, unlike Euclid who created a synthesis of existing knowledge of geometry in his time.

Scientific work. Opera Omnia. In 1907 the Swiss Academy of Sciences established the **Euler Commission** with the charge of publishing the complete body of work consisting of all of his papers, manuscripts, and correspondence. This project, known as *Opera Omnia*, began in 1911 and is still in progress. His scientific publications, not counting his correspondence, run to over 70 volumes, each between approximately 300 and 600 pages. Thousands of pages of handwritten manuscripts are still not in print.

Here is a list, which is at best partial, of topics he worked on in his life, many of which were founded by him and in almost all of which his work was pioneering:

Differential and integral Calculus
Logarithmic, exponential, and trigonometric functions
Differential equations, ordinary and partial
Elliptic functions and integrals
Hypergeometric integrals
Classical geometry
Number theory/Algebra
Continued fractions
Infinite series and products/Zeta and other products
Divergent series
Mechanics of particles/ solid bodies
Calculus of variations
Optics (theory and practice)
Hydrostatics/Hydrodynamics
Astronomy
Lunar and planetary motion
Topology
Graph theory

His books *Introductio in Analysin Infinitorum* and *Mechanica* which went through many editions, brought calculus and mechanics to the entire scientific world. There was no longer any necessity of reading the obscure and opaque papers of Leibniz or Newton. The *Introductio*, together with his books on calculus *Institutiones calculi differentialis* and *Institutiones calculi integralis* are the source of all modern books on analysis and calculus.

With its emphasis on concepts and structures, modern mathematics has mostly shied away from formulae. Nevertheless, many of the greatest peaks of mathematics are described by formulae of one sort or the other and Euler had his fair share of them.

$$e^{ix} = \cos x + \sin x$$
$$e^{i\pi} = -1$$
$$\log(-1) = i\pi + 2ki\pi (k = 0, \pm 1, \pm 2, \ldots)$$
$$\frac{\sin x}{x} = \prod_{n=1}^{\infty}\left(1 - \frac{x^2}{n^2\pi^2}\right) \qquad \cos x = \prod_{n=1}^{\infty}\left(1 - \frac{4x^2}{(2k+1)^2\pi^2}\right)$$
$$1 + \frac{1}{2^k} + \frac{1}{3^k} + \ldots + \frac{1}{n^{2k}} + \ldots = \frac{(-1)^{k-1}2^{2k-1}B_{2k}}{(2k)!}\ \pi^{2k}$$
$$\frac{\pi}{\sin s\pi} = \frac{1}{s} + \sum_{n=1}^{\infty}(-1)^n\left(\frac{1}{n+s} - \frac{1}{n-s}\right)$$
$$\pi \cot s\pi = \frac{1}{s} + \sum_{n=1}^{\infty}\left(\frac{1}{n+s} - \frac{1}{n-s}\right)$$
$$\frac{\pi}{3\sqrt{3}} = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{5} + \frac{1}{7} - \frac{1}{8} + \frac{1}{10} - \ldots$$
$$\frac{\pi}{2\sqrt{2}} = 1 + \frac{1}{3} - \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \frac{1}{11} - \ldots$$
$$\frac{\pi}{3} = 1 + \frac{1}{5} - \frac{1}{7} - \frac{1}{11} + \frac{1}{13} + \frac{1}{17} \ldots$$
$$\frac{\pi^2}{8\sqrt{2}} = 1 - \frac{1}{3^2} - \frac{1}{5^2} + \frac{1}{7^2} + \frac{1}{9^2} - \ldots$$
$$\frac{\pi^2}{6\sqrt{3}} = 1 - \frac{1}{5^2} - \frac{1}{7^2} + \frac{1}{11^2} + \frac{1}{13^2} - \ldots$$

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1}$$

$$\frac{1 - 2^{m-1} + 3^{m-1} - \text{etc}}{1 - 2^{-m} + 3^{-m} - \text{etc}} = -\frac{1.2.3.\ldots(m-1)(2^m - 1)}{(2^{m-1} - 1)\pi^m \cos\frac{m\pi}{2}}$$

$$\sum_{n>m>0} \frac{1}{n^2 m} = \sum_{n>0} \frac{1}{n^3}$$

$$\prod_{n=1}^{\infty}(1 - x^n) = 1 + \sum_{n=1}^{\infty}(-1)^n \left(x^{\frac{3n^2-n}{2}} + x^{\frac{3n^2+n}{2}}\right)$$

$$1 - 1!x + 2!x^2 - 3!x^3 + \ldots = \frac{1}{1+}\,\frac{x}{1+}\,\frac{x}{1+}\,\frac{2x}{1+}\,\frac{2x}{1+}\,\frac{3x}{1+}\,\frac{3x}{1+}\ \text{etc}$$

$$1 - 1! + 2! - 3! + \ldots = 0.596347362123\ldots.$$

$$\begin{aligned}\sum_{k=0}^{m} f(k) = \int_0^m f(x)dx + \frac{1}{2}(f(0) + f(m)) \\ + \sum_{k\geq 1} \frac{B_{2k}}{(2k)!}\left(f^{(2k-1)}(m) - f^{(2k-1)}(0)\right)\end{aligned}$$

$$\frac{\partial}{\partial y}F(x, y, y') = \frac{d}{dx}\left(\frac{\partial}{\partial y'}F(x, y, y')\right)$$

Opera Omnia. It is in four series.

- I. Series prima: Opera Mathematica (29 vols), 14042 pages
- II. Series secunda: Opera mechanica et astronomica (30 vols), 10658 pages
- III. Series tertia: Opera Physica, Miscellena (12 vols), 4331 pages
- IV A. Series quarta A: Commercium Epistolicum (7 vols, 1 in preparation), 2498 pages
- IV B. Series quarta B: Manuscripta (unpublished manuscripts, notes, diaries, etc)

• Internet Sources: There is a monumental project at Dartmouth to bring the entire Opera Omnia into the net for universal accessibility. The URL for this is

http://www.math.dartmouth.edu/∼euler/index.html

Trigonometric and exponential functions. Euler was the first to define analytically the exponential and trigonometric functions, even for complex arguments. He obtained the famous formula

$$e^{ix} = \cos x + i \sin x.$$

By writing

$$e^{ix} = \lim_{N\to\infty} (1 + (ix/N))^N$$

we get

$$\cos x = \lim_{N\to\infty} \frac{1}{2}\Big((1 + (ix/N))^N + (1 - (ix/N))^N\Big)$$
$$\sin x = \lim_{N\to\infty} \frac{1}{2i}\Big((1 + (ix/N))^N - (1 - (ix/N))^N\Big).$$

He always wrote these formulae without the limit sign; for him N was already an infinitely large number, while in other formulae, especially in differential calculus, when one has to take limits as a parameter tends to 0, he wrote ϵ for an infinitely small number. The series expansions for the exponential and the trigonometric functions follow by binomial expansion.

Let z_n be a *bounded* sequence of complex numbers; this means that there is a constant $K > 0$ such that $|z_n| < K$ for all n. By the binomial theorem,

$$\left(1 + \frac{z_n}{n}\right)^n = 1 + \frac{z_n}{n} + \binom{n}{2}\frac{z_n^2}{n^2} + \dots.$$

The r^{th} term is

$$\binom{n}{r}\frac{z_n^r}{n^r} = \frac{n(n-1)\dots(n-r+1)}{r!n^r} z_n^r = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\left(1 - \frac{r-1}{n}\right)\frac{z_n^r}{r!}.$$

If we now assume that $z_n \to z$ then we get, for r *fixed and* $n \to \infty$,

$$\lim_{n\to\infty}\binom{n}{r}\frac{z_n^r}{n^r} = \frac{z^r}{r!}.$$

Hence, *formally* (since we have not justified the limit process)

$$e^z = 1 + z + \frac{z^2}{2!} + \dots + \frac{z^r}{r!} + \dots.$$

Taking $z = ix$ and splitting into real and imaginary parts we get the series for $\cos x$ and $\sin x$.

Infinite product formulae for $\sin x$ and $\cos x$. We have seen above that

$$\cos x = \lim_{N\to\infty} \frac{1}{2}\Big((1 + (ix/N))^N + (1 - (ix/N))^N\Big)$$
$$\sin x = \lim_{N\to\infty} \frac{1}{2i}\Big((1 + (ix/N))^N - (1 - (ix/N))^N\Big).$$

By actually *factorizing* the polynomials on the right side for *finite* N as product of linear factors he found his famous *infinite product formula* for sin and cos:

$$\frac{\sin x}{x} = \prod_{n=1}^{\infty}\left(1 - \frac{x^2}{n^2\pi^2}\right) \qquad \cos x = \prod_{n=1}^{\infty}\left(1 - \frac{4x^2}{(2k+1)^2\pi^2}\right) \tag{1}$$

In spite of the modern proofs of these products by complex methods, Euler's proofs, with some added remarks on uniform convergence (which was understood only after *Weierstrass* (1815–1897)), are still the most direct and beautiful.

