Lecture 1: Mathematical roots

Similarly, as one has distinguished the **canons of rhetorics**: memory, invention, delivery, style, and arrangement, or combined the **trivium**: grammar, logic and rhetorics, with the **quadrivium**: arithmetic, geometry, music, and astronomy, to obtain the seven **liberal arts and sciences**, one has tried to **organize all mathematical activities**.

Historically, one has distinguished **eight ancient roots of mathematics**. Each of these 8 activities in turn suggest a key area in mathematics:

counting and sorting	**arithmetic**
spacing and distancing	**geometry**
positioning and locating	**topology**
surveying and angulating	**trigonometry**
balancing and weighing	**statics**
moving and hitting	**dynamics**
guessing and judging	**probability**
collecting and ordering	**algorithms**

To morph these 8 roots to the 12 mathematical areas covered in this class, we complemented the ancient roots with calculus, numerics and computer science, merge trigonometry with geometry, separate arithmetic into number theory, algebra and arithmetic and turn statics into analysis.

Lets call this modern adaptation the

12 modern roots of Mathematics:

counting and sorting	**arithmetic**
spacing and distancing	**geometry**
positioning and locating	**topology**
dividing and comparing	**number theory**
balancing and weighing	**analysis**
moving and hitting	**dynamics**
guessing and judging	**probability**
collecting and ordering	**algorithms**
slicing and stacking	**calculus**
operating and memorizing	**computer science**
optimizing and planning	**numerics**
manipulating and solving	**algebra**

While relating **mathematical areas** with **human activities** is useful, it makes sense to select specific topics in each of this area. These 12 topics will be the 12 lectures of this course.

Arithmetic	numbers and number systems
Geometry	invariance, symmetries, measurement, maps
Number theory	Diophantine equations, factorizations
Algebra	algebraic and discrete structures
Calculus	limits, derivatives, integrals
Set Theory	set theory, foundations and formalisms
Probability	combinatorics, measure theory and statistics
Topology	polyhedra, topological spaces, manifolds
Analysis	extrema, estimates, variation, measure
Numerics	numerical schemes, codes, cryptology
Dynamics	differential equations, maps
Algorithms	computer science, artificial intelligence

Like any classification, this chosen division is rather arbitrary and a matter of personal preferences. The **2010 AMS classification** distinguishes 63 areas of mathematics. Many of the just defined main areas are broken off into even finer pieces. Additionally, there are fields which relate with other areas of science, like economics, biology or physics:

00 General
01 History and biography
03 Mathematical logic and foundations
05 Combinatorics
06 Lattices, ordered algebraic structures
08 General algebraic systems
11 Number theory
12 Field theory and polynomials
13 Commutative rings and algebras
14 Algebraic geometry
15 Linear/multi-linear algebra; matrix theory
16 Associative rings and algebras
17 Non-associative rings and algebras
18 Category theory, homological algebra
19 K-theory
20 Group theory and generalizations

22 Topological groups, Lie groups
26 Real functions
28 Measure and integration
30 Functions of a complex variable
31 Potential theory
32 Several complex variables, analytic spaces
33 Special functions
34 Ordinary differential equations
35 Partial differential equations
37 Dynamical systems and ergodic theory
39 Difference and functional equations
40 Sequences, series, summability
41 Approximations and expansions
42 Fourier analysis
43 Abstract harmonic analysis
44 Integral transforms, operational calculus

45 Integral equations
46 Functional analysis
47 Operator theory
49 Calculus of variations, optimization
51 Geometry
52 Convex and discrete geometry
53 Differential geometry
54 General topology
55 Algebraic topology
57 Manifolds and cell complexes
58 Global analysis, analysis on manifolds
60 Probability theory and stochastic processes
62 Statistics
65 Numerical analysis
68 Computer science
70 Mechanics of particles and systems

74 Mechanics of deformable solids
76 Fluid mechanics
78 Optics, electromagnetic theory
80 Classical thermodynamics, heat transfer
81 Quantum theory
82 Statistical mechanics, structure of matter
83 Relativity and gravitational theory
85 Astronomy and astrophysics
86 Geophysics
90 Operations research, math. programming
91 Game theory, Economics Social and Behavioral Sciences
92 Biology and other natural sciences
93 Systems theory and control
94 Information and communication, circuits
97 Mathematics education

What are

fancy developments

in mathematics today? Michael Atiyah identified in the year 2000 the following **six hot spots**:

local	and	global
low	and	high dimension
commutative	and	non-commutative
linear	and	nonlinear
geometry	and	algebra
physics	and	mathematics

Also this choice is of course highly personal. One can easily add 12 other **polarizing** quantities which help to distinguish or parametrize different parts of mathematical areas, especially the ambivalent pairs which produce a captivating gradient:

regularity	and	randomness
integrable	and	non-integrable
invariants	and	perturbations
experimental	and	deductive
polynomial	and	exponential
applied	and	abstract

discrete	and	continuous
existence	and	construction
finite dim	and	infinite dimensional
topological	and	differential geometric
practical	and	theoretical
axiomatic	and	case based

An other possibility to refine the fields of mathematics is to **combine** different of the 12 areas. Examples are **probabilistic number theory**, **algebraic geometry**, **numerical analysis**, **geometric number theory**, **numerical algebra**, **algebraic topology**, **geometric probability**, **algebraic number theory**, **dynamical probability = stochastic processes**. Almost every pair is an actual field. Finally, lets give a short answer to the question: What is Mathematics?

Mathematics is the science of structure.

The goal is to illustrate some of these structures from a historical point of view.

Lecture 2: Arithmetic

The oldest mathematical discipline is **Arithmetic**, the theory of constructing and manipulating numbers. The earliest steps were done by **Babylonian**, **Egyptian**, **Chinese**, **Indian** and **Greek** thinkers. Building up the number system starts with the **natural numbers** $1, 2, 3, 4...$ where one can add and multiply. While addition is natural: when adding 3 sticks to 5 sticks to get 8 sticks, the multiplicative operation $*$ is more subtle: $3 * 4$ can be read that we take 3 copies of 4 and get $4+4+4 = 12$. And $4 * 3$ means we take 4 copies of 3 to get $3+3+3+3 = 12$. The first number counts the number of operations while the second counts objects. To motivate $3 * 4 = 4 * 3$, spacial insight can help: we can arrange the 12 objects in a rectangle. Realizing an addition and multiplication structure on the natural numbers is a **great moment** in mathematics. It leads naturally to more general numbers. There are two major motivations to **to build new numbers**:

1. **invert operations** and still get results. 2. **solve equations**.

For 1. to find an inverse of 3 means finding a number such that $x + 3 = 0$ we need $x = -3$, a negative number, to find a number such that $x3 = 1$, we need $x = 1/3$, a rational number. For 2., in order to solve $x + 3 = 1$ one needs integers, to solve $3x = 4$ one needs fractions, to solve $x^2 = 2$ one needs real numbers, to solve $x^2 = -2$ one needs complex numbers.

Numbers	Operation to complete	Examples of equations to solve
Natural numbers	addition and multiplication	$5 + x = 9$
Positive fractions	addition and division	$5x = 8$
Integers	also subtraction	$5 + x = 3$
Rational numbers	also division	$3x = 5$
Algebraic numbers	taking positive roots	$x^2 = 2$, $2x + x^2 - x^3 = 2$
Real numbers	taking limits	$x = 1 - 1/3 + 1/5 - +...$,$\cos(x) = x$
Complex numbers	take any roots	$x^2 = -2$
Surreal numbers	transfinite limits	$x^2 = \omega$, $1/x = \omega$
Surreal complex	any operation	$x^2 + 1 = -\omega$

The development and history of arithmetic can be summarized as follows: humans started with natural numbers, dealt with positive fractions, reluctantly introduced negative numbers and zero to get integers, struggled to "realize" real numbers, were scared to introduce complex numbers, hardly accepted surreal numbers and most do not even know about surcomplex numbers. Ironically, as simple but impossibly difficult questions in number theory show, the modern point of view is the opposite to Kronecker's **"God made the integers; all else is the work of man"**:

The **surreal complex** numbers are the most **natural** numbers;
The **natural** numbers are the most **complex, surreal** numbers.

Natural numbers. Counting can be realized by sticks, bones, quipu knots, pebbles or wampum knots. The **tally stick** concept is still used when playing card games: where bundles of fives are formed, maybe by crossing 4 "sticks" with a fifth. An old stone age tally stick, the **wolf radius bone** contains 55 notches, with 5 groups of 5. It is probably more than 30'000 years old. An other famous paleolithic tally stick is the **Ishango bone**, the fibula of a baboon. It could be 20'000 - 30'000 years old. Others date it to 9000-6500 BC. Earlier counting could have been done by assembling **pebbles**, tying **knots** in a string, making **scratches** in dirt or bark but no such traces have

survived the thousands of years. The **Roman system** improved the tally stick concept by introducing new symbols for larger numbers like $V = 5, X = 10, L = 40, C = 100, D = 500, M = 1000$. in order to avoid bundling too many single sticks. The system is unfit for computations as simple calculations $VIII + VII = XV$ show. **Clay tablets**, some as early as 2000 BC and others from 600 - 300 BC are known. They feature **Akkadian arithmetic** using the base 60. The hexadecimal system with base 60 is convenient becuase of many factors. It survived: we use 60 minutes per hour. **The Egyptians** used the base 10. The most important source on Egyptian mathematics is the **Rhind Papyrus** of 1650 BC. Hieratic numerals were used to write on papyrus from 2500 BC on. **Egyptian numerals** are hieroglyphics. They were found in carvings on tombs and monuments and are 5000 years old. The modern way to write numbers like 2015 is the **Hindu-Arab system** which diffused to the West only during the late Middle ages. It replaced the more primitive **Roman system**. Greek arithmetic used a primitive number system with no place values: 9 Greek letters for $1, 2, \dots 9$, nine for $10, 20, \dots, 90$ and nine for $100, 200, \dots, 900$.

Integers. **Indian Mathematics** morphed the place-value system into a modern method of writing numbers. Hindu astronomers used words to represent digits, but the numbers would be written in the opposite order. Sometimes after 500, the Hindus changed to a digital notation which included the symbol 0. Negative numbers were introduced around 100 BC in the **Chinese** text "Nine Chapters on the Mathematica art". Also the **Bakshali manuscript**, written around 300 AD subtracts numbers carried out additions with negative numbers, where + was used to indicate a negative sign. In Europe, negative numbers were avoided until the 15'th century.

Fractions: **Babylonians** could handle fractions. The **Egyptians** also used fractions, but wrote every fraction a as a sum of fractions with unit numerator and distinct denominators, like $4/5 = 1/2 + 1/4 + 1/20$ or $5/6 = 1/2 + 1/3$. Maybe because of such cumbersome computation techniques, Egyptian mathematics failed to progress beyond a primitive stage. The modern decimal fractions used nowadays for numerical calculations were adopted only in 1595 in Europe.

Real numbers: The Greeks who noticed first that the diagonal of the square is not a rational number. It produced a crisis. Only much later, it became clear that "most" numbers are not rational. **Georg Cantor** realized first that the cardinality of all real numbers is much larger than the cardinality of the integers: while one can count all rational numbers but not enumerate all real numbers. One consequence is that most real numbers are transcendental: they do not occur as solutions of polynomial equations with integer coefficients. The number π is an example. The concept of real numbers is related to th **concept of limit**. Sums like $1+1/4+1/9+1/16+1/25+...$ approach real numbers which are not rational any more.

Complex numbers: Some polynomials have no real root. To solve $x^2 = -1$ for example, we need new numbers. One idea is to use pairs of numbers (a, b) where $(a, 0) = a$ are the usual numbers and extend addition and multiplication $(a, b)+(c, d) = (a+c, b+d)$ and $(a, b)\cdot(c, d) = (ac-bd, ad+bc)$. With this multiplication, the number $(0, 1)$ has the property that $(0, 1) \cdot (0, 1) = (-1, 0) = -1$. It is more convenient to write $a+ib$ where $i = (0, 1)$ satisfies $i^2 = -1$. One can now use the common rules of addition and multiplication.