The zeta values. *What is the value of*

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots ?$$

This problem, known as the *Basel problem* because the Bernoullis had made a first attempt at its solution, was the great unsolved question when Euler was at the start of his career. Following Dirichlet and Riemann (who came after Euler) we write

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \ldots$$

and call it the *zeta function.* So the question is the value of $\zeta(2)$. Even numerical *bounds* for the sum were hard to come by because of its slow convergence; one has to take 100 terms to get a result accurate to two decimals. By a brilliant transformation of the series Euler could obtain a value accurate to 6 decimals but the exact value eluded him for a while.

Evaluation of $\zeta(2)$ and $\zeta(4)$. It is easy to see that

$$\frac{1}{n+1} < R_n = \frac{1}{(n+1)^2} + \frac{1}{(n+1)^2} + \ldots < \frac{1}{n}.$$

So if one wants to calculate $\zeta(2)$ accurately up to 6 decimals we must take $n > 10^6$, i.e., we must take a million terms. This must have been daunting even to Euler who was an indefatigable calculator. In his very first contribution to the subject in 1731 entitled *De Summatione Innumerabilium Progressionum*, I-14, 25-41, Euler discovered a transformation of this series that allowed him to calculate $\zeta(2)$ to 6 decimal places, and even more accurately had he wanted to do so. Euler proved that

$$\zeta(2) = \log u \log(1-u) + \sum_{n=1}^{\infty} \frac{u^n}{n^2} + \sum_{n=1}^{\infty} \frac{(1-u)^n}{n^2} \qquad (0 < u < 1).$$

In particular, taking $u = 1/2$, we get

$$\zeta(2) = (\log 2)^2 + \sum_{n=1}^{\infty} \frac{1}{n^2.2^{n-1}}.$$

Euler's argument is the following. Start with

$$-\frac{\log(1-x)}{x} = 1 + \frac{x}{2} + \frac{x^2}{3} + \ldots$$

Integrating from 0 to 1 we have,

$$\zeta(2) = \int_0^1 \frac{-\log(1-x)}{x} dx.$$

We split the integration from 0 to u and u to 1. So for any u with $0 < u < 1$,

$$\zeta(2) = \int_0^u \frac{-\log(1-x)}{x} dx + \int_u^1 \frac{-\log(1-x)}{x} dx.$$

In the second integral we change x to $1-x$ to get

$$\zeta(2) = \int_0^u \frac{-\log(1-x)}{x} dx + \int_0^{1-u} \frac{-\log x}{1-x} dx = I_1 + I_2.$$

For I_1 we expand $\log(1-x)$ as a series and integrate term by term from 0 to u. So

$$I_1 = \sum_{n=1}^{\infty} \frac{u^n}{n^2}.$$

To evaluate I_2 we integrate by parts first to get

$$I_2 = \log x \log(1-x)\Big|_0^{1-u} + \int_0^{1-u} \frac{-\log(1-x)}{x} dx$$

The second term is like I_1. For the first term note that

$$\log(1-x) = -x - \frac{x^2}{2} - \ldots \sim -x \qquad (x \to 0)$$

so that

$$\log x \log(1-x) \sim -x \log x \to 0 \qquad (x \to 0).$$

Hence we obtain

$$\zeta(2) = \log u \log(1-u) + \sum_{n=1}^{\infty} \frac{u^n}{n^2} + \sum_{n=1}^{\infty} \frac{(1-u)^n}{n^2}$$

which is Euler's result. For numerical evaluation Euler used the formula for $\log 2$ obtained earlier,

$$\log 2 = \sum_{n=1}^{\infty} \frac{1}{n.2^n}.$$

Using this Euler computes

$$(\log 2)^2 = 0.480453\ldots, \qquad \sum_{n=1}^{\infty} \frac{1}{n^2.2^{n-1}} = 1.164481\ldots$$

to obtain

$$\zeta(2) = 1.644934\ldots.$$

The above transformation used by Euler is the first instance of the appearance of the *dilogarithm.* Let

$$\mathrm{Li}_2(x) := \int_0^x \frac{-\log(1-t)}{t} dt = \int_{x>t_1>t_2>0} \frac{dt_1 dt_2}{t_1(1-t_2)}.$$

Then

$$\mathrm{Li}_2(x) = \sum_n \frac{x^n}{n^2}, \qquad \mathrm{Li}_2(1) = \zeta(2).$$

The discussion above is essentially the proof of the functional equation

$$\mathrm{Li}_2(x) + \mathrm{Li}_2(1-x) = -\log x \log(1-x) + \mathrm{Li}_2(1) \qquad (0 < x < 1)$$

which also is in Euler (loc. cit). The dilogarithm has obvious generalizations to several variables and these play an important role in the theory of zeta values.

So matters stood till suddenly and unexpectedly, around 1735, Euler had a stroke of inspiration that led him to the exact value of $\zeta(2), \zeta(4)$, etc. Here is the translation by A. Weil of the opening lines of Euler's paper:

So much work has been done on the series $\zeta(n)$ that it seems hardly likely that anything new about them may still turn up... I, too, in spite of repeated efforts, could achieve nothing more than approximate values for their sums... Now, however, quite unexpectedly, I have found an elegant formula for $\zeta(2)$, depending on the quadrature of a circle, namely, upon π.

Across a gulf of centuries, this passage, and indeed the whole paper, still conveys the excitement Euler must have felt on his discovery.

Euler's method was an audacious, one might say, even reckless, generalization of Newton's theorem on the symmetric functions of the roots of a polynomial, to an *infinite power series.* If we have a polynomial $f(x)$ written in the form

$$f(x) = 1 + a_1 x + a_2 x^2 + \ldots + a_k x^k$$

and if $r_1, r_2, \ldots, r_k$ are the roots of the equation $f(x) = 0$, we have, since none of these can be 0 (why?),

$$f(x) = \left(1 - \frac{x}{r_1}\right)\left(1 - \frac{x}{r_2}\right) \ldots \left(1 - \frac{x}{r_k}\right)$$

Euler assumed that if we replace $f(x)$ by a power series

$$f(x) = 1 + a_2 x^2 + a_4 x^4 + \ldots \tag{1}$$

and if $\pm r_1, \pm r_2, \ldots$ are the roots of the equation $f(x) = 0$, then

$$f(x) = \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{r_k^2}\right) \qquad (2).$$

Hence we may conclude that

$$\sum_{k=1}^{\infty} \frac{1}{r_k}^2 = -a_2.$$

Moreover, we get from (2) that

$$\sum_{j \neq k} \frac{1}{r_j^2} \frac{1}{r_k^2} = a_4.$$

But

$$2\sum_{j\neq k}\frac{1}{r_j^2}\frac{1}{r_k^2}=\left(\sum_k \frac{1}{r_k^2}\right)^2-\sum_{k=1}^{\infty}\frac{1}{r_k^4}.$$

This gives

$$\sum_{k=1}^{\infty}\frac{1}{r_k^4}=a_2^2-2a_4.$$

By using this argument Euler succeeded in determining the exact value of $\zeta(2)$ and $\zeta(4)$. In a letter to Daniel Bernoulli he communicated his formulae

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots=\frac{\pi^2}{6} \qquad 1+\frac{1}{2^4}+\frac{1}{3^4}+\ldots=\frac{\pi^4}{90}.$$

The essence of Euler's argument is to take

$$f(x)=\frac{\sin x}{x}$$

in the above discussion. It vanishes for

$$x=\pm k\pi \qquad (k=1,2,\ldots)$$

and has the infinite series expansion

$$1-\frac{1}{3!}x^2+\frac{1}{5!}x^5-\ldots.$$

By the above discussion we can then write

$$\frac{\sin x}{x}=\prod_{n=1}^{\infty}\left(1-\frac{x^2}{n^2\pi^2}\right) \tag{3}$$

Hence

$$\sum_{k=1}^{\infty}\frac{1}{k^2\pi^2}=\frac{1}{6}$$

or

$$\sum_{k=1}^{\infty}\frac{1}{k^2\pi^2}=\frac{\pi^2}{6}!!$$

In our special case this means that

$$\sum\frac{1}{k^4\pi^4}=\frac{1}{36}-\frac{1}{60}=\frac{1}{90}$$

or

$$\sum_{n=1}^{\infty}\frac{1}{n^4}=\frac{\pi^4}{90}.$$

Among the many objections raised was this: how does one know that the roots $\pm k\pi$ are the only ones for the function $\sin x/x$, and even more damagingly, how does one know that there is is not a factor that vanishes nowhere, like the exponential function, in the right side of (2)? Although these questions raised by the Bernoulli brothers and others regarding the validity of Euler's method in deriving these results were substantial, the numerical calculations essentially confirmed his results and Euler's reputation as a mathematician of the first rank was established. Euler himself was aware that the objections to his derivation were grave and legitimate and so he continued to work on meeting these objections and making his arguments more solid, at least by the standards of his era (actually, even by modern standards once we have the notion of uniform convergence, see earlier comment). He succeeded in doing this around 1742 when he proved the infinite product expansion

$$\frac{\sin x}{x} = \prod_{n=1}^{\infty}\left(1 - \frac{x^2}{n^2\pi^2}\right) \tag{3}$$

from which, by a generalization of the above calculations for $\zeta(2), \zeta(4)$, he could justify completely the values calculated earlier on :

$$\zeta(2) = \frac{\pi^2}{6}\ \zeta(4) = \frac{\pi^4}{90}\ \zeta(6) = \frac{\pi^6}{945}\ \zeta(8) = \frac{\pi^8}{9450}\ \zeta(10) = \frac{\pi^{10}}{93555}\ \zeta(12) = \frac{\pi^{12}}{6825 \times 93555}.$$

It is clear from (1) that the objections have been answered. Later he evaluated all the zeta values for *even positive integers* and tried all his life to evaluate the zeta values for odd positive integers (≥ 3). He was not successful. Only in recent times was it possible even to prove that $\zeta(3)$ is *irrational*, as was done by *Roger Apéry*(1916–1994). The question of the odd zeta values is still open, even the irrationality of $\zeta(5)$.

Factorization using roots of unity. The roots of the equation

$$T^n - 1 = 0$$

are

$$e^{\frac{2r\pi i}{n}} \qquad (r = 0, 1, 2, \ldots, n-1).$$

However this equation has real coefficients and so the roots can be arranged in pairs where each pair consists of conjugate quantities. Thus, if $n = 2p + 1$, then the roots can be written as

$$e^{\frac{2r\pi i}{n}} \qquad (r = 0, \pm 1, \pm 2, \ldots, \pm p).$$

On the other hand

$$(T - e^{\frac{2r\pi i}{n}})(T - e^{\frac{-2r\pi i}{n}}) = T^2 - 2T\cos\frac{2r\pi}{n} + 1$$

from which we obtain the factorization

$$T^n - 1 = (T - 1)\prod_{k=1}^{p}\left(T^2 - 2T\cos\frac{2k\pi}{n} + 1\right) \qquad (n = 2p + 1).$$

Writing $T = X/Y$ and clearing of fractions we get

$$X^n - Y^n = (X - Y)\prod_{k=1}^{p}\left(X^2 - 2XY\cos\frac{2k\pi}{n} + Y^2\right).$$

Let us now take

$$X = 1 + \frac{ix}{n}, \qquad Y = 1 - \frac{ix}{n}$$

and write

$$q_n(x) = \frac{\left(1 + \frac{ix}{n}\right)^n - \left(1 - \frac{ix}{n}\right)^n}{2ix}.$$

Then we get, since $X - Y = \frac{2ix}{n}$, after a little calculation,

$$q_n(x) = C_n \prod_{k=1}^{p} \left(1 - \frac{x^2}{n^2} \frac{1 + \cos\frac{2k\pi}{n}}{1 - \cos\frac{2k\pi}{n}}\right)$$

where C_n is a numerical constant. Since

$$q_n(x) = 1 + \ldots$$

while the product on the right side of the equation above takes the value C_n at $x = 0$, we must have $C_n = 1$. It is an elementary exercise to verify this directly also. Thus we finally obtain

$$q_n(x) = \prod_{k=1}^{p} \left(1 - \frac{x^2}{n^2} \frac{1 + \cos\frac{2k\pi}{n}}{1 - \cos\frac{2k\pi}{n}}\right).$$

It only remains to let n go to ∞. We know that

$$\lim_{n\to\infty} q_n(x) = \frac{e^{ix} - e^{-ix}}{2ix} = \frac{\sin x}{x}.$$

On the other hand let us examine the k^{th} term of the product on the right side, *for a fixed* k. From the power series for $\cos u$ it follows that

$$\lim_{u\to 0} \cos u = 1, \qquad \lim_{u\to 0} \frac{1 - \cos u}{u^2} = \frac{1}{2}.$$

Hence

$$\lim_{n\to\infty} 1 + \cos\frac{2k\pi}{n} = 2, \qquad \lim_{n\to\infty} n^2 \left(1 - \cos\frac{2k\pi}{n}\right) = \frac{2k^2}{\pi^2}.$$

Thus

$$\lim_{n\to\infty} 1 - \frac{x^2}{n^2} \frac{1 + \cos\frac{2k\pi}{n}}{1 - \cos\frac{2k\pi}{n}} = 1 - \frac{x^2}{k^2\pi^2}.$$

From this Euler deduced finally his famous infinite product expression:

$$\frac{\sin x}{x} = \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2\pi^2}\right).$$

The passage to the limit above was not scrutinized for nearly a hundred years. It was only after *Weierstrass* laid the foundations of modern analysis that it was realized that one needed the notion of *uniform convergence* to justify the Eulerian passage to the limit. The point is that the number of terms in the finite product is also going to infinity, and so one cannot apply the rule that the limit of a product is the product of limits. However. with the criterion supplied by Weierstrass for uniform convergence, it is a very easy matter to show that the limit is proper. There are many modern proofs for the infinite product for $\sin x/x$, but none can match Euler's for simplicity and directness.

Lagrange. The other great mathematician in the Eulerian era was *Joseph Louis Lagrange*. (1736–1813). He was the successor to Euler as the Director of the Prussian Academy of Sciences. His work was monumental and wide-ranging. He created the modern theory of Mechanics based on the notion of the *Lagrangian*, which is the basis of all theoretical physics now. Its essence consists of the *Euler-Lagrange equations* which are derived by the *Calculus of Variations*, a subject he developed, at first independently, and then by building on the work of Euler. He made fundamental contributions to number theory and astronomy, not to mention calculus, and carried these subjects far beyond what Euler did.

Taylor series with remainder. In calculus Lagrange clarified exactly how the Taylor series of a function is related to the function itself. If f is a function which is differentiable any number of times (we say f is *smooth*), then its Taylor or Maclaurin series is

$$f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^(n)(0)}{n!}x^n + \ldots.$$

Instead of 0 we can speak of the Taylor series at a point a. This is

$$f(a) + f'(0)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^(n)(a)}{n!}(x-a)^n + \ldots.$$

This topic was developed by *Brook Taylor* (1685–1731) as well as *Colin Maclaurin* (1698–1746), both contributors to the calculus as developed by Newton. But there was no discussion of how the series was related to the function.

To motivate the Taylor series expansion, let us view $f(x+h)$ as obtained from $f(x)$ by *translating the function f by h*. Let us write T_h for this operator that takes f to its translate by h. Then

$$T_h T_k = T_{h+k}$$

which suggests that

$$T_h = e^D \qquad D = \left(\frac{d}{dh}\right)_{h=0} T_h.$$

Since

$$D = \frac{d}{dx}$$

we can write formally

$$T_h = I + hD + \frac{h^2}{2!}D^2 + \ldots$$

giving

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \ldots$$

which is the Taylor series expansion. These ideas are due to Euler.

Let us write the series as

$$f(a+h) = f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \ldots + \frac{f^(n)(a)}{n!}h^n + R_n(h)$$

where $R_n(h)$, *the remainder after n terms*, is defined by this equation. There are *two* ways to understand this remainder

$$R_n(h) = f(a+h) - \left(f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \ldots + \frac{f^(n)(a)}{n!}h^n \right) \tag{4.}$$

First we *fix* n and let $h \to 0$. Then we can show that

$$\lim_{n \to \infty} \frac{R_n(h)}{h^n} = 0 \tag{5.}$$

This is often expressed by saying that $R_n(h)$ is of *smaller order than* n, or

$$R_n(h) = o(h^n).$$

For the second we *fix* h and let $n \to \infty$. No one before Lagrange had considered this. The issue is whether $R_n(h) \to 0$. That $R_n(h)$ may not go to 0 as $n \to \infty$ was discovered by Cauchy when he pointed out that for the function

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases} \tag{6}$$

all derivatives at 0 are 0 so that its Taylor series is identically zero, although the function is not zero for $x \neq 0$!! Lagrange proved the first precise result about $R_n(h)$ and used it to clarify when the Taylor series of a function converges to the function.

Theorem (Lagrange). *For a function with $n+1$ continuous derivatives in an interval containing $[a, a+h]$,*

$$R_n(h) = \frac{f^{(n+1)}(u)}{(n+1)!}h^{n+1} \tag{7}$$

where u lies between a and $a+h$.