Surreal numbers: Similarly as real numbers fill in the gaps between the integers, the surreal numbers fill in the gaps between Cantors ordinal numbers. They are written as $(a, b, c, ...|d, e, f, ...)$ meaning that the "simplest" number is larger than $a, b, c...$ and smaller than $d, e, f, ...$ We have $(|) = 0, (0|) = 1, (1|) = 2$ and $(0|1) = 1/2$ or $(|0) = -1$. Surreals contain already transfinite numbers like $(0, 1, 2, 3...|)$ or infinitesimal numbers like $(0|1/2, 1/3, 1/4, 1/5, ...)$. They were introduced in the 1970'ies by John Conway. The late appearance confirms the pedagogical principle: **late human discovery manifests in increased difficulty to teach it**.

Lecture 3: Geometry

Geometry is the science of **shape, size and symmetry**. While arithmetic dealt with numerical structures, geometry deals with metric structures. Geometry is one of the oldest mathematical disciplines and early geometry has relations with arithmetics: we have seen that that the implementation of a commutative multiplication on the natural numbers is rooted from an interpretation of $n \times m$ as an area of a **shape** that is invariant under rotational **symmetry**. Number systems built upon the natural numbers inherit this. Identities like the **Pythagorean triples** $3^2+4^2 = 5^2$ were interpreted geometrically. The **right angle** is the most "symmetric" angle apart from 0. Symmetry manifests itself in quantities which are **invariant**. Invariants are one the most central aspects of geometry. Felix Klein's **Erlanger program** uses symmetry to classify geometries depending on how large the symmetries of the shapes are. In this lecture, we look at a few results which can all be stated in terms of invariants. In the presentation as well as the worksheet part of this lecture, we will work us through smaller miracles like **special points in triangles** as well as a couple of gems: **Pythagoras**, **Thales**,**Hippocrates**, **Feuerbach**, **Pappus**, **Morley**, **Butterfly** which illustrate the importance of symmetry.

Much of geometry is based on our ability to measure **length**, the **distance** between two points. A modern way to measure distance is to determine how long light needs to get from one point to the other. This **geodesic distance** generalizes to curved spaces like the sphere and is also a practical way to measure distances, for example with lasers. It bypasses the problem to determine first the underlying nature of the space in which we do geometry. Having a distance $d(A, B)$ between any two points A, B, we can look at the next more complicated object, which is a set A, B, C of 3 points, a **triangle**. Given an arbitrary triangle ABC, are there relations between the 3 possible distances $a = d(B, C), b = d(A, C), c = d(A, B)$? If we fix the scale by $c = 1$, then $a + b \geq 1, a + 1 \geq b, b + 1 \geq a$. For any pair of (a, b) in this region, there is a triangle. After an identification, we get an abstract space, which represent all triangles uniquely up to similarity. Mathematicians call this an example of a **moduli space**.

A **sphere** $S_r(x)$ is the set of points which have distance r from a given point x. In the plane, the sphere is called a **circle**. A natural problem is to find the circumference $L = 2\pi$ of a unit circle, or the area $A = \pi$ of a unit disc, the area $F = 4\pi$ of a unit sphere and the volume $V = 4 = \pi/3$ of a unit sphere. Measuring the length of segments on the circle leads to new concepts like **angle** or **curvature**. Because the circumference of the unit circle in the plane is $L = 2\pi$, angle questions are tied to the number π, which Archimedes already approximated by fractions.

Also **volumes** were among the first quantities, Mathematicians wanted to measure and compute. A problem on **Moscow papyrus** dating back to 1850 BC explains the general formula $h(a^2+ab+b^2)/3$ for a truncated pyramid with base length a, roof length b and height h. Archimedes achieved to compute the **volume of the sphere**: place a cone inside a cylinder. The complement of the cone inside the cylinder has on each height h the area $\pi - \pi h^2$. The half sphere cut at height h is a disc of radius $(1 - h^2)$ which has area $\pi(1 - h^2)$ too. Since the slices at each height have the same area, the volume must be the same. The complement of the cone inside the cylinder has volume $\pi - \pi/3 = 2\pi/3$, half the volume of the sphere.

The first geometric playground was **planimetry**, the geometry in the flat two dimensional space. Highlights are **Pythagoras theorem**, **Thales theorem**, **Hippocrates theorem**, and **Pappus**

theorem. Discoveries in planimetry have been made later on: an example is the Feuerbach theorem from the 19th century or the Sadov theorem for quadrilaterals. Greek Mathematics is closely related to history. It starts with **Thales** goes over Euclid's era at 300 BC, and ends with the threefold destruction of Alexandria 47 BC by the Romans, 392 by the Christians and 640 by the Muslims. Geometry was also a place, where the **axiomatic method** was brought to mathematics: theorems are proved from a few statements which are called axioms like the 5 axioms of Euclid:

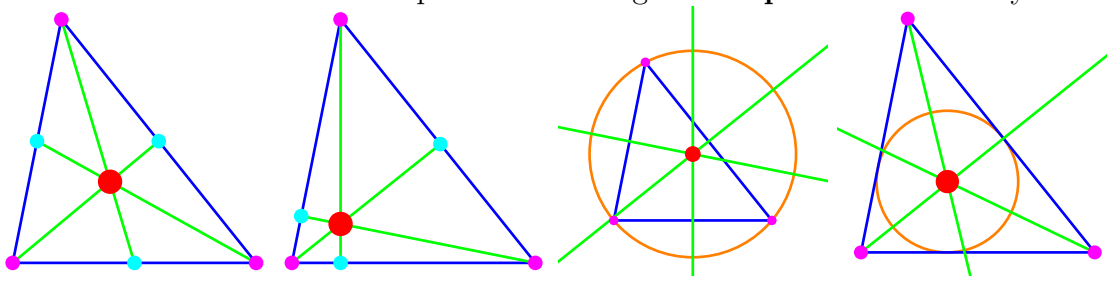
1. Any two distinct points A, B determines a line through A and B.
2. A line segment $[A, B]$ can be extended to a straight line containing the segment.
3. A line segment $[A, B]$ determines a circle containing B and center A.
4. All right angles are congruent.
5. If lines L, M intersect with a third so that inner angles add up to $< \pi$, then L, M intersect.

Euclid wondered whether the fifth postulate can be derived from the first four and called theorems derived from the first four the "absolute geometry". Only much later, with **Karl-Friedrich Gauss** and **Janos Bolyai** and **Nicolai Lobachevsky** in the 19'th century in **hyperbolic space** the 5'th axiom does not hold. Indeed, geometry can be generalized to non-flat, or even much more abstract situations. Basic examples are geometry on a sphere leading to **spherical geometry** or geometry on the Poincare disc, a **hyperbolic space**. Both of these geometries are non-Euclidean. **Riemannian geometry**, which is essential for **general relativity theory** generalizes both concepts to a great extent. An example is the geometry on an arbitrary surface. Curvatures of such spaces can be computed by measuring length alone, which is how long light needs to go from one point to the next.

An important moment in mathematics was the **merge of geometry with algebra**: this giant step is often attributed to **René Descartes**. Together with algebra, the subject leads to algebraic geometry which can be tackled with computers: here are some examples of geometries which are determined from the amount of symmetry which is allowed:

Euclidean geometry	Properties invariant under a group of rotations and translations
Affine geometry	Properties invariant under a group of affine transformations
Projective geometry	Properties invariant under a group of projective transformations
Spherical geometry	Properties invariant under a group of rotations
Conformal geometry	Properties invariant under angle preserving transformations
Hyperbolic geometry	Properties invariant under a group of Möbius transformations

Here are four pictures about the 4 special points in a triangle and with which we will begin. We will see why in each of these cases, the 3 lines intersect in a common point. It is a manifestation of a **symmetry** present on the space of all triangles. **size** of the distance of intersection points is constant 0 if we move on the space of all triangular **shapes**. It's Geometry!



Lecture 4: Number Theory

Number theory studies the structure of integers like prime numbers and solutions to Diophantine equations. Gauss called it the "Queen of Mathematics". Here are a few theorems and open problems.

An integer larger than 1 which is divisible by 1 and itself only is called a **prime number**. The number $2^{57885161}-1$ is the largest known prime number. It has 17425170 digits. **Euclid** proved that there are infinitely many primes: [Proof. Assume there are only finitely many primes $p_1 < p_2 < \ldots < p_n$. Then $n = p_1 p_2 \cdots p_n + 1$ is not divisible by any $p_1, \ldots, p_n$. Therefore, it is a prime or divisible by a prime larger than p_n.] Primes become more sparse as larger as they get. An important result is the **prime number theorem** which states that the n'th prime number has approximately the size $n\log(n)$. For example the $n = 10^{12}$'th prime is $p(n) = 29996224275833$ and $n\log(n) = 27631021115928.545...$ and $p(n)/(n\log(n)) = 1.0856...$ Many questions about prime numbers are unsettled: Here are four problems: the third uses the notation $(\Delta a)_n = |a_{n+1} - a_n|$ to get the absolute difference. For example: $\Delta^2(1,4,9,16,25...) = \Delta(3,5,7,9,11,...) = (2,2,2,2,...)$. Progress on prime gaps has been done recently: a paper which just appears showed $p_{n+1} - p_n$ is smaller than 100'000'000 eventually (Yitang Zhang April 2013) $p_{n+1} - p_n$ is smaller than 600 eventually (Maynard) . The largest known gap is 1476 which occurs after $p = 1425172824437699411$.

Twin prime	there are infinitely many primes p such that $p+2$ is prime.
Goldbach	every even integer $n > 2$ is a sum of two primes.
Gilbreath	If p_n enumerates the primes, then $(\Delta^k p)_1 = 1$ for all $k > 0$.
Andrica	The prime gap estimate $\sqrt{p_{n+1}} - \sqrt{p_n} < 1$ holds for all n.

If the sum of the proper divisors of a n is equal to n, then n is called a **perfect number**. For example, 6 is perfect as its proper divisors $1, 2, 3$ sum up to 6. All currently known perfect numbers are even. The question whether odd perfect numbers exist is probably the oldest open problem in mathematics and not settled. Perfect numbers were familiar to Pythagoras and his followers already. Calendar coincidences like that we have 6 work days and the moon needs "perfect" 28 days to circle the earth could have helped to promote the "mystery" of perfect number. **Euclid of Alexandria** (300-275 BC) was the first to realize that if $2^p - 1$ is prime then $k = 2^{p-1}(2^p - 1)$ is a perfect number: [Proof: let $\sigma(n)$ be the sum of **all** factors of n, including n. Now $\sigma(2^n - 1)2^{n-1}) = \sigma(2^n - 1)\sigma(2^{n-1}) = 2^n(2^n - 1) = 2 \cdot 2^n(2^n - 1)$ shows $\sigma(k) = 2k$ and verifies that k is perfect.] Around 100 AD, **Nicomachus of Gerasa** (60-120) classified in his work "Introduction to Arithmetic" numbers on the concept of perfect numbers and lists four perfect numbers. Only much later it became clear that Euclid got all the even perfect numbers: Euler showed that all even perfect numbers are of the form $(2^n - 1)2^{n-1}$, where $2^n - 1$ is prime. The factor $2^n - 1$ is called a **Mersenne prime**. [Proof: Assume $N = 2^k m$ is perfect where m is odd and $k > 0$. Then $2^{k+1}m = 2N = \sigma(N) = (2^{k+1} - 1)\sigma(m)$. This gives $\sigma(m) = 2^{k+1}m/(2^{k+1} - 1) = m(1 + 1/(2^{k+1} - 1)) = m + m/(2^{k+1} - 1)$. Because $\sigma(m)$ and m are integers, also $m/(2^{k+1} - 1)$ is an integer. It must also be a factor of m. The only way that $\sigma(m)$ can be the sum of only two of its factors is that m is prime and so $2^{k+1} - 1 = m$.] The first 39 **known Mersenne primes** are of the form $2^n - 1$ with n = 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213, 19937, 21701, 23209, 44497, 86243, 110503, 132049, 216091, 756839, 859433, 1257787, 1398269, 2976221, 3021377, 6972593, 13466917. There are 8 more known from which one does not know the rank of the corresponding Mersenne prime: n = 20996011, 24036583, 25964951, 30402457, 32582657,

37156667, 42643801,43112609,57885161. The last was found in January 2013 only. It is unknown whether there are infinitely many.