Lagrange described those functions for which this remainder $\to 0$ as $n \to \infty$ for each a in the interval where th function is defined, as *analytic.* A sufficient condition for f to be analytic in some *open interval* I is that for each closed subinterval $J \subset I$ there exists a constant $M > 0$ such that

$$\sup_{x \in J} |f^{(n)}(x)| \leq M^n n! \tag{8}$$

for all n. For such a function its Taylor series converges to it. One can prove by this method that all the functions that one encounters normally are analytic. This condition is also necessary.

The proof of the *Lagrange's form of remainder* is as follows. First we need a lemma, usually known as *Rolle's theorem.*

Lemma. *If g is a function which is continuous on $[a,b]$ and differentiable in (a,b), and if $g(a) = g(b) = 0$, then ther is c, with $a < c < b$ such that $g'(c) = 0$.*

Proof. If g is constant, $g' \equiv 0$ and there is nothing to prove. Otherwise g reaches a maximum or a minimum inside (a,b) , say at c. Then $g'(c) = 0$.

Proof of Lagrange's theorem. We write $b = a + h$ and consider the function

$$F(x) = f(b) - f(x) - (b-x)f'(x) - \frac{(b-x)^2}{2!}f''(x) - \ldots - \frac{(b-x)^n}{n!}f^{(n)}(x).$$

Then a simple calculation gives

$$F'(x) = -\frac{(b-x)^n}{n!}f^{(n+1)}(x).$$

Write

$$g(x) = F(x) - \frac{(b-x)^{n+1}}{(b-a)^{n+1}}F(a).$$

We have

$$g(a) = g(b) = 0$$

so that by the Lemma, there is u between a and $a+h$ such that $g'(u) = 0$. But

$$g'(x) = (n+1)\frac{(b-x)^n}{(b-a)^{n+1}}\left(F(a) - \frac{(b-a)^{n+1}}{(n+1)!}f^{(n+1)}(x)\right).$$

Hence, as $F(a) = R_n(h)$, we get

$$R_n(h) = \frac{h^{n+1}}{(n+1)!}f^{(n+1)}(u) \qquad u \in (a, a+h).$$

L'hospital's theorem. This well-known recipe for calculating limits of the form

$$\lim_{x \to a} \frac{f(x)}{g(x)}$$

when both f and g go to 0 as $x \to a$ as

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}$$

arises as follows, in an even more general form. Suppose that

$$f^{(r)}(a) = g^{(r)}(a) = 0, 0 \leq r \leq n-1 \quad g^{(n)}(a) \neq 0.$$

Then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f^{(n)}(a)}{g^{(n)}(a)}.$$

This follows by Taylor series. We have, as $h \to 0$,

$$\frac{f(a+h)}{g(a+h)} = \frac{f^{(n)}(u)}{g^{(n)}(v)}$$

where u, v are in $(a, a+h)$. As $h \to 0$, both u and v tend to a and the result follows.

Integral form of remainder. If we want to use the Lagrange's form of remainder to prove some familiar series we encounter a surprise. Indeed, consider the function $f(x) = (1-x)^{-1}$. We have

$$f^{(n)}(x) = \frac{n!}{(1-x)^n}$$

Hence the Taylor series at $x = 0$ is

$$1 + x + x^2 + x^3 + \dots$$

and the remainder is

$$R_n(x) = \frac{x^{n+1}}{(1+\theta x)^{n+1}}$$

for some θ, with $0 < \theta < 1$. If $0 < x < 1$ we have

$$|R_n(x)| \le x^{n+1} \to 0 \qquad (n \to \infty).$$

But if $-1 < x < 0$ we only have, writing $x = -y, 0 < y < 1$,

$$|R_n(x)| \le \frac{y^{n+1}}{(1-\theta y)^{n+1}}$$

and we cannot assert that the estimate goes to 0, as we cannot be sure that $y/1 - \theta y < 1$. The situation is the same for other functions like $\log(1+x), (1+x)^\alpha$ etc.

We shall now describe a form of the remainder which is exact and which will allow us to prove that the Taylor series for the elementary functions converge to the functions in the natural ranges of values of the independent variable.

Let f be a smooth function. In text books (Hardy's for example) a lot is made of exact conditions for the theorem with remainder, like how many derivatives etc. I shall simply assume that f has $n+1$ continuous derivatives on a closed interval $[-a, a]$ containing 0 in its interior. Then for x in the interval,

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f'''(0) + \dots + \frac{x^n}{n!}f^{(n)}(0) + R_n(x)$$

where $R_n(x)$ is defined by this equation. Then

$$R_0(x) = f(x) - f(0) = \int_0^x f'(t)dt.$$

Replacing $f'(t)$ by $f'(0) + [f'(t) - f'(0)]$ we get

$$R_0(x) = xf'(0) + \int_0^x dx \int_0^t f''(u)du$$

so that

$$R_2(x) = \int_0^x \left(\int_0^t f''(u)du \right) dt.$$

We invert the order of integration to get

$$R_2(x) = \int_0^x f''(u)du \int_u^x dt = \int_0^x (x-u)f''(u)du.$$

Replacing $f''(u)$ by $f''(0) + [f''(u) - f''(0)]$ we get

$$R_3(x) = \int_0^x (x-u)\left(\int_0^u f^{(3)}(v)dv \right) du = \int_0^x (x-t)\left(\int_0^t f^{(3)}(u)du \right) dt.$$

Once again we invert the order of integration to get

$$R_2(x) = \int_0^x \frac{(x-t)^2}{2} f^{(3)}(t)dt.$$

The pattern is clear and we have in general the following theorem which we shall prove by induction.

Theorem. *We have for all n*

$$R_n(x) = \int_0^x \frac{(x-u)^n}{n!} f^{(n+1)}(u)du = \frac{x^{n+1}}{n!} \int_0^1 (1-t)^n f^{(n+1)}(tx)dt.$$

Proof. Assume this for $n-1$. Then writing, as before, $f^{(n)}(t) = f^{(n)}(0) + [f^{(n)}(t) - f^{(n)}(0)]$ we get

$$R_n(x) = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} \left(\int_0^t f^{(n+1)}(u)du \right) dt.$$

We invert the order of integration as before to get

$$R_n(x) = \int_0^x \frac{(x-u)^n}{n!} f^{(n+1)}(u)du.$$

By writing $u = tx$ we get the second formula.

By changing x to $a+h$ we get

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!} f''(a) + \ldots + \frac{h^n}{n!} f^{(n)}(0) + R_n(h)$$

where

$$R_n(h) = \int_a^{a+h} \frac{(a+h-u)^n}{n!} f^{(n+1)}(u)du$$

which reduces to

$$R_n(h) = \int_0^h \frac{(h-v)^n}{n!} f^{(n+1)}(a+v)dv$$

so that

$$R_n(h) = \frac{h^{(n+1)}}{n!} \int_0^1 (1-t)^n f^{(n+1)}(a+th)dt.$$

As an illustration let us consider the case

$$f(x) = (1+x)^a \qquad (|x| < 1).$$

Then

$$f^{(n)}(x) = a(a-1)(a-2)\dots(a-n+1)(1+x)^{a-n}$$

so that

$$\left({\tfrac{f^{(n)}(0)}{n!} = a \atop n} \right)$$

and the Taylor series is the binomial series. To show that it converges to $(1+x)^a$ we use the integral form of the remainder derived above. We have

$$R_n(x) = a(a-1)\dots(a-n)\frac{x^{n+1}}{n!} \int_0^1 (1-t)^n (1+tx)^{a-n-1} dx.$$

Let k be a positive integer such that $|a| \leq k$. Then for $n > k$,

$$\left| \frac{a(a-1)\dots(a-n)}{n!} \right| \leq \frac{k(k+1)\dots(k+n)}{n!} \leq \frac{(n+1)\dots(n+k)}{(k-1)!} \leq Cn^k$$

where C is a constant independent of n. Further

$$|(1+tx)^a| \leq \begin{cases} 2^a & \text{if } a > 0 \\ (1-|x|)^{-|a|} & \text{if } a < 0 \end{cases}$$

so that for fixed x and a constant $C > 0$

$$|R_n(x)| \leq Cn^k |x|^{n+1} \int_0^1 (1-t)^n |(1+tx)|^{-(n+1)}| dt.$$

We consider two cases. First $x \geq 0$. In this case $(1+tx) \geq 1$ and so

$$|R_n(x)| \leq Cn^k |x|^{n+1} \to 0$$

as $n \to \infty$. Suppose now $-1 < x < 0$ so that $x = -y$ where $0 < y < 1$. Then

$$(1-t)^n (1-ty)^{-(n+1)} \leq \left(\frac{1-t}{1-ty} \right)^n \frac{1}{1-ty} \leq \left(\frac{1-t}{1-ty} \right)^n \frac{1}{1-y} \leq \frac{1}{1-y}$$

since $1 - ty \geq 1 - y$. Hence

$$|R_n(x)| \leq Cn^k y^{(n+1)} (1-y)^{-1} \to 0$$

as $n \to \infty$.