A polynomial equations for which all coefficients and variables are integers is called a **Diophantine equation**. The first Diophantine equation studied already by Babilonians is $x^2 + y^2 = z^2$. A solution (x, y, z) of this equation in positive integers is called a **Pythagorean triple**. For example, $(3, 4, 5)$ is a Pythagorean triple. Since 1600 BC, it is known that all solutions to this equation are of the form $(x, y, z) = (2st, s^2 - t^2, s^2 + t^2)$ or $(x, y, z) = (s^2 - t^2, 2st, s^2 + t^2)$, where s, t are different integers. [Proof. Either x or y has to be even because if both are odd, then the sum $x^2 + y^2$ is even but not divisible by 4 but the right hand side is either odd or divisible by 4. Move the even one, say x^2 to the left and write $x^2 = z^2 - y^2 = (z - y)(z + y)$, then the right hand side contains a factor 4 and is of the form $4s^2t^2$. Therefore $2s^2 = z - y, 2t^2 = z + y$. Solving for z, y gives $z = s^2 + t^2, y = s^2 - t^2$, $x = 2st$.]
Analyzing Diophantine equations can be difficult. Only 10 years ago, one has established that the **Fermat equation** $x^n + y^n = z^n$ has no solutions with $xyz \neq 0$ if $n > 2$. Here are some **open problems** for Diophantine equations. Are there nontrivial solutions to the following Diophantine equations?

$x^6 + y^6 + z^6 + u^6 + v^6 = w^6$	$x, y, z, u, v, w > 0$
$x^5 + y^5 + z^5 = w^5$	$x, y, z, w > 0$
$x^k + y^k = n!z^k$	$k \geq 2, n > 1$
$x^a + y^b = z^c, a, b, c > 2$	$\gcd(a, b, c) = 1$

The last equation is called **Super Fermat**. A Texan banker **Andrew Beals** once sponsored a prize of 100′000 dollars for a proof or counter example to the statement: "If $x^p + y^q = z^r$ with $p, q, r > 2$, then $\gcd(x, y, z) > 1$."
Given a prime like 7 and a number n we can add or subtract multiples of 7 from n to get a number in $\{0, 1, 2, 3, 4, 5, 6\}$. We write for example $19 = 12 \bmod 7$ because 12 and 19 both leave the rest 5 when dividing by 7. Or $5 * 6 = 2 \bmod 7$ because 30 leaves the rest 2 when dividing by 7. The most important theorem in elementary number theory is **Fermat's little theorem** which tells that if a is an integer and p is prime then $a^p - a$ is divisible by p. For example $2^7 - 2 = 126$ is divisible by 7. [Proof: use induction. For $a = 0$ it is clear. The binomial expansion shows that $(a + 1)^p - a^p - 1$ is divisible by p. This means $(a + 1)^p - (a + 1) = (a^p - a) + mp$ for some m. By induction, $a^p - a$ is divisible by p and so $(a + 1)^p - (a + 1)$.] An other beautiful theorem is **Wilson's theorem** which allows to characterize primes: It tells that $(n - 1)! + 1$ is divisible by n if and only if n is a prime number. For example, for $n = 5$, we verify that $4! + 1 = 25$ is divisible by 5. [Proof: assume n is prime. There are then exactly two numbers $1, -1$ for which $x^2 - 1$ is divisible by n. The other numbers in $1, \dots, n - 1$ can be paired as (a, b) with $ab = 1$. Rearranging the product shows $(n - 1)! = -1$ modulo n. Conversely, if n is not prime, then $n = km$ with $k, m < n$ and $(n - 1)! = ...km$ is divisible by $n = km$.]
The solution to systems of linear equations like $x = 3 \pmod 5, x = 2 \pmod 7$ is given by the **Chinese remainder theorem**. To solve it, continue adding 5 to 3 until we reach a number which leaves rest 2 to 7: on the list $3, 8, 13, 18, 23, 28, 33, 38$, the number 23 is the solution. Since 5 and 7 have no common divisor, the system of linear equations has a solution.
For a given n, how do we solve $x^2 - yn = 1$ for the unknowns y, x? A solution produces a square root x of 1 modulo n. For prime n, only $x = 1, x = -1$ are the solutions. For composite $n = pq$, more solutions $x = r \cdot s$ where $r^2 = -1 \bmod p$ and $s^2 = -1 \bmod q$ appear. Finding x is equivalent to factor n, because the greatest common divisor of $x^2 - 1$ and n is a factor of n. **Factoring is difficult** if the numbers are large. It assures that **encryption algorithms** work and that bank accounts and communications stay safe. Number theory, once the least applied discipline of mathematics has become one of the most applied one in mathematics.

Lecture 4: Number Theory

Twin prime conjecture

There are infinitely many prime twins $p, p+2$.

The first twin prime is $(3,5)$. The largest known prime twins $(p, p+2)$ have been found in 2011. It is $3756801695685 \cdot 2^{666669} \pm 1$. There are analogue problems for **cousin primes** $p, p+4$, **sexy primes** $p, p+6$ or **Germaine** primes, where $p, 2p+1$ are prime. Progress: we know that prime gaps of order 600 or smaller appear infinitely often. (Work of Zhang,Maynard,Tao)

Goldbach conjecture

Every even integer $n > 2$ is a sum of two primes.

The Goldbach conjecture has been verified numerically until $4 \cdot 10^{18}$. It is known that every sufficiently large odd number is the sum of 3 primes. One believes this "weak Goldbach conjecture" for 3 primes is true for every odd integer larger than 7.

Andrica conjecture

The prime gap estimate $\sqrt{p_{n+1}} - \sqrt{p_n} < 1$ holds.

For example $\sqrt{p_{1000}} - \sqrt{p_{999}} = \sqrt{7919} - \sqrt{7907} = 0.067...$. An other prime gap estimate conjectures is **Polignac's conjecture** claiming that there are infinitely many prime gaps for every even number n. It is stronger than the twin prime conjecture. It includes for example the claim that there are infinitely many cousin primes or sexy primes. **Legendre's conjecture** claims that there exists a prime between any two perfect squares. Between $16 = 4^2$ and $25 = 5^2$, there is the prime 23 for example.

Odd perfect numbers

Probably the oldest open problem in mathematics is the question

There is an odd perfect number.

A perfect number is equal to the sum of all its proper positive divisors. Like $6 = 1 + 2 + 3$. The search for perfect numbers is related to the search of large prime numbers. The largest prime number known today is $p = 2^{43112609} - 1$. It is called a Mersenne prime. Every even perfect number is of the form $2^{n-1}(2^n - 1)$ where $2^n - 1$ is prime.

Diophantine equations

Many problems about Diophantine equations, equations with integer solutions are unsettled. Here is an example:

Solve $x^5 + y^5 + z^5 = w^5$ for $x, y, z, w \in N$.

Also $x^5 + y^5 = u^5 + v^5$ has no nontrivial solutions yet. Probabilistic considerations suggest that there are no solutions. The analogue equation $x^4 + y^4 + z^4 = w^4$ had been settled by Noam Elkies in 1988 who found the identity $2682440^4 + 15365639^4 + 18796760^4 = 20615673^4$.

ABC Conjecture

The abc conjecture is:

If $a + b = c$, then $c \leq (\prod_{p|abc} p)^2$.

For example, for $10+22 = 32$, the prime factors of $abc = 7040$ are $2, 5, 11$ and indeed $32 \leq (2 * 5 * 11)^2 = 12100$. The abc-conjecture is open but implies Fermat's theorem for $n \geq 6$: assume $x^n + y^n = z^n$ with coprime x, y, z. Take $a = x^n, b = y^n, c = z^n$. The abc-conjecture gives $z^n \leq (\prod_{p|abc} p) \leq (abc)^2 < z^6$ establishing Fermat for $n \geq 6$. The cases $n = 3, 4, 5$ to Fermat have been known for a long time. In August 2012, there were rumors of an attack by Shinichi Mochizuki. During 2013 various mathematicians have tried to understand and verify the theory.

Lecture 5: Algebra

Algebra studies **algebraic structures** like "groups" and "rings". The theory allows to solve polynomial equations, characterize objects by its symmetries and is the heart and soul of many puzzles. Lagrange claims **Diophantus** to be the inventor of Algebra, others argue that the subject started with solutions of **quadratic equation** by **Mohammed ben Musa Al-Khwarizmi** in the book Al-jabr w'al muqabala of 830 AD. Solutions to equation like $x^2 + 10x = 39$ are solved there by **completing the squares**: add 25 on both sides go get $x^2 + 10x + 25 = 64$ and so $(x + 5) = 8$ so that $x = 3$.
The use of **variables** introduced in school in **elementary algebra** were introduced later. Ancient texts only dealt with particular examples and calculations were done with concrete numbers in the realm of **arithmetic**. **Francois Viete** (1540-1603) used first letters like A, B, C, X for variables.

The search for formulas for polynomial equations of degree 3 and 4 lasted 700 years. In the 16'th century, the cubic equation and quartic equations were solved. **Niccolo Tartaglia** and **Gerolamo Cardano** reduced the cubic to the quadratic: [first remove the quadratic part with $X = x - a/3$ so that $X^3 + aX^2 + bX + c$ becomes the **depressed cubic** $x^3 + px + q$. Now substitute $x = u - p/(3u)$ to get a quadratic equation $(u^6 + qu^3 - p^3/27)/u^3 = 0$ for u^3.] **Lodovico Ferrari** shows that the quartic equation can be reduced to the cubic. For the **quintic** however no formulas could be found. It was **Paolo Ruffini**, **Niels Abel** and **Évariste Galois** who independently realized that there are no formulas in terms of roots which allow to "solve" equations $p(x) = 0$ for polynomials p of degree larger than 4. This was an amazing achievement and the birth of "group theory".

Two important algebraic structures are **groups** and **rings**.

In a **group** G one has an operation $*$, an inverse a^{-1} and a one-element 1 such that $a * (b * c) = (a * b) * c, a * 1 = 1 * a = a, a * a^{-1} = a^{-1} * a = 1$. For example, the set Q^* of nonzero fractions p/q with multiplication operation $*$ and inverse $1/a$ form a group. The integers with addition and inverse $a^{-1} = -a$ and "1"-element 0 form a group too. A **ring** R has two compositions $+$ and $*$, where the plus operation is a group satisfying $a + b = b + a$ in which the one element is called 0. The multiplication operation $*$ has all group properties on R^* except the existence of an inverse. The two operations $+$ and $*$ are glued together by the **distributive law** $a * (b + c) = a * b + a * c$. An example of a ring are the **integers** or the **rational numbers** or the **real numbers**. The later two are actually **fields**, rings for which the multiplication on nonzero elements is a group too. The ring of integers are no field because an integer like 5 has no multiplicative inverse. The ring of rational numbers however form a field.

Why is the theory of groups and rings not part of arithmetic? First of all, a crucial ingredient of algebra is the appearance of **variables** and computations with these algebras without using concrete numbers. Second, the algebraic structures are not restricted to "numbers". Groups and rings are general structures and extend for example to objects like the set of all possible symmetries of a geometric object. The set of all **similarity operations** on the plane for example form a group. An important example of a ring is the **polynomial ring** of all polynomials. Given any ring R and a variable x, the set $R[x]$ consists of all polynomials with coefficients in R. The addition and multiplication is done like in $(x^2 + 3x + 1) + (x - 7) = x^2 + 4x - 7$. The problem to

for example can be written as $(x+1)(x-2)$ have a number theoretical flavor. Because symmetries of some structure form a group, we also have intimate connections with geometry. But this is not the only connection with geometry. Geometry also enters through the polynomial rings with several variables. Solutions to $f(x,y) = 0$ leads to geometric objects with shape and symmetry which sometimes even have their own algebraic structure. They are called **varieties**, a central object in **algebraic geometry**.