The argument is similar for $f(x) = \log(1+x)$ in the range $|x| < 1$.

Homework # 8 : Due March 4, 2011

1. Ex. 5, page 275.

2. Ex, 7, page 277.

3. Ex. 8, page 277.

4. Let $n = 2p + 1$. Show that

$$T^n - 1 = (T-1)\prod_{k=1}^{p}\left(T^2 - 2T\cos\frac{2k\pi}{n} + 1\right).$$

   Writing $T = X/Y$ and clearing fractions show that

$$X^n - Y^n = (X-Y)\prod_{k=1}^{p}\left(X^2 - 2XY\cos\frac{2k\pi}{n} + Y^2\right).$$

   Hence deduce that

$$\frac{(1+\frac{ix}{n})^n - (1-\frac{ix}{n})^n}{2ix} = \prod_{k=1}^{p}\left(1 - \frac{x^2}{n^2}\frac{1+\cos\frac{2k\pi}{n}}{1-\cos\frac{2k\pi}{n}}\right).$$

   Letting $n \to \infty$ deduce formally that

$$\frac{\sin x}{x} = \prod_{k=1}^{\infty}\left(1 - \frac{x^2}{k^2\pi^2}\right).$$

5. Show that $e^x, \sin x, \cos x$ are analytic everywhere, $\log x$ analytic whenever $x > 0$, and $(1+x)^\alpha$ is analytic for $x \neq -1$.

6. Prove that

$$\frac{f(a+h) - 2f(a) + f(a-h)}{h^2} \to f''(a).$$

7. Prove that

$$\frac{3\sin 2\theta}{2(2+\cos\theta)} = \theta + \frac{4}{45}\theta^5 + o(\theta^5)$$

   where $o(\theta^5)$ denotes a quantity $f(\theta)$ such that $f(\theta)/\theta^5 \to 0$. (*Hint*: Write g for this function so that $4g = 3\sin 2\theta - 2g\cos\theta$. Differentiate repeatedly to show that $g'(0) = 1, g^{(r)}(0) = 0$ for $r = 2, 3, 4$ and $4/45$ for $r = 5$.)

8. Show that

$$\frac{1 - 4\sin^2\frac{1}{6}\pi x}{1 - x^2} \to \frac{1}{6}\pi\sqrt{3} \quad \text{when } x \to 1.$$

9. (From Professor Grossman's notes of winter 2007) (a) Prove that

$$\frac{22}{7} - \pi = \int_0^1 \frac{x^4(1-x)^4}{1+x^2} dx.$$

Hence deduce that $\frac{2}{7} > \pi$. (*Hint*: By long division show that $x^4(1-x)^4 = (1+x^2)(x^6 - 4x^5 + 5x^4 - 4x^2 + 4) - 4$.)

(b) Using $2 < 1 + x^2 < 2$ for $0 < x < 1$ in the above integral show that

$$\frac{1}{1260} < \frac{22}{7} - \pi < \frac{1}{630}.$$

(c) Show that the Archimedes bounds are equivalent to

$$0 < \frac{22}{7} - \pi < \frac{1}{497}.$$

Deduce that the bounds in 9(b) are better.

Note: There are integrals more general than the above giving better approximations to π, such as

$$\int_0^1 \frac{x^{4n}(1-x)^{4n}}{1+x^2} dx.$$

See the Wikipedia article on $22/7 - \pi$. See also *Irresistible Integrals* by *George Boros and Victor Moll*, Cambridge University Press, 2004, 6.7., for an expression for $x^{4n}(1-x)^{4n}(1+x^2)^{-1}$.

10. (From Professor Grossman's notes of winter 2007) (a) Use the formula $\sin x = 2\cos\frac{x}{2}\sin\frac{x}{2}$ repeatedly to prove that

$$\frac{\sin x}{x} = \cos\frac{x}{2}\cos\frac{x}{2^2}\ldots = \lim_{N\to\infty}\prod_{n=1}^{N}\cos\frac{x}{2^n} = \prod_{n=1}^{\infty}\cos\frac{x}{2^n}.$$

(b) Taking $x = \frac{\pi}{2}$ obtain Viete's product formula for π:

$$\frac{2}{\pi} = \frac{\sqrt{2}}{2}\frac{\sqrt{2+\sqrt{2}}}{2}\frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2}\cdots.$$

For the ambitious student

1. Examine the difficulty in justifying the limit process in obtaining the series for e^z.

2. Examine the difficulty in passing from the finite factorization to its limit in deriving the infinite product for $\sin x$ and $\cos x$.

3. Using logarithmic differentiation derive from the infinite product for $\sin x$ the famous formula of Euler:

$$\pi \cot \pi x = \frac{1}{x} + \sum_{k=1}^{\infty} \left(\frac{1}{x+k} + \frac{1}{x-k} \right)$$

which can be formally (and loosely) written as

$$\pi \cot \pi x = \sum_{k \in \mathbf{Z}} \frac{1}{x+k}.$$

Why is the previous expression to be preferred?

4. Show using Taylor's formula with Lagrange's remainder that

$$\cos y = 1 - \frac{y^2}{2} \cos(\theta y) \quad (0 < \theta < 1)$$

and hence deuce that the infinite product in problem 10 above for $\frac{\sin x}{x}$ converges for $x = \frac{\pi}{2}$.

5. (See Boros and Moll, loc. cit) Expanding $(1+t^2)^{-1}$ a a power series in t^4 show that if L_n is the sum of the first n terms of the Leibniz series for $\frac{\pi}{4}$,

$$\frac{1}{2(2n+1)} < \left| \frac{\pi}{4} - L_n \right| < \frac{1}{(2n+1)}.$$

Weeks 9–10 : Additional remarks

Cauchy. *Augustin-Louis Cauchy* (1789–1857) was one of the first mathematicians to insist on rigor in analysis. He was universal like Euler, and made fundamental discoveries in differential equations, and complex analysis. The whole of complex analysis is based on the *Cauchy formula* which says that the integral

$$\oint_C f(z)dz = 0$$

if f is an analytic function in a simply connected domain and C is a closed curve in that domain. The *Cauchy-Riemann equations* characterize analytic functions, the *Cauchy problem* is the initial value problem in partial differential equations, and so on. He wrote the first modern books on calculus.

Integration. The foundations for integration were unsatisfactory for mathematicians who came after Euler and Lagrange. They wanted to integrate functions with discontinuities as they were increasingly occurring in applications. For instance Fourier encountered such functions in his theory of heat conduction. Eventually Riemann developed a theory of integration that was, and is, adequate for most applications. But his theory was not sufficient for many *theoretical* purposes and a more complete notion of the integral was developed by Lebesgue at the beginning of the twentieth century. Lebesgue's integral is the basic one for all applications, theoretical as well as practical. In the modern theories of probability and quantum mechanics where integration has to be performed in *infinite dimensional spaces*, a generalization of the Lebesgue integral, in an abstract form liberated from its euclidean confinement, was introduced by Wiener. It is called the *Wiener integral* and is a supremely flexible tool.

Riemann. *Georg Friedrich Bernhard Riemann* (1826–1866) is the successor to Gauss as the greatest mathematician of his era. Certainly, along with Euler and Gauss, he inaugurated the modern era. A large part of twentieth century mathematics consists in understanding Riemann's work and pushing it along the lines that Riemann had foreseen, and then going beyond. Riemann created what is now called *Riemannian geometry.* It unified euclidean and non-euclidean geometry under one umbrella. It is the foundation for Einstein's *theory of general relativity.* He created the theory of *Riemann surfaces*. In his famous inaugural lecture at Göttingen in 1854 he suggested that understanding space in its very minutest parts may require models in which the usual coordinates may not be enough. This has been justified by the physicists' discovery of super geometry as a tool for studying spacetime in regions of energy in the TeV range.

The Riemann integral. Riemann's idea for the integral is very simple. Let f be a *bounded* function defined on the interval $[a, b]$. Then we divide this interval by points

$$x_0 = a < x_1 < \ldots < x_n = b, \qquad \Delta_i = [x_i, x_{i+1}]$$

and introduce two sums

$$I^- = \sum_{0 \le i \le n-1} m_i(x_{i+1} - x_i), \qquad I^+ = \sum_{0 \le i \le n-1} M_i(x_{i+1} - x_i)$$

where

$$m_i = \inf_{x \in \Delta_i} f(x), \quad M_i = \sup_{x \in \Delta_i} f(x).$$

Notice that $I^{\pm}$ depend also on the division points, a dependence that has been suppressed in our notation. I^+, I^- are called *upper and lower Riemann sums*. One then shows that $I^- \le I^+$ and, even more generally, if $I_1^{\pm}$ and $I_2^{\pm}$ are the upper and lower sums associated to *any two* divisions of $[a, b]$, we have

$$I_1^- \le I_2^+.$$

If

$$A \le f(x) \le B \qquad (a \le x \le b)$$

then

$$A \le I_1^- \le I_2^+ \le B \tag{1.}$$

What Riemann wanted was to take more and more division points such that the *mesh* of the divisons, namely the maximum length $|\Delta_i|$, becomes very small. This is the essence of the *method of exhaustion* that goes back to Archimedes, as we have seen.