Arithmetic introduces addition and multiplication of numbers. Both form a group. The operations can be written additively or multiplicatively. Lets look at this a bit closer:
For integers, fractions and reals and the addition $+$, the 1 element 0 and inverse $-g$, we have a group. Many groups are written multiplicatively where the 1 element is 1. In the case of fractions or reals, 0 is not part of the multiplicative group because it is not possible to divide by 0. The nonzero fractions or the nonzero reals form a group. In all these examples the groups satisfy the commutative law $g * h = h * g$.
Here is a group which is not commutative: let G be the set of all rotations in space, which leave the unit cube invariant. There are 3*3=9 rotations around each major coordinate axes, then 6 rotations around axes connecting midpoints of opposite edges, then 2*4 rotations around diagonals. Together with the identity rotation e, these are 24 rotations. The group operation is the composition of these transformations.
An other example of a group is S_4, the set of all permutations of four numbers $(1,2,3,4)$. If $g : (1,2,3,4) \to (2,3,4,1)$ is a permutation and $h : (1,2,3,4) \to (3,1,2,4)$ is an other permutation, then we can combine the two and define $h * g$ as the permutation which does first g and then h. We end up with the permutation $(1,2,3,4) \to (1,2,4,3)$. The rotational symmetry group of the cube happens to be the same than the group S_4. To see this "isomorphism", label the 4 space diagonals in the cube by $1,2,3,4$. Given a rotation, we can look at the induced permutation of the diagonals and every rotation corresponds to exactly one permutation. The symmetry group can be introduced for any geometric object. For shapes like the triangle, the cube, the octahedron or tilings in the plane.

Symmetry groups describe geometric shapes by algebra.

Many **puzzles** are groups. A popular puzzle, the **15-puzzle** was invented in 1874 by **Noyes Palmer Chapman** in the state of New York. If the hole is given the number 0, then the task of the puzzle is to order a given random start permutation of the 16 pieces. To do so, the user is allowed to transposes 0 with a neighboring piece. Since every step changes the signature s of the permutation and changes the taxi-metric distance d of 0 to the end position by 1, only situations with even $s + d$ can be reached. It was **Sam Loyd** who suggested to start with an impossible solution and as an evil plot to offer 1000 dollars for a solution. The 15 puzzle group has 16!/2 elements and the "god number" is between 152 and 208. The **Rubik cube** is an other famous puzzle, which is a group. Exactly 100 years after the invention of the 15 puzzle, the Rubik puzzle was introduced in 1974. Its still popular and the world record is to have it solved in 5.55 seconds. Cubes 2x2x2 to 7x7x7 have been solved in a total time of 6 minutes. For the 3x3x3 cube, the god number is now known to be 20: one can always solve it in 20 or less moves.

Many puzzles are groups.

A small rubik type game is the "floppy", which is a third of the rubik and which has only 192 elements. An other example is the **Meffert's great challenge**. Probably the simplest example of a Rubik type puzzle is the **pyramorphix**. It is a puzzle based on the tetrahedron. Its group has only 24 elements. It is the group of all possible permutations of the 4 elements. It is the same group as the group of all reflection and rotation symmetries of the cube in three dimensions and also is relevant when understanding the solutions to the quartic equation discussed at the

Lecture 6: Calculus

Calculus formalizes the process of **taking differences** and **taking sums**. Differences measure **change**, sums explore how things **accumulate**. The process of taking differences has a limit called **derivative**. The process of taking sums will lead to the **integral**. These two processes are related in an intimate way. In this lecture, we look at these two processes in the simplest possible setup, where functions are evaluated on integers and where we do not take any limits.
Several dozen thousand years ago, numbers were represented by units like

$$1, 1, 1, 1, 1, 1, \ldots$$

for example carved in the Ishango bone. It took thousands of years until numbers were represented with symbols like

$$0, 1, 2, 3, 4, \ldots .$$

Using the modern concept of function, we can say $f(0) = 0, f(1) = 1, f(2) = 2, f(3) = 3$ and mean that the **function** f assigns to an input like 1001 an output like $f(1001) = 1001$. Lets call $Df(n) = f(n+1) - f(n)$ the **difference** between two function values. We see that the function f satisfies $Df(n) = 1$ for all n. We can also formalize the summation process. If $g(n) = 1$ is the function which is constant 1, then $Sg(n) = g(0) + g(1) + \ldots + g(n-1) = 1 + 1 + \cdots + 1 = n$. We see that $Df = g$ and $Sg = f$. Lets start with $f(n) = n$ and apply **summation** on that function:

$$Sf(n) = f(0) + f(1) + f(2) + \cdots + f(n-1) .$$

In our example, we get the values:

$$0, 1, 3, 6, 10, 15, 21, \ldots .$$

The new function $g = Sf$ satisfies $g(1) = 1, g(2) = 3, g(2) = 6$, etc. These numbers are called **triangular numbers**. From g we can get back f by taking difference:

$$Dg(n) = g(n+1) - g(n) = f(n) .$$

For example $Dg(5) = g(6) - g(5) = 15 - 10 = 5$ which indeed is $f(5)$. Finding a formula for the sum $Sf(n)$ is not so easy. Can you do it? When **Karl-Friedrich Gauss** was a 9 year old school kid, his teacher, a Mr. Büttner gave him the task to sum up the first 100 numbers $1+2+\cdots+100$. Gauss found the answer immediately by pairing things up: to add up $1 + 2 + 3 + \ldots + 100$ he would write this as $(1 + 100) + (2 + 99) + \cdots + (50 + 51)$ leading to 50 terms of 101 to get for $n = 101$ the value $g(n) = n(n-1)/2 = 5050$. Taking differences again is easier $Dg(n) = n(n+1)/2 - n(n-1)/2 = n = f(n)$.
Lets add now the triangular numbers up compute $h = Sg$. We get the sequence

$$0, 1, 4, 10, 20, 35, ...$$

called the **tetrahedral numbers**. One can $h(n)$ balls to build a tetrahedron of side length n. For example, $h(4) = 20$ golf balls are needed to build a tetrahedron of side length 4. The formula which holds for h is $\boxed{h(n) = n(n-1)(n-2)/6}$. Here is the fundamental theorem of calculus, which is the core of calculus:

$$\boxed{SDf(n) = f(n) - f(0), \quad DSf(n) = f(n) .}$$

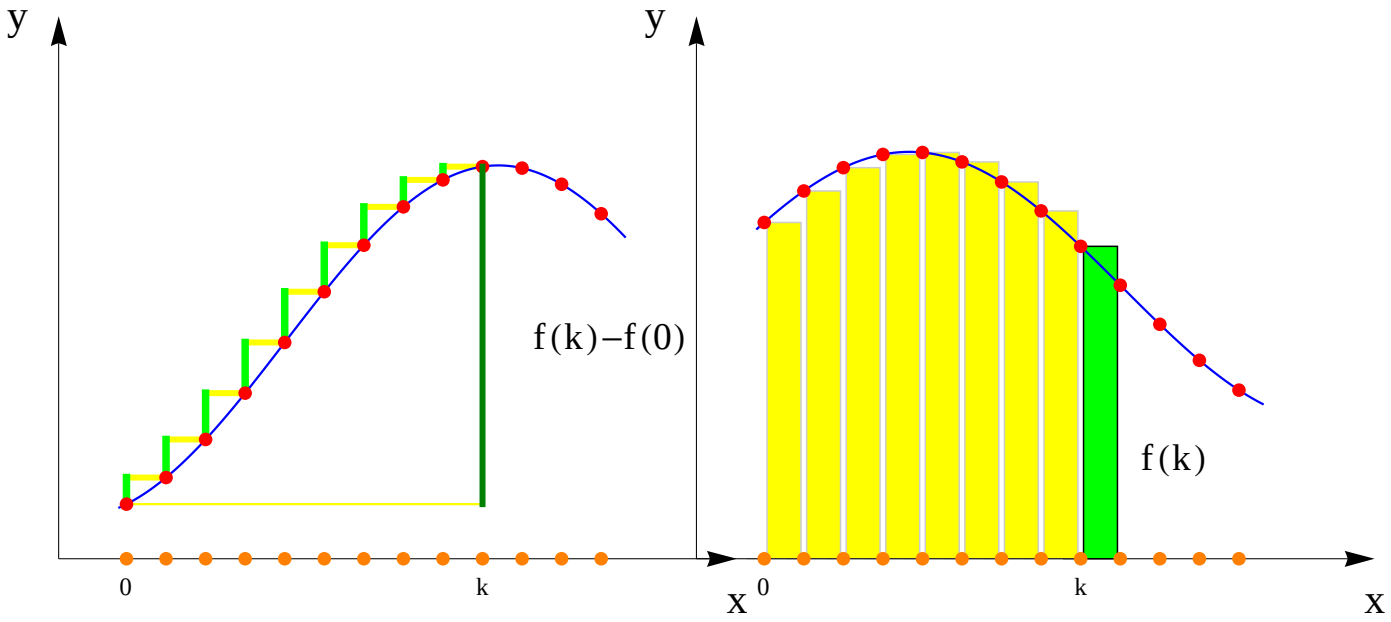
Proof.

$$SDf(n) = \sum_{k=0}^{n-1} [f(k+1) - f(k)] = f(n) - f(0) \ ,$$

$$DSf(n) = [\sum_{k=0}^{n-1} f(k+1) - \sum_{k=0}^{n-1} f(k)] = f(n) \ .$$

The process of adding up numbers will lead to the **integral** $\boxed{\int_0^x f(x)\, dx}$. The process of taking differences will lead to the **derivative** $\boxed{\frac{d}{dx} f(x)}$.

$$\boxed{\int_0^x \frac{d}{dt} f(t)\, dt = f(x) - f(0), \qquad \frac{d}{dx} \int_0^x f(t)\, dt = f(x)}$$



Theorem: Sum the differences and get $\boxed{SDf(kh) = f(kh) - f(0)}$

Theorem: Difference the sum and get $\boxed{DSf(kh) = f(kh)}$

If we define $[n]^0 = 1, [n]^1 = n, [n]^2 = n(n-1)/2, [n]^3 = n(n-1)(n-2)/6$ then $D[n] = [1], D[n]^2 = 2[n], D[n]^3 = 3[n]^2$ and in general

$$\boxed{\frac{d}{dx}[x]^n = n[x]^{n-1}}$$

The calculus you have just seen, contains the essence of single variable calculus. This core idea will become more powerful and natural if we use it together with the concept of limit.

1 **Problem:** The Fibonnacci sequence $1, 1, 2, 3, 5, 8, 13, 21, \ldots$ satisfies the rule $f(x) = f(x-1) + f(x-2)$. It defines a function on the positive integers. For example, $f(6) = 8$. What is the function $g = Df$, if we assume $f(0) = 0$? We take the difference between successive numbers and get the sequence of numbers

$$0, 1, 1, 2, 3, 5, 8, ...$$

which is the same sequence again. We can deduce from this recursion that f has the property that $\boxed{Df(x) = f(x-1)}$.

2 **Problem:** Take the same function f given by the sequence $1, 1, 2, 3, 5, 8, 13, 21, ...$ but now compute the function $h(n) = Sf(n)$ obtained by summing the first n numbers up. It gives the sequence $1, 2, 4, 7, 12, 20, 33, ...$ What sequence is that?
Solution: Because $Df(x) = f(x-1)$ we have $f(x) - f(0) = SDf(x) = Sf(x-1)$ so that $Sf(x) = f(x+1) - f(1)$. Summing the Fibonnacci sequence produces the Fibonnacci sequence shifted to the left with $f(2) = 1$ is subtracted. It has been relatively easy to find the sum, because we knew what the difference operation did. This example shows:

We can study differences to understand sums.