The bounds (1) show that there is a *least upper bound J^- for all lower sums*, and a *greatest lower bound J^+ for all upper sums.* We have

$$A \le J^- \le J^+ \le B.$$

Riemann defines the function f to be *integrable* if

$$J^- = J^+ \tag{2.}$$

The common value is then written

$$\int_a^b f(x)dx \tag{3}$$

and is called the *integral of f over* $[a, b]$. In this case, if we take a *sequence of divisions* such that the mesh of the divisions goes to 0, then

$$\lim \sum_{0 \le i \le n-1} f(\xi_i)(x_{i+1} - x_i) = \int_a^b f(x)dx \tag{4}$$

where ξ is any point of Δ_i. Finally, if f is continuous or even has a finite number of discontinuities, then it is integrable, and if

$$F(t) = \int_a^t f(x)dx$$

then F is differentiable and

$$\frac{dF}{dt}(x) = f(x)$$

whenever f is continuous at x. In particular, given $f(x)$, if G is any function whose derivative is f, then

$$\int_a^b f(x)dx = [G(x)]_a^b = G(b) - G(a) \tag{5.}$$

These results of Riemann unify *all classical results of integration and its relation to differentiation.*

Higher dimensions. The same method applies in two dimensions, and in fact to all dimensions. One has to replace the Δ_i by squares, cubes, etc. The domain of integration can be a square, cube, etc or even a region bounded by smooth curves, surfaces, or finite unions of these. Riemann showed that even functions with *countably* many discontinuities can be integrable, by constructing some beautiful examples. For a convex domain D in the plane with a smooth simple closed curve C as a boundary

$$\int\int_D f(x,y)dxdy = \int dx \left(\int_{a(x)}^{b(x)} f(x,y)dy\right) \tag{6}$$

where $a(x)$ and $b(x)$ are the points where the vertical line with abscissa x meets C. For example, for D the unit disk $\{(x,y) : |x^2+y^2 \leq 1\}$,

$$\int\int_{x^2+y^2\leq 1} f(x,y)dxdy = \int_{-1}^{+1} dx \left(\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}} f(x,y)dy\right).$$

Lebesgue. *Henri Lebesgue* (1875–1941) created the *Lebesgue integral* when the Riemann integral proved inadequate, especially in the theory of Fourier series. For a bounded function f which is Riemann integrable on $[0,2\pi]$ one can define its Fourier series as the series

$$\sum_{n\in\mathbf{Z}} a_n e^{inx} \qquad a_n = \int_0^{2\pi} f(x)e^{-inx}dx.$$

The a_n are the *Fourier coefficients* of f. One can then show that

$$\int_0^{2\pi} |f(x)|^2 dx = \sum_{n\in\mathbf{Z}} |a_n|^2.$$

The question naturally arose that if we are given a sequence of complex numbers a_n with $\sum_n |a_n|^2 < \infty$, whether there is a function whose Fourier coefficients are the a_n. It turned out that this is impossible in full generality *unless one enlarges the class of integrable functions far beyond Riemann's definition.* Lebesgue's definition does precisely this and answers the above question *affirmatively.* Give the a_n there is a function f such that $|f|^2$ is Lebesgue integrable whose Fourier coefficients are the a_n, and that this function is unique in a certain essential sense.

In Lebesgue's theory, one proceeds exactly as in Riemann's but introduce a wider class of subdivisions than the intervals:

$$[a,b] = \bigsqcup_{0\leq i\leq n-1} S_i \quad \bigsqcup = \text{ disjoint union} \tag{7.}$$

However, before introducing the corresponding lower and upper sums, one has to extend the notion of length from the intervals to a much wider class of sets, the so-called *measurable sets*. The length function is called now the *measure* and it has remarkable properties on the class of measurable sets, especially its *countable additivity*: if (S_i) is a sequence of *mutually disjoint* measurable sets and $S = \cup_i S_i$, then

$$\ell(S) = \sum_i \ell(S_i) \qquad \ell(A) = \text{ measure of } A.$$

A function f is called *measurable* if the sets obtained from it such as $\{x|f(x) < a\}$ are all measurable. Then if we take the S_i in (7) to be measurable sets, we can define the sums

$$I^- = \sum_i m_i \ell(S_i), \qquad I^+ = \sum_i M_i \ell(S_i) \qquad m_i = \inf_{x\in S_i} f(x), \quad M_i = \sup_{x\in S_i} f(x).$$

The theory of these sums is exactly the same as in Riemann's theory; but now one has the result that *all bounded measurable functions are integrable.* The class of measurable functions is enormously wider than the continuous functions. It contains all continuous functions and is closed under pointwise sequential limits, and so, passing repeatedly to limits, contains functions whose discontinuities can be very complex. Riemann's theory is encompassed by the result that all bounded Riemann integrable functions are Lebesgue integrable and the two integrals are equal; moreover a bounded function is Riemann integrable if and only if its set of discontinuities form a set of measure 0.

The countable additivity of the Lebesgue measure function leads to remarkable properties for the Lebesgue integral. The basic limit theorem is the following: if (f_n) arer bounded measurable and $|f_n| \leq C$ for all n where C is a constant (uniformly bounded sequence), and if $f_n(x) \to f(x)$ for all x (or almost all x, meaning that the limit holds except for a set of points of measure 0), then $|f| \leq C$ (obvious) except for a set of measure 0 and

$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx.$$

the integral is unaltered if the value of the function is changed on a set of measure 0. The theory can then be extended very simply to include unbounded functions. The theory extends to all dimensions where length is replaced by area, volume, etc. The measure so obtained is translation and rotation invariant and so expresses a basic property of the euclidean spaces. The Lebesgue integral is the basis for all modern analysis. However, for most applications the Riemann integral is perfectly adequate.

Wiener. *Norbert Wiener* (1894–1964) discovered a profound generalization of the Lebesgue measure and the Lebesgue integral when the space on which they are defined is *infinite dimensional.* An example of such a space is the *space* $C[0,1]$ *of all continuous functions on* $[0,1]$. In such a space one has first to introduce the elementary sets like the intervals, squares, and cubes, in euclidean space, and then introduce their measures. Wiener did this and his measure and integration theory has been remarkable for applications, such as modeling *Brownian Motion* and *Quantum Mechanics.* The *Feynman integral*, familiar to the physicists, can be defined via the Wiener integral. The Feynman integral is the basis of all quantum field theories.

Infinitesimals. To return to the foundations of calculus and the Leibniz infinitesimals dx, dy, etc. Euler worked very successfully with infinitesimals but there remained an aura of mystery surrounding them. In the fall of 1960, *Abraham Robinson* (1918–1974) broke through this veil of mystery and discovered a beautiful and fully rigorous way of dealing with them, creating thereby a new generalization of analysis, nowadays called *The Non-Standard Analysis.* He showed that the entire theory of Leibniz can now be done in a completely rigorous manner, and nothing has to be changed in the proofs of Leibniz.

Calculus of variations. There are many problems in classical analysis and geometry which are very easy to state but whose solution depends on the ideas from a new subject, the *Calculus of Variations.* Historically some of these problems were already posed by several people like John and Jacob Bernoulli and solved by many, including Newton, by ad hoc methods. But after a few years Euler recognized that these problems required the creation of a new calculus, and developed it systematically. A few years after Euler's publications in this area appeared, Lagrange found a new approach and communicated it to Euler. Euler recognized the merits of Lagrange's approach, and from then on, adopted it for his own work (and treated Lagrange as his equal). The fundamental equations of the nw theory as the famous *Euler-Lagrange equations.* They dominate all of mechanics, and in modern times, all of field theory, classical, quantum, or supersymmetric.

Some examples. 1. *Find the shortest paths between two points on a surface.* If the surface is the Euclidean plane these are the straight lines joining the points, while for the surface of a sphere these are the great circles joining the points. For other surfaces they are more complicated, for instance, an ellipsoid. This the problem of *geodesics.* An interesting case is the upper half plane of complex numbers $z = x + iy$ with $y > 0$ and

$$ds^2 = \frac{1}{y^2}(dx^2 + dy^2).$$

2. *Brachistochrone.* Let A, B be two points in a vertical plane with B lower than A. The problem, originally posed by John Bernoulli, is to find the curve from A to B with the property that a small particle rolling along the curve under the influence of gravity takes the least amount of travel time. The solution curve is called the *Brachistochrone.* It is a *cycloid.*

3. *The isoperimetric problem.* Among all closed curves of a fixed length ℓ find the curve that encloses the maximum area. This was solved by Euler who showed that it is the circle.