The next problem illustrates this too:

3 **Problem:** Find the next term in the sequence
2 6 12 20 30 42 56 72 90 110 132 . **Solution:** Take differences

2	6	12	20	30	42	56	72	90	110	132
2	4	6	8	10	12	14	16	18	20	22
2	2	2	2	2	2	2	2	2	2	2
0	0	0	0	0	0	0	0	0	0	0

.

Now we can add an additional number, starting from the bottom and working us up.

2	6	12	20	30	42	56	72	90	110	132	156
2	4	6	8	10	12	14	16	18	20	22	24
2	2	2	2	2	2	2	2	2	2	2	2
0	0	0	0	0	0	0	0	0	0	0	0

4 **Problem:** The function $f(n) = 2^n$ is called the **exponential function**. We have for example $f(0) = 1, f(1) = 2, f(2) = 4, \ldots$. It leads to the sequence of numbers

n=	0	1	2	3	4	5	6	7	8	...
f(n)=	1	2	4	8	16	32	64	128	256	...

We can verify that f satisfies the equation $\boxed{Df(x) = f(x)}$. because $Df(x) = 2^{x+1} - 2^x = (2-1)2^x = 2^x$.
This is an important special case of the fact that

The derivative of the exponential function is the exponential function itself.

The function 2^x is a special case of the exponential function when the Planck constant is equal to 1. We will see that the relation will hold for any $h > 0$ and also in the limit $h \to 0$, where it becomes the classical exponential function e^x which plays an important role in science.

Calculus has many applications: computing areas, volumes, solving differential equations. It even has applications in arithmetic. Here is an example for illustration. It is a proof that π is irrational. This is especially appropriete since next Friday is π day!

We show here the proof by Ivan Niven is given in a book of Niven-Zuckerman-Montgomery. It originally appeared in 1947 (Ivan Niven, Bull.Amer.Math.Soc. 53 (1947),509). The proof illustrates how calculus can help to get results in arithmetic.
Proof. Assume $\pi = a/b$ with positive integers a and b. For any positive integer n define

$$f(x) = x^n(a - bx)^n/n! \ .$$

We have $f(x) = f(\pi - x)$ and

$$0 \leq f(x) \leq \pi^n a^n/n! (*)$$

for $0 \leq x \leq \pi$. For all $0 \leq j \leq n$, the j-th derivative of f is zero at 0 and π and for $n <= j$, the j-th derivative of f is an integer at 0 and π.

The function

$$F(x) = f(x) - f^{(2)}(x) + f^{(4)}(x) - ... + (-1)^n f^{(2n)}(x)$$

has the property that $F(0)$ and $F(\pi)$ are integers and $F + F'' = f$. Therefore, $(F'(x)\sin(x) - F(x)\cos(x))' = f\sin(x)$. By the fundamental theorem of calculus, $\int_0^\pi f(x)\sin(x)\,dx$ is an integer. Inequality (*) implies however that this integral is between 0 and 1 for large enough n. For such an n we get a contradiction.

Lecture 7: Set Theory and Logic

We will mostly focus on the work of two mathematicians: **Georg Cantor** and **Kurt Gödel**. Their mathematics changed our way we think about mathematics. In both cases, the mathematics community needed time to absorb the implications of the revolutions.
Hilbert said about Cantor "Nobody will drive us from the paradise that Cantor has created for us". Cantor clarified the term "cardinality" is, showed that certain infinities like that the cardinality of points in the plane or points in space are the same and most importantly showed that different infinities exist. Gödels theorems show that mathematics and knowledge in general can not be exhausted by listing a sequence of basic truths from which everything follows. Whenever we make such a list, there are statements which are independent of the system. It would be a mistake to take this as a limitation of mathematics, in contrary it shows that mathematics is inexhaustible: there is always something more to explore.

Counting: Set theory

We first demonstrate that one can compute with sets like with numbers. There is an addition, the symmetric difference and a multiplication, the intersection. With these two operations, we prove the familiar rules of arithmetic

$$A + B = B + A, A \cdot B = B \cdot A, A \cdot (B + C) = A \cdot B + A \cdot C$$

hold. This is a Boolean algebra. There is a set which plays the role of 0. Which one is it? There is also a set which plays the role of 1. Which one is it?

Counting: Hilbert's Hotel

Hilbert's hotel is located on route 8. It has countably many rooms numbered $1, 2, 3, \ldots$. The hotel is fully booked. As a newcomer arrives. David, the hotel manager is mortified. David has an idea and moves guest in room i to room $i + 1$ and gives the newcomer the first room 1.

An other day, the hotel is empty but a large group arrives. They are the "fractions" on their way to a cardinal match with the "squares". Can David accommodate them? He thinks hard and finally manages.

In the summer, the "reals" appear. David is not there but has George, the apprentice is in the office. The group consists of all real numbers between 0 and 1. Can George accomodate them? As much as he tries to shift and renumber, he can not do it.

Counting: the interval

The interval $(-1,1)$ has the same cardinality than the real line. The function $f(x) = \tan(\pi x/2)$ maps the interval $(-1,1)$ onto the real line, one to one.

The square $(0,1)\times(0,1)$ has the same cardinality than the real line. A bijection can be constructed by $f(0.a_1a_2a_3a_4...) = (0.a_1a_3a_5, 0.a_2a_4a_6...)$.

Arithmetic with sets

One can calculate with sets as with numbers. They form a "Boolean ring".

Addition: $A + B = A\Delta B$ with the zero element $\emptyset$
Multiplication: $A \cdot B = A \cap B$ with the one element Ω.

All the rules of the real numbers apply but there are additional consequences which appear a bit strange $A + A = 0$ and $A^2 = A \cdot A = A$. This means that A is its own additive inverse.

Paradoxa

We have seen a few paradoxa like the **Liars paradox** "I lie", the **barbers paradox** "the barber is the person who shaves everybody who does not shave him or herself", the **surprise exam problem** "it is impossible to make a surprise exam problem", the **heap problem** "take a grain away from a heap keeps it a heap", the **biographer's problem** "who needs one year to write one day of his biography", Here is an other, the **Berry paradox** which comes somehow close to the Goedel numbering:

The smallest integer not definable in less than 11 words.

The problem is that this number is defined with 10 words. This looks like a stupid example but it illustrates that there are properties of numbers like "the shortest way to describe the number" which is not computable.

Axiom of choice

The **axiom of choice (C)** has a nonconstructive nature which can lead to seemingly paradoxical results like the **Banach Tarski paradox**: one can cut the unit ball into 5 pieces, rotate and translate the pieces to assemble two identical balls of the same size than the original ball. Gödel and Cohen showed that the axiom of choice is logically independent of the other axioms ZF.

Lecture 8: Probability theory

Probability theory is the science of chance. It starts with **combinatorics** and leads to a theory of **stochastic processes**. Historically, probability theory initiated from gambling problems as in **Girolamo Cardano's** gamblers manual in the 16th century. A great moment of mathematics occurred, when **Blaise Pascal** and **Pierre Fermat** jointly laid a foundation of mathematical probability theory.

It took mathematicians longer to formalize "randomness" precisely. Here is the setup as which it had been put forward by **Andrey Kolmogorov**: all possible experiments of a situation are modeled by a set Ω, the "laboratory". A measurable subset of experiments is called an "event". Measurements are done by real-valued functions X. These functions are called **random variables** and are used to **observe the laboratory**.

As an example, lets model the process of throwing a coin 5 times. An experiment is a word like $httht$, where h stands for "head" and t represents "tail". The laboratory consists of all such 32 words. We could look for example at the event A that the first two coin tosses are tail. It is the set $A = \{ttttt, tttth, tttht, ttthh, tthtt, tthth, tthht, tthhh\}$. We could look at the random variable which assigns to a word the number of heads. For every experiment, we get a value, like for example, $X[tthht] = 2$.

In order to make statements about randomness, the concept of a **probability measure** is needed. This is a function P from the set of all events to the interval $[0, 1]$. It should have the property that $P[\Omega] = 1$ and $P[A_1 \cup A_2 \cup ...] = P[A_1] + P[A_2] + ...$, if A_i are disjoint events.

The most natural probability measure on a finite set Ω is $P[A] = \|A\|/\|\Omega\|$, where $\|A\|$ stands for the number of elements in A. It is the "number of good cases" divided by the "number of all cases". For example, to count the probability of the event A that we throw 3 heads during the 5 coin tosses, we have $|A| = 10$ possibilities. Since the entire laboratory has $|\Omega| = 32$ possibilities, the probability of the event is $10/32$. In order to study these probabilities, one needs **combinatorics**:

How many ways are there to:	**The answer is:**
rearrange or permute n elements	$n! = n(n-1)...2 \cdot 1$
choose k from n with repetitions	n^k
pick k from n if order matters	$\frac{n!}{(n-k)!}$
pick k from n with order irrelevant	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

The **expectation** of a random variable $E[X]$ is defined as the sum $m = \sum_{\omega \in \Omega} X(\omega) P[\{\omega\}]$. In our coin toss experiment, this is $5/2$. The **variance** of X is the expectation of $(X - m)^2$. In our coin experiments, it is $5/4$. Its square root is called the **standard deviation**. This is the expected deviation from the mean. An event happens **almost surely** if the event has probability 1.

An important case of a random variable is $X(\omega) = \omega$ on $\Omega = R$ equipped with probability $P[A] = \int_A \frac{1}{\sqrt{\pi}} e^{-x^2} \, dx$, the **standard normal distribution**. Analyzed first by **Abraham de**

Moivre in 1733, it was studied by **Carl Friedrich Gauss** in 1807 and therefore also called **Gaussian distribution**.

Two random variables X, Y are called **decorrelated**, if $E[XY] = E[X] \cdot E[Y]$. If for any functions f, g also $f(X)$ and $g(Y)$ are decorrelated, then X, Y are called **independent**. Two random variables are said to have the same distribution, if for any $a < b$, the events $\{a \leq X \leq b\}$ and $\{a \leq Y \leq b\}$ are independent. If X, Y are decorrelated, then the relation $\mathrm{Var}[X] + \mathrm{Var}[Y] = \mathrm{Var}[X + Y]$ holds which is just **Pythagoras theorem**, because decorrelated can be understood geometrically: $X - E[X]$ and $Y - E[Y]$ are orthogonal. A common problem is to study the sum of independent random variables X_n with identical distribution. One abbreviates this IID. Here are the three most important theorems which we formulate in the case, whee all random variables are assumed to have expectatation 0 and standard deviation 1. Let $S_n = X_1 + ... + X_n$ be the n'th sum of the IID random variables. It is also called a **random walk**.

LLN **Law of Large Numbers** assures that S_n/n converges to 0.

CLT **Central Limit Theorem**:$S_n/\sqrt{n}$ approaches the Gaussian distribution.

LIL **Law of Iterated Logarithm:** $S_n/\sqrt{2n \log\log(n)}$ accumulates in $[-1, 1]$.

The LLN shows that one can find out about the expectation by averaging experiments. The CLT explains why one sees the standard normal distribution so often. The LIL finally gives us a precise estimate how fast S_n grows. Things become interesting if the random variables are no more independent. Generalizing LLN,CLT,LIL to such situations is part of ongoing research.

Here are two open questions in probability theory:

Are $\pi, e, \sqrt{2}$... normal: do all digits appear with the same frequency?
What growth rates Λ_n can occur in S_n/Λ_n having limsup 1 and liminf -1?

For the second question, there are examples for $\Lambda_n = 1, \lambda_n = \log(n)$ and of course $\lambda_n = \sqrt{n \log\log(n)}$ from LIL if the random variables are independent. Examples of random variables which are not independent are $X_n = \cos(n\sqrt{2})$.

Statistics is the science of modeling random events in a probabilistic setup. Given data points, we want to find a **model** which fits the data best. This allows to **understand the past**, **predict the future** or **discover laws of nature**. The most common task is to find the **mean** and the **standard deviation** of some data. The mean is also called the **average** and given by $m = \frac{1}{n}\sum_{k=1}^{n} x_k$. The variance is $\sigma^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - m)^2$ with standard deviation σ.

A sequence of random variables X_n define a so called **stochastic process**. Continuous versions of such processes are where X_t is a curve of random random variables. An important example is **Brownian motion**, which is a model of a random particles.

Besides gambling and analyzing data, also **physics** was an important motor to develop probability theory. An example is statistical mechanics where laws of nature are studied with probabilistic methods. A famous physical law is **Ludwig Boltzmann's** relation $S = k\log(W)$ for entropy, a formula which decorates Boltzmann's tombstone. The **entropy** of a probability measure $P[\{k\}] = p_k$ on a finite set $\{1, ..., n\}$ is defined as $S = -\sum_{i=1}^{n} p_i \log(p_i)$. Today, we would reformulate Boltzmann's law and say that it is the expectation $S = E[\log(W)]$ of the logarithm of the "Wahrscheinlichkeit" random variable $W(i) = 1/p_i$ on $\Omega = \{1, ..., n\}$. Entropy is important because nature tries to maximize it

Lecture 9: Topology

Topology studies properties of geometric objects which do not change under continuous reversible deformations. For a topologist, a coffee cup with 1 handle is the same as a doughnut. One can deform one into the other without punching any holes in it or ripping things apart. Similarly, a plate and a croissant are the same. But a croissant is not equivalent to a bagle. On a bagle, there are closed curves which can not be deformed to a point. For a topologist the letters O and P are the same but different from the letter B. The mathematical setup is beautiful: a **topological space** is a set X with a set O of subsets of X containing both $\emptyset$ and X such that finite intersections and arbitrary unions in O are in O. Sets in O are called **open sets** and O is called a **topology**. The complement of an open set is called **closed**. Examples of topologies are the **trivial topology** $O = \{\emptyset, X\}$, where no open sets besides the empty set and X exist or the discrete topology $O = \{A \subset X\}$,, where every subset is open. But these are in general not interesting. An important example on the plane X is the collection O of sets U in the plane X for which every point is the center of a small disc still contained in U. A special class of topological spaces are **metric spaces**, where a set X is equipped with a **distance function** $d(x, y) = d(y, x) \geq 0$ which satisfies the **triangle inequality** $d(x, y) + d(y, z) \geq d(x, z)$ and for which $d(x, y) = 0$ if and only if $x = y$. A set U in a metric space is open if to every x in U, there is a **ball** $B_r(x) = \{y | d(x, y) < r\}$ of positive radius r contained in U. Metric spaces are topological spaces but not all topological spaces are metric: the trivial topology for example is not in general. For doing **calculus** on a topological space X, each point has a neighborhood called **chart** which is topologically equivalent to a disc in Euclidean space. Finitely many such neighborhoods covering X form an **atlas** of X. If the charts are glued together with identification maps on the intersection one obtains a **manifold**. Two dimensional examples are the **sphere**, the **torus**, the projective plane or the **Klein bottle**. Topological spaces X, Y are called **homeomorphic** meaning "topologically equivalent" if there is an invertible map from X to Y which is also induces an invertible map on the corresponding topologies. A basic task is to decide whether two spaces are equivalent in this sense or not. The surface of the coffee cup for example is equivalent in this sense to the surface of a doughnut but it is not equivalent to the surface of a sphere.

Many properties of geometric spaces can be understood by replacing them with **graphs** forming a skeleton of the space. A graph is a finite collection of vertices V together with a finite set of edges E, where each edge connects two points in V. For example, the set V of cities in the US where the edges are pairs of cities connected by a street is a graph. The **Königsberg bridge problem** was a trigger puzzle for the study of graph theory. **Polyhedra** were an other start in graph theory. It study is losely related to the analysis of surfaces. The reason is that one can see polyhedra as discrete versions of surfaces. In computer graphics for example, surfaces are rendered as finite graphs, using triangularizations.

The **Euler characteristic** of a convex polyhedron is a remarkable topological invariant. It is

$$V - E + F = 2\,,$$

where V is the number of vertices, E the number of edges and F the number of **faces**. This number is equal to 2 for connected polyhedra in which every closed loop can be pulled together to a point. This formula for the Euler characteristic is also called **Euler's gem**, a fact which comes with a rich history. **René Descartes** seems have stumbled upon it and written it down in a secret notebook. It was Leonard Euler in 1752 was the first to proved the formula for convex polyhedra. A convex polyhedron is called a **platonic solid**, if all vertices are on the unit sphere, all edges have the same length and all faces are congruent polygons. A theorem of Theaetetus states that there are only 5 platonic solids: [Proof: Assume the faces

are regular n-gons and m of them meet at each vertex. Beside the Euler relation $V+E+F=2$, a polyhedron also satisfies the relations $nF=2E$ and $mV=2E$ which are obvious from counting vertices or edges in different ways. This gives $2E/m-E+2E/n=2$ or $1/n+1/m=1/E+1/2$. From $n \geq 3$ and $m \geq 3$ we see that it is impossible that both m and n are larger than 3. There are now nly two possibilities: either $n=3$ or $m=3$. In the case $n=3$ we have $m=3,4,5$ in the case $m=3$ we have $n=3,4,5$. The five possibilities $(3,3),(3,4),(3,5),(4,3),(5,3)$ represent the 5 platonic solids.] The pairs (n,m) are called the **Schläfly symbol** of the polyhedron:

Name	V	E	F	V-E+F	Schläfli
tetrahedron	4	6	4	2	$\{3,3\}$
hexahedron	8	12	6	2	$\{4,3\}$
octahedron	6	12	8	2	$\{3,4\}$

Name	V	E	F	V-E+F	Schläfli
dodecahedron	20	30	12	2	$\{5,3\}$
icosahedron	12	30	20	2	$\{3,5\}$

The Greeks proved this more geometrically: Euclid showed in his "Elements" that at each vertex, we can attach 3,4 or 5 equilateral triangles, 3 squares or 3 regular pentagons. (6 triangles, 4 squares or 4 pentagons would lead to a total angle which is too large because each corner must have at least 3 different edges). **Simon Antoine-Jean L'Huilier** refined in 1813 Euler's formula to situations with holes: $V-E+F=2-2g$, where g is the number of holes. For a doughnut with one hole we have $V-E+F=0$. . Cauchy first proved that there are exactly 4 non-convex regular **Kepler-Poinsot** polyhedra. Their Euler characteristic can be different.

Name	V	E	F	V-E+F	Schläfli
small stellated dodecahedron	12	30	12	-6	$\{5/2,5\}$
great dodecahedron	12	30	12	-6	$\{5,5/2\}$
great stellated dodecahedron	20	30	12	2	$\{5/2,3\}$
great icosahedron	12	30	20	2	$\{3,5/2\}$

If two different face types are allowed but each vertex still look the same, one obtains 13 **semi-regular polyhedra.** They were first studied by **Archimedes** in 287 BC. Since his work is lost, **Johannes Kepler** is considered the first person since antiquity to describe the whole set of thirteen in his "Harmonices Mundi". The Euler characteristic $\chi=2-2g$ is also useful for surfaces. One can reduce the question to graphs, triangularizations of the surface. The Euler characteristic completely characterizes smooth compact surfaces if they are orientable. A non-orientable surface, the **Klein bottle** can be obtained by gluing ends of the Möbius strip. Classifying higher dimensional manifolds is more difficult and finding more invariants is part of modern research. Higher analogues of polyhedra are called **polytopes** (Alicia Boole Stott). **Regular polytopes** are the analogue of the platonic solids in higher dimensions. Here they are for the first few dimensions:

dimension	name	Schläfli symbols
2:	Regular polygons	$\{3\},\{4\},\{5\},...$
3:	Platonic solids	$\{3,3\},\{3,4\},\{3,5\},\{4,3\},\{5,3\}$
4:	Regular 4D polytopes	$\{3,3,3\},\{4,3,3\},\{3,3,4\},\{3,4,3\},\{5,3,3\},\{3,3,5\}$
≥ 5:	Regular polytopes	$\{3,3,3,\ldots,3\},\{4,3,3,\ldots,3\},\{3,3,3,\ldots,3,4\}$

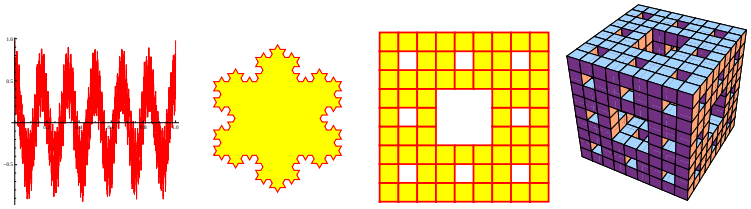
Ludwig Schläfly found in 1852 that there are exactly six convex regular convex 4-polytopes or **polychora**. The expression "choros" is Greek for "space". Schlaefli's polyhedral formula tells that for any **convex polytope** in four dimensions, the relation $V-E+F-C=0$ holds, where C is the number of 3-dimensional **chambers**. In dimensions 5 and higher, there are only 3 types of polytopes: the higher dimensional analogues of the tetrahedron, octahedron and the cube. A general formula $\sum_{i=0}^{d-1}(-1)^i V_i = 1-(-1)^d$ gives the Euler characteristic of a convex polytop in d dimensions with i-dimensional parts V_i.

Lecture 10: Analysis

Analysis is the science of measure and optimization. As a collection of mathematical fields, it contains **real and complex analysis**, **functional analysis**, **harmonic analysis** and **calculus of variations**. Analysis has relations to calculus, geometry, topology, probability theory and dynamics. We will focus mostly on "the geometry of fractals" today. Examples are Julia sets which belong to the subfield of "complex analysis" of "dynamical systems". "Calculus of variations" is illustrated by the Kakeya needle set in "geometric measure theory", a glimpse of "Fourier analysis" is seen by looking at functions which have fractal graphs, "spectral theory" as part of functional analysis is represented by the "Hofstadter butterfly". As we take a tabloid approach and describe the topic with gossip about some "pop icons" in each field, consider this page the center fold page of the "Analytical Enquirer".

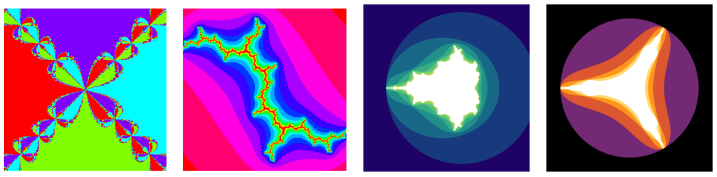
A **fractal** is a set with non-integer dimension. An example is the **Cantor set**, as discovered in 1875 by Henry Smith. Start with the unit interval. Cut the middle third, then cut the middle third from both parts then the middle parts of the four parts etc. The limiting set is the Cantor set. The mathematical theory of fractals belongs to **measure theory** and can also be thought of a playground for real analysis or topology. The term **fractal** had been introduced by Benoit Mandelbrot in 1975. Dimension can be defined in different ways. The simplest is the **box counting definition** which works for most household fractals: if we need n squares of length r to cover a set, then $d = -\log(n)/\log(r)$ converges to the dimension of the set with $r \to 0$. A curve of length L for example needs L/r squares of length r so that its dimension is 1. A region of area A needs A/r^2 squares of length r to be covered and its dimension is 2. The Cantor set needs to be covered with $n = 2^m$ squares of length $r = 1/3^m$. Its dimension is $-\log(n)/\log(r) = -m\log(2)/(m\log(1/3)) = \log(2)/\log(3)$.
Examples of fractals (for the first, the dimension is not yet known):

Weierstrass function	1872
Koch snowflake	1904
Sierpinski carpet	1915
Menger sponge	1926



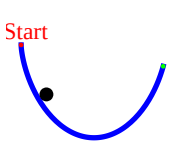
Complex analysis extends calculus to the complex. It deals with functions $f(z)$ defined in the complex plane. Integration is done along paths. Complex analysis completes the understanding about functions. It also provides more examples of fractals by iterating functions like the **quadratic map** $f(z) = z^2 + c$:

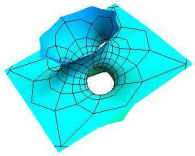
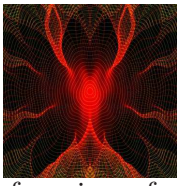
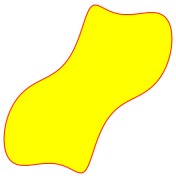
Newton method	1879
Julia sets	1918
Mandelbrot set	1978
Mandelbar set	1989



Particularly famous are the **Douady rabbit** and the **dragon**, the **dendrite**, the **airplane**.
Calculus of variations is calculus in infinite dimensions. Taking derivatives is called taking "variations". Historically, it started with the problem to find the curve of fastest fall leading to the **Brachistochrone** $r(t) = (t - \sin(t), 1 - \cos(t))$. In calculus, we find maxima and minima of functions. In calculus of variations, we extremize on much larger spaces. Here are some examples of problems:

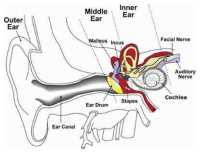
Brachistochrone	1696
Minimal surface	1760
Geodesics	1830
Isoperimetric problem	1838
Kakeya Needle problem	1917



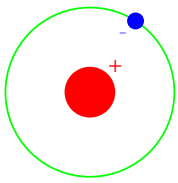
Fourier theory decomposes a function into basic components of various frequencies $f(x) = a_1 \sin(x) + a_2 \sin(2x) + a_3 \sin(3x)...$. The numbers a_i are called Fourier coefficients. Our ear does such a decomposition, when we listen to music. By distinguish different frequencies, our ear produces a Fourier analysis.