Problems of maxima and minima in finite number of variables. As a guide to the study of the above problems and other similar to them, let us first review the calculus problems of extrema of functions of s finite number of variables. Consider the US Post Office problem of finding a rectangular box of dimensions ℓ, b, h of maximum volume subject to the constraint $\ell + b + h = C$ on its linear dimension, where $C > 0$ is a constant. Then the volume is

$$V = \ell b(C - \ell - b) \qquad (0 \leq \ell, b \leq C, \ell + b < C).$$

At an extremum, we must have

$$\frac{\partial V}{\partial \ell} = 0, \quad \frac{\partial V}{\partial b} = 0$$

giving the equations

$$Cb - 2b\ell - b^2 = 0, \quad C\ell - \ell^2 - 2b\ell = 0.$$

Subtraction of the first from the second gives

$$(\ell - b)(C - \ell - b) = 0 \Longrightarrow \ell = b.$$

Hence there is a unique extremum

$$\ell = b = h = \frac{C}{3}.$$

The required solution is a cube of side $C/3$. To show that this is a local maximum we must show that the matrix of second derivatives at the extremum is negative definite, or at least ≤ 0. The matrix is

$$H = \begin{pmatrix} V_{\ell\ell} & V_{\ell b} \\ V_{b\ell} & V_{bb} \end{pmatrix} = \begin{pmatrix} -2b & C - 2\ell - 2b \\ C - 2\ell - 2b & -2b \end{pmatrix} = \begin{pmatrix} \frac{-2C}{3} & \frac{-C}{3} \\ \frac{-C}{3} & \frac{-3C}{3} \end{pmatrix}.$$

This is negative definite clearly. This does not prove that this is a *global maximum.* For that we must give an additional argument.

As another very interesting example we consider a torus hanging above the xy-plane and the function defined on the torus whose value at a point of the torus is its height. There are 4 extrema, the top and bottom and the inner ones (draw a figure). The top is a max, the bottom is a min, and the two inside ones are *saddle points*; in the neighborhoods of these two, the torus looks like a saddle, and there are directions along which we have a min and other directions along which we have a max.

In the general case we first determine the *critical points* of a function f which are given by the equations

$$\frac{\partial f}{\partial x_i} = 0 \quad (1 \leq i \leq n).$$

Once these are determined, we must look at each one of them and determine the index of the matrix of second derivatives (the *Hessian*).

In the case where the solution is a function rather than a vector in a finite dimensional space, the first step is to understand what is meant by a critical point. Since a cyrve requires an *infinite number* of parameters to determine it, the critical point would be determined by an infinite number of wquations. Roughly speaking, we must have one equation for each direction in the space of curves. Given a curve $y(x)$ where the end points are fixed, namely,

$$y(a) = A, y(b) = B,$$

we define a neighboring family of curves to be the curves y_ε,

$$y_\varepsilon(x) = y(x) + \varepsilon h(x) \qquad (h(a) = h(b) = 0).$$

Note that the fact that h vanishes at a and b means that the y_ε also pass through (a, A) and (b, B). We think of $h(x)$ as a *direction* in the space of curves. As ε varies the curves give us a family containing y (at $\varepsilon = 0$).

To encompass the examples we discussed above we assume that we have a function L of three variables. It is called the *Lagrangian.* Then for any x we have the function $L(x, y(x), y'(x))$ and hence

$$J[y] = \int_a^b L(x, y(x), y'(x))dx.$$

The quantity J depends on the curve y and so is a function on the space of curves. It is called a *functional.* For brevity we often write

$$J[y] = \int_a^b L(x, y, y')dx.$$

When we write for instance $\partial L/\partial y'$ we mean the partial with respect to the third variable in the definition of L but with the third variable replaced by y'. In examples we consider we can see what this means. Similar remarks apply to $\partial L/\partial y$. Then an *extremal curve* is defined as the curve $y(x)$ such that

$$\left.\frac{dJ[y + \varepsilon h]}{d\varepsilon}\right|_{\varepsilon=0} = 0 \qquad \text{for all } h \text{ with } h(a) = h(b) = 0.$$

This is natural because this says that *the derivative of the functional J at y in any direction is* 0. This is the essence of the calculus of variations–to reduce concepts and calculations in the space of functions to the corresponding ones in functions of a finite number of variables by working with families of curves. The quantity

$$\delta J[y](h) = \frac{dJ[y+\varepsilon h]}{d\varepsilon}\bigg|_{\varepsilon=0}$$

is called the *first variation of J in the direction h.* It is also a functional, but of h, with y fixed. Note that $\delta J[y](h)$ is *linear* in h.We can similarly define the *second variation* of J by

$$\delta^2 J[y](h) = \frac{d^2 J[y+\varepsilon h]}{d\varepsilon}\bigg|_{\varepsilon=0}.$$

We say that an extremal y is a local minimum (or maximum) if

$$\delta^2 J[y](h) \geq 0(\text{ resp. } \leq 0) \text{ for all } h \text{ with } h(a) = h(b) = 0.$$

It is easy to show that

$$\delta^2 J[y](th] = t^2 \delta^2 J[y](h) \qquad (t \in \mathbf{R})$$

so that the second variation is a *quadratic functional* in h. So the sign conditions above are consistent.

The Euler-Lagrange equations. Let us now calculate the first variation. Differentiating under the integral sign we have

$$\delta J[y](h) = \int_a^b \left(h\frac{\partial L}{\partial y} + h'\frac{\partial L}{\partial y'} \right) dx.$$

Since $h(a) = h(b) = 0$ we can integrate by parts the second term of the integral and obtain

$$\delta J[y](h) = \int_a^b \left(\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} \right) h dx.$$

For y to be an extremal this quantity has to vanish for all h with $h(a) = h(b) = 0$. We now have the easily proved classical lemma.

Lemma. *If f is a continuous function on $[a, b]$ such that*

$$\int_a^b f(x)h(x)dx = 0 \qquad (\text{for all } h \text{ with } h(a) = h(b) = 0)$$

then $f = 0$ on $[a, b]$.

We then have

Theorem. *For y to be an extremal it must satisfy the equations*

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = 0.$$

These are the famous *Euler-Lagrange equations* for an extremal with the Lagrangian L.

The question of when the extremal is a local minimum or maximum is much deeper. Legendre worked on it but despite much progress he was unable to completely solve it. It was solved eventually by Jacobi and Weierstrass.

let us now look at the examples of geodesics in the plane and the Brachistochrone.

Geodesics in the plane. Here the length of a curve joining (a, A) and (b, B) is

$$\int_a^b \sqrt{1 + y'^2}dx.$$

Hence

$$L = L(y') = \sqrt{1 + y'^2}.$$

The E-L equation is

$$\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = 0.$$

A simple calculation gives

$$\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = \frac{y''}{(1 + y'^2)^{3/2}}$$

so that

$$y'' = 0$$

is the equation of the extremal. The extremal is the unique line joining the two points.

Brachistochrone. We take the two points to be $A = (0, 0)$ and $B = (b, c)$ where $c < 0$. We regard y as a smooth function f with

$$f(0) = 0, \quad f(b) = c \quad f(x) < 0(a < x < b).$$

Write g for the constant gravitational acceleration. Then the transit time is

$$J[y] = \int_a^b \frac{ds}{v}$$

where v is the velocity and $ds = \sqrt{1 + y'^2}\ dx$. By the conservation of energy and the fact that the total energy is the kinetic + potential energy, we get, with mass =1,

$$\frac{1}{2}v^2 + gy = 0$$

or

$$v = \sqrt{g(-y)}.$$

Hence

$$J[y] = \int_a^b \frac{\sqrt{1 + y'^2}}{\sqrt{-2gy}}dx.$$

Writing $z = -y$ we have

$$J[z] = \int_1^b l(z, z')dx$$

where

$$L(z, z') = \frac{\sqrt{1 + z'^2}}{\sqrt{2gz}}$$

where z is now positive between a and b. The key is now a lemma which is a special case of Noether's theorem.

Lemma. *If L is independent of x, then the Euler-Lagrange equations have a first integral $L - y'L_{y'}$ so that*

$$L - y'L_{y'} = C$$

where C is a constant.

Proof. Indeed, the Euler-Lagrange equations say

$$L_y - (L_{y'})' = 0.$$

Thus

$$L_y - L_{y'y}y' - L_{y'y'}y'' = 0.$$

Multiply by y' and we have

$$y'L_y - L_{y'y'}y'^2 - L_{y'y'}y'y'' = 0.$$

This is just

$$\frac{d}{dx}\left(L - y'L_{y'}\right) = 0$$

giving

$$L - y'L_{y'} = C$$

where C is a constant.