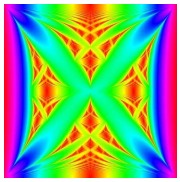
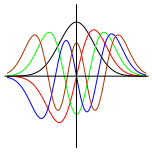
Fourier series	1729
Fourier transform (FT)	1811
Discrete FT	Gauss?
Wavelet transform	1930


The Weierstrass function mentioned above is given as the Fourier series $\sum_n a^n \cos(\pi b^n x)$ with $0 < a < 1, ab > 1 + 3\pi/2$. The dimension of its graph is believed to be $2 + \log(a)/\log(b)$.
Spectral theory analyzes linear maps L. The **spectrum** are the real numbers E such that $L - E$ is not invertible. A Hollywood celebrity among all linear maps is the **Matthieu operator** $L(x)_n = x_{n+1} + x_{n-1} + (2 - 2\cos(cn))x_n$: if we draw the spectrum for for each c, we see the **Hofstadter butterfly**. For fixed c the map describes the behavior of an electron in an almost periodic crystal. An other famous system is the **quantum harmonic oscillator**, $L(f) = f''(x) + f(x)$, the **vibrating drum** $L(f) = f_{xx} + f_{yy}$, where f is the amplitude of the drum and $f = 0$ on the boundary of the drum.

Hydrogen atom	1914
Hofstadter butterfly	1976
Harmonic oscillator	1900
Vibrating drum	1680




All these examples in analysis look unrelated at first. Fractal geometry ties many of them together: spectra are often fractals, minimal configurations have fractal nature, like in solid state physics or in **diffusion limited aggregation** or in other critical phenomena like **percolation** phenomena, **cracks** in solids or the formation of **lighting bolts** In Hamiltonian mechanics, minimal energy configurations are often fractals like **Mather theory**. And solutions to minimizing problems lead to fractals in a natural way like when you have the task to turn around a needle on a table by 180 degrees and minimize the area swept out by the needle. The minimal turn leads to a Kakaya set, which is a fractal. Finally, lets mention some unsolved problems in analysis: does the **Riemann zeta function** $f(z) = \sum_{n=1}^{\infty} 1/n^z$ have all nontrivial roots on the axis $Re(z) = 1/2$? This question is called the **Riemann hypothesis** and is the most important open problem in mathematics. It is an example of a question in **analytic number theory** which also illustrates how analysis has entered into number theory. Some mathematicians think that spectral theory might solve it. Also the Mandelbrot set M is not understood yet: the "holy grail" in the field of complex dynamics is the problem whether it M is locally connected. From the Hofstadter butterfly one knows that it has measure zero. What is its dimension? An other open question in spectral theory is the "can one hear the sound of a drum" problem which asks whether there are two convex drums which are not congruent but which have the same spectrum. In the area of calculus of variations, just one problem: how long is the shortest curve in space such that its convex hull (the union of all possible connections between two points on the curve) contains the unit ball.

Lecture 11: Cryptography

Cryptography is the theory of **codes**. Two important aspects of the field are the **encryption** rsp. **decryption** of information and **error correction**. Both are crucial in daily life. When getting access to a computer, viewing a bank statement or when taking money from the ATM, encryption algorithms are used. When phoning, surfing the web, accessing data on a computer or listening to music, error correction algorithms are used. Since our lives have become more and more digital: music, movies, books, journals, finance, transportation, medicine, and communication have become digital, we rely on strong error correction to avoid errors and encryption to assure things can not be tempered with. Without error correction, airplanes would crash: small errors in the memory of a computer would produce glitches in the navigation and control program. In a computer memory every hour a couple of bits are altered, for example by cosmic rays. Error correction assures that this gets fixed. Without error correction music would sound like a 1920 gramophone record. Without encryption, everybody could intrude electronic banks and transfer money. Medical history shared with your doctor would all be public. Before the digital age, error correction was assured by extremely redundant information storage. Writing a letter on a piece of paper displaces billions of billions of molecules in ink. Now, changing any single bit could give a letter a different meaning. Before the digital age, information was kept in well guarded safes which were physically difficult to penetrate. Now, information is locked up in computers which are connected to other computers. Vaults, money or voting ballots are secured by mathematical algorithms which assure that information can only be accessed by authorized users. Also life needs error correction: information in the genome is stored in a **genetic code**, where a error correction makes sure that life can survive. A cosmic ray hitting the skin changes the DNA of a cell, but in general this is harmless. Only a larger amount of radiation can render cells cancerous.

How can an encryption algorithm be safe? One possibility is to invent a new method and keep it secret. An other is to use a well known encryption method and rely on the **difficulty of mathematical computation tasks** to assure that the method is safe. History has shown that the first method is unreliable. Systems which rely on "security through obfuscation" usually do not last. The reason is that it is tough to keep a method secret if the encryption tool is distributed. Reverse engineering of the method is often possible, for example using plain text attacks. Given a map T, a third party can compute pairs $x, T(x)$ and by choosing specific texts figure out what happens.

The **Caesar cypher** permutes the letters of the alphabet. We can for example replace every letter A with B, every letter B with C and so on until finally Z is replaced with A. The word "Mathematics" becomes so encrypted as "Nbuifnbujdt". Caesar would shift the letters by 3. The right shift just discussed was used by his Nephew Augustus. **Rot13** shifts by 13, and **Atbash cypher** reflects the alphabet, switch A with Z, B with Y etc. The last two examples are involutive: encryption is decryption. More general cyphers are obtained by permuting the alphabet. Because of $26! = 403291461126605635584000000 \sim 10^{27}$ permutations, it appears first that a brute force attack is not possible. But Cesar cyphers can be cracked very quickly using statistical analysis. If we know the frequency with which letters appear and match the frequency of a text we can figure out which letter was replaced with which. The **Trithemius cypher** prevents this simple analysis by changing the permutation in each step. It is called a polyalphabetic substitution cypher. Instead of a simple permutation, there are many permutations. After transcoding a letter, we also change the key. Lets take a simple example. Rotate for the first letter the alphabet by 1, for the second

letter, the alphabet by 2, for the third letter, the alphabet by 3 etc. The word "Mathematics" becomes now "Ncwljshbrmd". Note that the second "a" has been translated to something different than a. A frequency analysis is now more difficult. The **Viginaire cypher** adds even more complexity: instead of shifting the alphabet by 1, we can take a key like "BCNZ", then shift the first letter by 1, the second letter by 3 the third letter by 13, the fourth letter by 25 the shift the 5th letter by 1 again. While this cypher remained unbroken for long, a more sophisticated frequency analysis which involves first finding the length of the key makes the cypher breakable. With the emergence of computers, even more sophisticated versions like the German **enigma** had no chance.

Diffie-Hellman key exchange allows Ana and Bob want to agree on a secret key over a public channel. The two palindromic friends agree on a prime number p and a base a. This information can be exchanged publically. Ana choses now a secret number x and sends $X = a^x$ modulo p to Bob over the channel. Bob choses a secret number y and sends $Y = a^y$ modulo p to Ana. Ana can compute Y^x and Bob can compute X^y but both are equal to a^{xy}. This number is their common secret. The key point is that eves dropper Eve, can not compute this number. The only information available to Eve are X and Y, as well as the base a and p. Eve knows that $X = a^x$ but can not determine x. The key difficulty in this code is the **discrete log problem**: getting x from a^x modulo p is believed to be difficult for large p.
The **Rivest-Shamir-Adleman public key system** uses a **RSA public key** (n, a) with an integer $n = pq$ and $a < (p-1)(q-1)$, where p, q are prime. Also here, n and a are public. Only the factorization of n is kept secret. Ana publishes this pair. Bob who wants to email Ana a message x, sends her $y = x^a \bmod n$. Ana, who has computed b with $ab = 1 \bmod (p-1)(q-1)$ can read the secrete email y because $y^b = x^{ab} = x^{(p-1)(q-1)} = x \bmod n$. But Eve, has no chance because the only thing Eve knows is y and (n, a). It is believed that without the **factorization** of n, it is not possible to determine x. The message has been transmitted securely. The core difficulty is that **taking roots** in the ring $Z_n = \{0, \ldots, n-1\}$ is difficult without knowing the factorization of n. With a factorization, we can quickly take arbitrary roots. If we can take square roots, then we can also factor: assume we have a product $n = pq$ and we know how to take square roots of 1. If x solves $x^2 = 1 \bmod n$ and x is different from 1, then $x^2 - 1 = (x-1)(x+1)$ is zero modulo n. This means that p divides $(x-1)$ or $(x+1)$. To find a factor, we can take the greatest common divisor of $n, x-1$. Take $n = 77$ for example. We are given the root 34 of 1. ($34^2 = 1156$ has reminder 1 when divided by 34). The greatest common divisor of $34-1$ and 77 is 11 is a factor of 77. Similarly, the greatest common divisor of $34+1$ and 77 is 7 divides 77. Finding roots modulo a composite number and factoring the number is equally difficult.

Cipher	Used for	Difficulty	Attack
Cesar	transmitting messages	many permutations	Statistics
Viginere	transmitting messages	many permutations	Statistics
Enigma	transmitting messages	no frequency analysis	Plain text
Diffie-Helleman	agreeing on secret key	discrete log mod p	Unsafe primes
RSA	electronic commerce	factoring integers	Factoring

The simplest **error correcting code** uses 3 copies of the same information. A single error can be corrected. With 3 watches for example, you know the time even if one of the watches fails. Cockpits of airplanes have three copies important instruments. But this basic error correcting code is not efficient. It can correct single errors by tripling the size. Its efficiency is 33 percent. A cheap way to make it more efficient is to compress the data first and then make three copies. **Data compression** is a topic by itself. Here is a simple example, the **dictionary compression**. Take dictionary with $65'536 = 2^{16}$ words for example. Every word can be encode by two bytes. Assuming an average word length of 6, we can encode every word with 2 bytes instead of 6. There are better error correcting codes using linear algebra or algebraic geometry.

Lecture 12: Dynamical systems

Dynamical systems theory is the science of time evolution. If time is **continuous** the evolution is defined by a **differential equation** $\dot{x} = f(x)$. If time is **discrete** then we look at the **iteration of a map** $x \to T(x)$.

The goal is to **predict the future** of the system when the present state is known. A **differential equation** is an equation of the form $d/dtx(t) = f(x(t))$, where the unknown quantity is a path $x(t)$ in some "phase space". We know the **velocity** $d/dtx(t) = \dot{x}(t)$ at all times and the initial configuration $x(0)$), we can to compute the **trajectory** $x(t)$. What happens at a future time? Does $x(t)$ stay in a bounded region or escape to infinity? Which areas of the phase space are visited and how often? Can we reach a certain part of the space when starting at a given point and if yes, when. An example of such a question is to predict, whether an asteroid located at a specific location will hit the earth or not. An other example is to predict the weather of the next week.