We can now solve the brachistochrone. Here L is independent of x and so

$$L - z'L_{z'} = C.$$

For

$$L = \frac{\sqrt{1 + z'^2}}{\sqrt{2gz}}$$

this yields

$$\sqrt{2gz}\sqrt{1 + z'^2} = \frac{1}{C}$$

or

$$z(1 + z'^2) = \frac{1}{2gC^2} = k^2.$$

Solving for z' we have

$$z' = \frac{\sqrt{k^2 - z}}{\sqrt{z}}$$

which leads to

$$\frac{\sqrt{z}}{\sqrt{k^2 - z}}dz = dx.$$

Make the substitution $z = k^2 \sin^2 \theta$ to get

$$k^2(1 - \cos 2\theta) = dx$$

giving

$$x = \frac{k^2}{2}(2 - \sin 2\theta + A, \quad z = \frac{k^2}{2}(1 - \cos 2\theta).$$

Changing 2θ to t and noting that the start of the curve is $(0, 0)$ and that $y = -z$ we finally get the solution in parametric form:

$$x = a(t - \sin t), \quad y = -a(1 - \cos t) \quad (a = \frac{k^2}{2}).$$

It is the *mirror image of the usual cycloid.*

Discovery of Calculus in Kerala in the fourteenth century by Madhavan and his disciples. The first inkling to the western world of science that some very sophisticated mathematics had been done in Kerala, India, in the 14^{th} century, anticipating by more than two centuries the discovery of the calculus by Newton and Leibniz, was in an account read in 1832 before the Royal Asiatic Society of Great Britain and Ireland, by one Charles M. Whish, an employee of the East India Company. But it was only in the middle of the 20^{th} century that a serious attempt was begun to really understand the scope and limits of this remarkable development in India, predating the discovery of the calculus in Europe and England. For more details see:

1. A preprint by P. P. Diwakaran, entitled *Natural and Artficial in the Language of the Malayalam Text Yuktibhaṣa̱*, 2007. This text is hereafter abbreviated as *YB*.

2. Two papers by C. T. Rajagopal and M. S. Rangachari: (a) *On an untapped source of medieval Keralese mathematics*, Archive for History of exact sciences, **18**(1978), 89–101. (b) *On Medieval Kerala Mathematics*, Archive for History of exact sciences, **35**(1986), 91–99.

Kerala is a state in India, in the deep southwest coastal part of the country. It has an unbroken tradition of respect and support for learning going back to the 10^{th} century. It enjoys the highest rate of literacy in the country and its people are extremely sophisticated, both intellectually and politically, in comparison with any part of the world. It was in this milieu, fully a century before the Portugese landing on the west coast of India, that a remarkable school of mathematics and astronomy was created around the period of years 1350–1425, and flourished for almost two centuries. The first member of this school, and apparently its deepest and most original, was one *Madhavan* of the village of *Sangamagramam*.. His disciples and their students formed the core of this school for the next two centuries. In order of succession they are

$$\textit{Madhavan} \Rightarrow \textit{Paramesvaran} \Rightarrow \begin{cases} Nilakanthan \\ \textit{Damodaran} \end{cases} \Rightarrow \textit{Jyeshtadevan} \Rightarrow \dots$$

Although the primary purpose of the school was to develop mathematics as applied to astronomy, the development of mathematics went beyond the narrow requirement of applicability to astronomy, and was driven by a profound and genuine intellectual curiosity. The discoveries of Madhavan and his school were eventually codified in what is in modern terms a set of lecture notes by *Jyeshtadevan* (1500–1610), called *YB*, or in full, *Yuktibhaṣa̱*. It is written in *malayalam*, the native language in Kerala. It is estimated that this was composed during the years 1550–1560. The first public appearance of a document giving an account (in malayalam) of the discoveries in *YB*, was in 1948, in a book by *Rama Varma Maru Tampuran* and *A. R. Akhilevara Ayyar*. The book is in malayalam and is an annotated transcription of *YB*. It is over 250 pages, half of which is a commentary on the work with diagrams and other additions helpful to the understanding of the text. The original *YB* was in two parts and the above account concerns only the first part dealing with mathematics, the second part devoted to astronomy not being translated. By now full annotated translations in English with commentaries of both parts are available, translated and edited by the late *K. V. Sarma*, in cooperation with *M. D. Srinivas, M. S. Sriram*, and *K. Ramasubramanian*, published by the Hindustan Book Company. Unlike almost all Indian texts in mathematics which are in verse form in *Sanskrit*, which like Latin, was the language of learning and study, *YB* is in malayalam prose, in the common everyday language. It is very discursive and colloquial, a fact that was responsible for its preservation as well as for the obstacles it posed for the further continuation of the original discoveries. Whereas the verse form is very abbreviated and needs the presence of a teacher for elaboration, the more leisurely and prose form of the *YB* makes it possible that its contents can be learnt by anyone in private, without the assistance of a teacher, thus essentially a set of lecture notes in modern parlance. Furthermore, these notes contain detailed proofs of statements, a matter of great help to any prospective reader.

The mathematical content of *YB* consists of a study of the circle and the trigonometric functions that arise from it. Among other things the *YB* achieves the following.

• A detailed approximation to π.

The Leibniz series

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

is derived with a detailed treatment of the error estimates after truncation at an arbitrary term, and improvements for more accurate results by adding correction terms.

• More generally the series for $\arctan x$, $\sin x$ and $\cos x$ are obtained.

These are

$$\begin{aligned} \arctan &= x - \frac{x^2}{3} + \frac{x^2}{5} - \frac{x^7}{7} + \dots \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots . \\ \cos x &= 1 - \frac{x^2}{x!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \end{aligned}$$

Furthermore in the process of deriving these results integrals of the power functions x^k are obtained recursively, by a formula which is essentially the integration by parts for them. It is remarkable that the *general term* of each of these series is described precisely (recall that Newton could only guess the general term of the series for $\sin x$.)

It is in these derivations that *YB* contains the most original material. The method is based on the discovery of tangent lines to the circular arcs and their relation to derivatives, and the summing of the infinitesimal difference quotients to get the integral. *There is no doubt that the YB contains the discovery of the differential and integral calculus together with the fundamental theorem of calculus, at least as far as the circular functions are concerned.*

There are other remarkable aspects to these results. The question is raised as to why one seeks approximate formulae for π instead of an exact expression. It is pointed out that the partial sums of the series for $\pi/4$ never coincide with $\pi/4$ and there is an implicitly recognized problem whether π is not rational at all.

What is not in *YB* is a recognition that the ideas discussed are actually extremely general and apply to more or less general functions. It is not clear why the discoverers of these results stopped short. One reason may be that the language they used was not sophisticated enough for generalization. Another may be that there was no interest in studying other functions as they did not arise in their astronomical calculations.

In any case there is no question that the first steps for calculus as we know it today were taken in Kerala by Madhavan in the second half of the fourteenth century, and his insights and discoveries were written down in a form that is perfectly understandable for any one today.

Homework # 9 : Due March 11, 2011

1. The area $a(D)$ of any domain D is defined as

$$a(D) = \int\int_D f(x,y)dxdy.$$

   Use the formula (6) to show that the area of the domain enclosed by the ellipse$\frac{x^2}{a^2} + \frac{y^2}{b62} = 1$ is πab.

2. Use (6) to verify the Archimedes formula

$$a(D) = \frac{4}{3}a(\Delta APB)$$

   for the area of the domain D parabola $Y^2 = X$ cut off by a line segment AB, P being the vertex of the parabola above AB.

3. Use (6) to show that the area enclosed by the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ is πab.

For the ambitious student

1. Let $f_1, f_2, \ldots, f_n, \ldots$ be a sequence of Riemann integrable functions on $[a,b]$ and let

$$lim_{n\to\infty} f_n(x) = f(x) \quad \textit{uniformly in } x.$$

   Show that f is integrable and

$$\int_a^b f(x)dx = \lim_{n\to\infty} f_n(x))dx.$$

2. For any real number x let (x) be the *fractional part* of x, namely, $(x) = x - [x]$ where $[x]$ is the integral part of x, i.e., the largest integer $\leq x$. Prove that the function (x) is continuous everywhere except at the integers where it has a jump 1. (It looks like a saw tooth). Use this

to deduce that for any integer $n \geq 1$ the function (nx) is continuous everywhere at the points $\pm k/n (k = 0, 1, 2, \ldots)$ where its jump is 1.

3. Let
$$f(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n^2}.$$
prove that f is integrable over $[0, 1]$ although it is discontinuous at all rational points (and only at these).