An examples of a dynamical systems in one dimension is the differential equation

$$x'(t) = x(t)(2 - x(t)), x(0) = 1$$

It is called the **logistic system** and describes population growth. This system has the solution $x(t) = 2e^t/(1 + e^{2t})$ as you can see by computing the left and right hand side.

A **map** is a rule which assigns to a quantity $x(t)$ a new quantity $x(t + 1) = T(x(t))$. The state $x(t)$ of the system determines the situation $x(t + 1)$ at time $t + 1$. An example is is the **Ulam map** $T(x) = 4x(1 - x)$ on the interval $[0, 1]$. This is an example, where we have no idea what happens after a few hundred iterates even if we would know the initial position with the accuracy of the Planck scale.

Dynamical system theory has applications all fields of mathematics. It can be used to find roots of equations like for

$$T(x) = x - f(x)/f'(x) \ .$$

A system of number theoretical nature is the **Collatz map**

$$T(x) = \frac{x}{2} \text{ (evenx)}, 3x + 1 \text{ else} \ .$$

A system of geometric nature is the **Pedal map** which assigns to a triangle the pedal triangle.

About 100 years ago, **Henry Poincaré** was able to deal with **chaos** of low dimensional systems. While **statistical mechanics** had formalized the evolution of large systems with probabilistic methods already, the new insight was that simple systems like a **three body problem** or a **billiard map** can produce very complicated motion. It was Poincaré who saw that even for such low dimensional and completely deterministic systems, random motion can emerge. While physisists have dealt with chaos earlier by assuming it or artificially feeding it into equations like the **Boltzmann equation**, the occurrence of stochastic motion in geodesic flows or billiards or restricted three body problems was a surprise. These findings needed half a century to sink in and only with the emergence of computers in the 1960ies, the awakening happened. Icons like

wing of a butterfly can produce a tornado in Texas in a few weeks. The reason for this statement is that the complicated equations to simulate the weather reduce under extreme simplifications and truncations to a simple differential equation $\dot{x} = \sigma(y-x), \dot{y} = rx - y - xz, \dot{z} = xy - bz$, the **Lorenz system**. For $\sigma = 10, r = 28, b = 8/3$, Ed Lorenz discovered in 1963 an interesting long time behavior and an aperiodic "attractor". Ruelle-Takens called it a **strange attractor**. It is a **great moment** in mathematics to realize that attractors of simple systems can become fractals on which the motion is chaotic. It suggests that such behavior is abundant. What is chaos? If a dynamical system shows **sensitive dependence on initial conditions**, we talk about **chaos**. We will experiment with the two maps $T(x) = 4x(1-x)$ and $S(x) = 4x - 4x^2$ which starting with the same initial conditions will produce different outcomes after a couple of iterations.
The sensitive dependence on initial conditions is measured by how fast the derivative dT^n of the n'th iterate grows. The exponential growth rate γ is called the **Lyapunov exponent**. A small error of the size h will be amplified to $he^{\gamma n}$ after n iterates. In the case of the Logistic map with $c = 4$, the Lyapunov exponent is $\log(2)$ and an error of 10^{-16} is amplified to $2^n \cdot 10^{-16}$. For time $n = 53$ already the error is of the order 1. This explains the above experiment with the different maps. The maps $T(x)$ and $S(x)$ round differently on the level 10^{-16}. After 53 iterations, these initial fluctuation errors have grown to a macroscopic size.
Here is a famous open problem which has resisted many attempts to solve it: Show that the map $T(x,y) = (c\sin(2\pi x) + 2x - y, x)$ with $T^n(x,y) = (f_n(x,y), g_n(x,y))$ has sensitive dependence on initial conditions on a set of positive area. More precisely, verify that for $c > 2$ and all n $\frac{1}{n}\int_0^1 \int_0^1 \log|\partial_x f_n(x,y)|\, dxdy \geq \log(\frac{c}{2})$. The left hand side converges to the average of the Lyapunov exponents which is in this case also the **entropy** of the map. For some systems, one can compute the entropy. The logistic map with $c = 4$ for example, which is also called the **Ulam map**, has entropy $\log(2)$. The **cat map**

$$T(x,y) = (2x + y, x + y) \bmod 1$$

has entropy $\log|(\sqrt{5}+3)/2|$. This is the logarithm of the larger eigenvalue of the matrix.

While questions about simple maps look artificial at first, the mechanisms prevail in other systems: in astronomy, when studying planetary motion or electrons in the van Allen belt, in mechanics when studying coupled penduli or nonlinear oscillators, in fluid dynamics when studying vortex motion or turbulence, in geometry, when studying the evolution of light on a surface, the change of weather or tsunamis in the ocean. Dynamical systems theory started historically with the problem to understand the **motion of planets**. Newton realized that this is governed by a differential equation, the **n-body problem**

$$x_j''(t) = \sum_{i=1}^{n} \frac{c_{ij}(x_i - x_j)}{|x_i - x_j|^3} \ ,$$

where c_{ij} depends on the masses and the gravitational constant. If one body is the sun and no interaction of the planets is assumed and using the common center of gravity as the origin, this reduces to the **Kepler problem** $x''(t) = -Cx/|x|^3$, where planets move on **ellipses**, the radius vector sweeps equal area in each time and the period squared is proportional to the semi-major axes cubed. A great moment in astronomy was when Kepler derived these laws empirically. An other great moment in mathematics is Newton's theoretically derivation from the differential equations.

Lecture 13: Computing

Computing deals with algorithms and the praxis of programming. While the subject intersects with computer science, information technology, the theory is by nature very mathematical. But there are new aspects: computers have opened the field of **experimental mathematics** and serve now as the **laboratory** for new mathematics. Computers are not only able to **simulate** more and more of our physical world, they allow us to **explore** new worlds.
A mathematician pioneering new grounds with computer experiments does similar work than an experimental physicist. Computers have smeared the boundaries between physics and mathematics. According to Borwein and Bailey, experimental mathematics consists of:

Gain insight and intuition.	Explore possible new results
Find patterns and relations	Suggest approaches for proofs
Display mathematical principles	Automate lengthy hand derivations
Test and falsify conjectures	Confirm already existing proofs

When using computers to prove things, reading and verifying the computer program is part of the proof. If Goldbach's conjecture would be known to be true for all $n > 10^{18}$, the conjecture should be accepted because numerical verifications have been done until $2 \cdot 10^{18}$ until today. The first famous theorem proven with the help of a computer was the "4 color theorem" in 1976.
Here are some pointers in the history of computing:

2700BC	Sumerian Abacus	1935	Zuse 1 programmable	1973	Windowed OS
200BC	Chinese Abacus	1941	Zuse 3	1975	Altair 8800
150BC	Astrolabe	1943	Harvard Mark I	1976	Cray I
125BC	Antikythera	1944	Colossus	1977	Apple II
1300	Modern Abacus	1946	ENIAC	1981	Windows I
1400	Yupana	1947	Transistor	1983	IBM PC
1600	Slide rule	1948	Curta Gear Calculator	1984	Macintosh
1623	Schickard computer	1952	IBM 701	1985	Atari
1642	Pascal Calculator	1958	Integrated circuit	1988	Next
1672	Leibniz multiplier	1969	Arpanet	1989	HTTP
1801	Punch cards	1971	Microchip	1993	Webbrowser, PDA
1822	Difference Engine	1972	Email	1998	Google
1876	Mechanical integrator	1972	HP-35 calculator	2007	iPhone

We live in a time where technology explodes exponentially.**Moore's law** from 1965 predicted that semiconductor technology doubles in capacity and overall performance every 2 years. This has happened since. Some futurologists like Ray Kurzweil conclude from this technological singularity in which artificial intelligence might take over. Lets move to safer ground and discuss an important concept of computing, the question how to decide whether a computation is "easy" or "hard". In 1937, **Alan Turing** introduced the idea of a **Turing machine**, a theoretical model of a computer which allows to quantify complexity. It has finitely many states $S = \{s_1, ..., s_n, h\}$ and works on an tape of $0-1$ sequences. The state h is the "halt" state. If it is reached, the machine stops. The machine has rules which tells what it does if it is in state s and reads a letter a. Depending on s and a, it writes 1 or 0 or moves the tape to the left or right and moves into a new state. Turing showed that anything we know to compute today can be computed with Turing machines. For any

known machine, there is a polynomial p so that a computation done in k steps with that computer can be done in $p(k)$ steps on a Turing machine. What can actually be computed? Church's thesis of 1934 states that everything which can be computed can be computed with Turing machines. Similarly as in mathematics itself, there are limitations of computing. Turing's setup allowed him to enumerate all possible Turing machine and use them as input of an other machine. Denote by TM the set of all pairs (T, x), where T is a Turing machine and x is a finite input. Let $H \subset TM$ denote the set of Turing machines (T, x) which halt with the tape x as input. Turing looked at the decision problem: is there a machine which decides whether a given machine (T, x) is in H or not. An ingenious Diagonal argument of Turing shows that the answer is "no". [Proof: assume there is a machine $HALT$ which returns from the input (T, x) the output HALT(T, x) = true, if T halts with the input x and otherwise returns HALT(T, x) = false. Turing constructs a Turing machine DIAGONAL, which does the following:

1) Read x. 2) Define Stop=HALT(x,x) 3) While Stop=True repeat Stop:=True; 4) Stop

Now, DIAGONAL is either in H or not. If DIAGONAL is in H, then the variable Stop is true which means that the machine DIAGONAL runs for ever and DIAGONAL is not in H. But if DIAGONAL is not in H, then the variable Stop is false which means that the loop 3) is never entered and the machine stops. The machine is in H.]

Lets go back to the problem of distinguishing "easy" and "hard" problems: One calls **P** the class of decision problems that are solvable in polynomial time and **NP** the class of decision problems which can efficiently be tested if the solution is given. These categories do not depend on the computing model used. The question "N=NP?" is the most important open problem in theoretical computer science. It is one of the seven **millenium problems** and it is widely believed that $P \neq NP$. If a problem is such that every other NP problem can be reduced to it, it is called **NP-complete**. Popular games like Minesweeper or Tetris are NP-complete. If $P \neq NP$, then there is no efficient algorithm to beat the game. The intersection of NP-hard and NP is the class of NP-complete problems. An example of an NP-complete problem is the **balanced number partitioning problem**: given n positive integers, divide them into two subsets A, B, so that the sum in A and the sum in B are as close as possible. A first shot: chose the largest remaining number and distribute it to alternatively to the two sets.

We all feel that it is harder to **find a solution to a problem** rather than to **verify a solution**. If $N \neq NP$ there are one way functions, functions which are easy to compute but hard to verify. For some important problems, we do not even know whether they are in NP. Here are two examples: 1) **the integer factoring problem**: given n find the factors 2) **the merit factor problem**: minimize $\sum_{k=-n}^{n} c_k^2$, where $c_k = \sum_{j=0}^{n-k} a_j a_{j+k}$ An efficient algorithm for the first one would have enormous consequences for ou modern lives.

Finally, lets look at some mathematical problems in artificial intelligence AI:

problem solving	playing games like chess, performing algorithms, solving puzzles
pattern matching	speech, music, image, face, handwriting, plagiarism detection, spam
reconstruction	tomography, city reconstruction, body scanning
research	computer assisted proofs, discovering theorems, verifying proofs
data mining	knowledge acquisition, knowledge organization, learning
translation	language translation, porting applications to programming languages
creativity	writing poems, jokes, novels, music pieces, painting, sculpture
simulation	physics engines, evolution of bots, game development, aircraft design
inverse problems	earth quake location, oil depository, tomography
prediction	weather prediction, climate change, warming, epidemics, supplies

We had started with basic human activities defining mathematical fields, we end the course with mathematical activities defining some aspects of computing. Our journey through math is over.